UMTRI-2025-4 April 2025

The Little Book of Profiling
Second Edition

Basic Information about Measuring and Interpreting Road Profiles

Steven M. Karamihas, Ph.D.

Michael W. Sayers. Ph.D.

Technical Report Documentation Page

1. Report No. UMTRI-2025-4	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle The Little Book of Profiling. Second Edition.		5. Report Date April 2025 6. Performing Organization Code
7. Author(s) Karamihas, Steven M. (ORCID: 0000-0002-7480-3952) and Sayers, Michael W.		8. Performing Organization Report No.
9. Performing Organization Name and Address University of Michigan Transportation Research Institute 2901 Baxter Rd. Ann Arbor MI 48109		10. Work Unit No. (TRAIS) 11. Contract or Grant No. SD2021-09
12. Sponsoring Agency Name and Address South Dakota Department of Transportation Office of Research 700 East Broadway Avenue Pierre, SD 57501-2586		13. Type of Report and Period Covered January 2022-April 2025 14. Sponsoring Agency Code
15. Supplementary Notes The Contracting Officer's Technical Representative (COTR) was Thad Bauer. The work was supported by Pooled Fund Study TPF-5(354), Improving the Quality of Highway Profile Measurement.		

16. Abstract

The Little Book of Profiling was created in 1995 as an educational resource on the principles of longitudinal road profile measurement and interpretation for highway agency staff, highway researchers, profiling equipment manufacturers, highway construction contractors, and consultants.

The Second Edition addresses changes to the state of practice since the 1990s, which include: increased application of road profilers for construction quality assurance, increased use of road profilers to measure urban and low-speed roadways, improvements to profile measurement technology, better understandings of measurement errors, and the adoption of new measurement and analysis standards.

17. Key Word Road profile, longitudinal profile, International Roughness Index, Ride Number, inertial profilers, power spectral density, profile measurement and interpretation, cross correlation	18. Distribution Statement No restrictions.

19. Security Classification (of this report) Unclassified	20. Security Classification (of this page) Unclassified	21. No. of Pages 109	22. Price N/A

Table of Contents

List of Abbreviations

AASHO	American Association of State Highway Officials
AASHTO	American Association of State Highway and Transportation Officials
ASTM	American Society of Testing and Materials
BPR	Bureau of Public Roads
DMI	Distance Measurement Instrument
DOT	Department of Transportation
FHWA	Federal Highway Administration
FIR	Finite Impulse Response
GM	General Motors
GPS	Global Positioning System
HPMS	Highway Performance Monitoring System
HRI	Half-car Roughness Index
IRI	International Roughness Index
ISO	International Organization for Standardization
LTPP	Long-Term Pavement Performance
MPR	Mean Panel Rating
MRI	Mean Roughness Index
NCHRP	National Cooperative Highway Research Program
PSD	Power Spectral Density
PSI	Present Serviceability Index
PSR	Present Serviceability Rating
RMS	Root Mean Square
RN	Ride Number
RPUG	Road Profiler Users' Group
RTRRMS	Response-Type Road Roughness Measuring System
TPM	Transportation Performance Management
UMTRI	University of Michigan Transportation Research Institute

Acknowledgements

Mike Sayers developed the original version of *The Little Book of Profiling* with contributions from Tom Gillespie and assistance from Steve Karamihas. Steve would like to thank Mike for allowing him to be involved in the work, and for all that he learned from Mike.

The authors would like to thank the Federal Highway Administration; the National Highway Institute; and members of pooled fund study "Interpretation of Road Roughness Profile Data" for supporting the development of *The Little Book of Profiling*. The authors would also like to thank Dave Huft of the South Dakota Department of Transportation (DOT) for inspiring the work.

The authors would like to thank pooled fund study TPF5-(354) "Improving the Quality of Highway Profile Measurement" for supporting the development of the Second Edition. The authors would also like to thank the technical panel for their advice. The panel included Thad Bauer (South Dakota DOT), Frank Bell (Alabama DOT), John DeBoer (South Dakota DOT), Dan DeGraaf (Michigan Concrete Association), Dave Huft (South Dakota DOT), Colin McClenahen (Pennsylvania DOT), Steven Miles (Ohio DOT), Jason Omundson (Iowa DOT), Rohan Perera (SME, Inc.), Robert Rescot (ARA), Dulce Rufino Feldman (California DOT), Brian Schleppi (VIHSC LLC), John Senger (Illinois DOT), Amy Simpson (WSP USA), and Brett Stanton (Asphalt Pavement Association of Michigan). Thad Bauer, Frank Bell, Rohan Perera, and Brian Schleppi provided particularly detailed technical editing assistance.

The authors are grateful to Jack R. Baker (J. Raymond Baker Consulting LLC) for his professional editing of the document. The authors would also like to thank Emily N. Karamihas for helping to improve the clarity of selected sections.

Introduction

High-speed road profiling began in the 1960s when Elson Spangler and William Kelly developed the inertial profilometer at the General Motors (GM) Research Laboratory. The GM profilometer provided an efficient way to measure profiles of in-service roads for use as realistic inputs to vehicles on a driving simulator in the laboratory. Over the next two decades, several state highway agencies adopted the technology for pavement management applications. Inertial road profilers, which are based on the design of the GM profilometer, have become standard tools for measuring road roughness. Most state and federal highway agencies now use roughness as a measure of the quality of their roads. In fact, the Federal Highway Administration (FHWA) requires states to submit roughness data for a large portion of their federal aid highway network.

The widespread adoption of inertial profilers created demand for educational resources and technical information among the road-monitoring community. In response, road profiler users from different states met in 1989 and formed the Road Profiler Users' Group (RPUG). RPUG has met nearly every year since 1989 to provide a forum for discussing issues involving the measurement and interpretation of road profiles. From 1992 to 1995, the authors conducted a research project at The University of Michigan Transportation Research Institute (UMTRI) called "Interpretation of Road Roughness Profile Data." The research was funded by the FHWA and pooled state funds. At the 1994 RPUG meeting, representatives of the participating states suggested that a short course was needed to provide the necessary education on profiling. Dave Huft, inventor of the South Dakota Profilometer, proposed a "Little Golden Book of Profiling" in the spirit of the popular children's books and modeled after his educational presentations at early RPUG meetings.

The National Highway Institute of the FHWA supported the development of a short course called "Measuring and Interpreting Road Profiles." The first session was held in November 1995 in Ann Arbor, Michigan. *The Little Book of Profiling* was written for the course. The book was then revised and extended for additional sessions of the course held in conjunction with the 1996 RPUG meeting in Denver, Colorado and the 1997 RPUG meeting held in Overland Park, Kansas. The book was finalized in 1998.

Much of the material in the 1998 version of the *Little Book of Profiling* remains pertinent. However, an update was needed to address changes to the state of practice since the 1990s, which include: increased application of road profilers for construction quality assurance, increased use of road profilers to measure urban and low-speed roadways, improvements to profile measurement technology, better understandings of measurement errors, and the adoption of new measurement and analysis standards. In 2022, FHWA Pooled Fund Study TPF-5(354) supported an update to *The Little Book of Profiling* to address these changes.

The Little Book of Profiling exists because road profiler users recognized the need for background information and technical training to maximize the value of the measurements. Though a large body of knowledge exists pertaining to road profile measurement and interpretation, it is spread out over many sources. *The Little Book of Profiling* consolidates important background information about road profile measurement and interpretation into a single source. *The Little Book of Profiling* addresses three basic questions:

- How do profilers work?

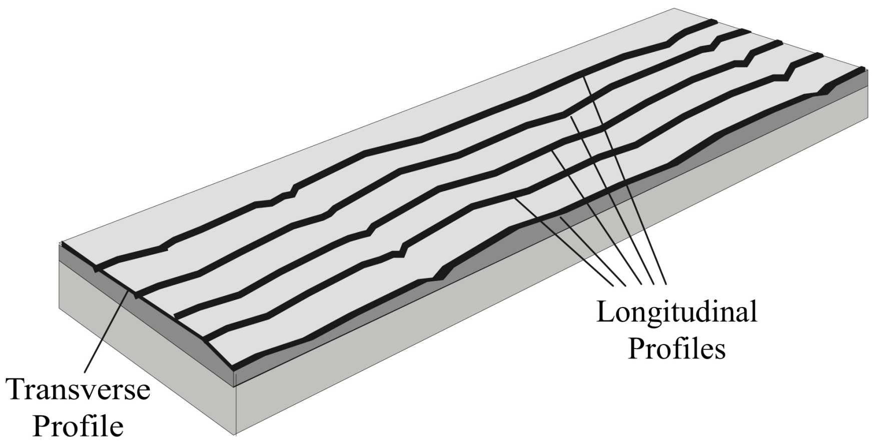
- What can be done with their measurements?
- What can be done to reduce measurement errors?

In this updated edition of *The Little Book of Profiling*, we have maintained the user-centric focus of the text. The material is meant to be no more technical than is necessary to address the three basic questions that define the scope.

What Is a Profile?

A profile is a two-dimensional slice of the road surface, taken along an imaginary line.

Transverse profiles, which are taken along a lateral line, include information about cross slope, superelevation, rutting, and other pavement distress. Longitudinal profiles include information about grade, roughness, and texture. This book addresses longitudinal profiles.



A profile of a road, or any other traveled surface, can be measured along any continuous line. It is possible to measure the profile along a curved line. Longitudinal profiles are typically measured along a line that is parallel to the road centerline or some other reference that follows the same path as the roadway.

The term “longitudinal profile” is really shorthand for “longitudinal roadway elevation profile,” since many things besides elevation can be expressed as a function of distance along a line or path of interest. In this book, the term “longitudinal roadway elevation profile” is abbreviated simply as “profile.” When the profile of something other than elevation is discussed, such as roughness, a modifier is added to the name.

Many profiles can be measured on the same road, each along a different line.

Often, longitudinal profile is measured along two lines per lane, one in each wheel track. In some cases, longitudinal profile is measured along several parallel paths at once, or as the byproduct of a three-dimensional surface measurement. If a profile measurement is repeated, the same output can only be expected if the same path is followed in each pass.

The width of the line depends on the measurement instrument.

Many road profilers detect the road surface using non-contacting sensors, such as lasers. Profiles measured with some non-contacting sensors cover a slice of the road that is very narrow (e.g., 0.04 inches wide), while others cover a slice that is several inches wide. Still other profilers detect the road surface by contacting it using supporting wheels or footpads. The footprint of a profiler is the area of the road surface detected by its sensors. The content within a profile and its relevance depends on a complex combination of the profiler's footprint, the weighting of asperities within the footprint, and sampling and processing procedures. In many cases, it is easier to repeat a measurement of roughness using profilers with a wider footprint.

For any line on the road, there is a true profile.

A true profile exists for any line along a surface. The true profile of a roadway contains a vast amount of information, from the large changes in absolute height over long distances down to minor fluctuations of height over very short distances. The requirements for measuring profile depend on the end use of the data.

For example, consider a profile measurement used to verify the vertical alignment of a newly built segment of roadway. The profile would be adequately described with elevation points taken at 25-ft intervals for the entire length of the project, with the individual measurements having a resolution of 1/8 inch. In a different application, consider a computer analysis used to characterize the macrotexture of a pavement based on profile. A typical analysis requires profile points spaced 0.04 inches apart over 3.28 ft, with a resolution of 0.002 inches. Each set of readings describes a part of the true profile. In this book, we are primarily interested in the content we must capture within the true profile to evaluate the roughness of the roadway.

What Is a Profiler?

A profiler is an instrument used in conjunction with appropriate test methods to produce a sequence of numbers related to the true profile for an imaginary line on the road.

A profiler produces a sequence of numbers related in a well-defined way to a true profile.

The numbers obtained from typical profilers are not necessarily equal to true elevation. A profiler does not always measure true profile, exactly. It measures the components of true profile that are needed for a specific purpose. However, the relationship between the true profile and the numbers produced by a profiler must meet specifications linked to the end use of the measurement. This book addresses specifications related to roughness measurement.

A profiler works by combining three ingredients.

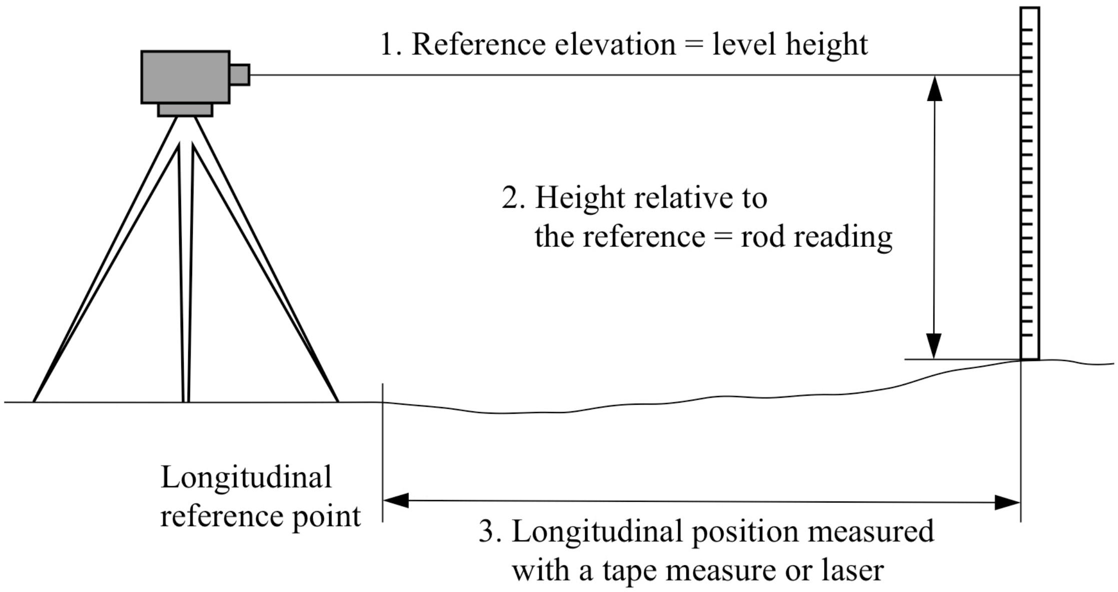
They are:

1. a reference elevation
2. a height relative to the reference
3. longitudinal distance

The three ingredients are combined in different ways, depending on the design of the profiler. Let's examine three profiler types.

Rod and Level

The rod and level are familiar surveying tools. The level provides the reference elevation, the readings from the rod provide the height relative to the reference, and a tape measure or a laser provides the longitudinal position associated with each elevation measurement.



The rod-and-level method is called static because the instruments are not moving when the readings are taken.

Although the rod and level are familiar to many engineers, the requirements for obtaining a profile that is valid for computing roughness are much different than for laying out a road. To measure roughness, one must take elevation measurements spaced only a few inches apart, and the individual height measurements must be accurate to 0.02 inches or less. These requirements are much more stringent than is normal for surveying and get even more stringent on very smooth roads. However, the absolute height of the instrument is not of interest when measuring a profile for roughness, even though it is normally very important when using rod and level for other applications.

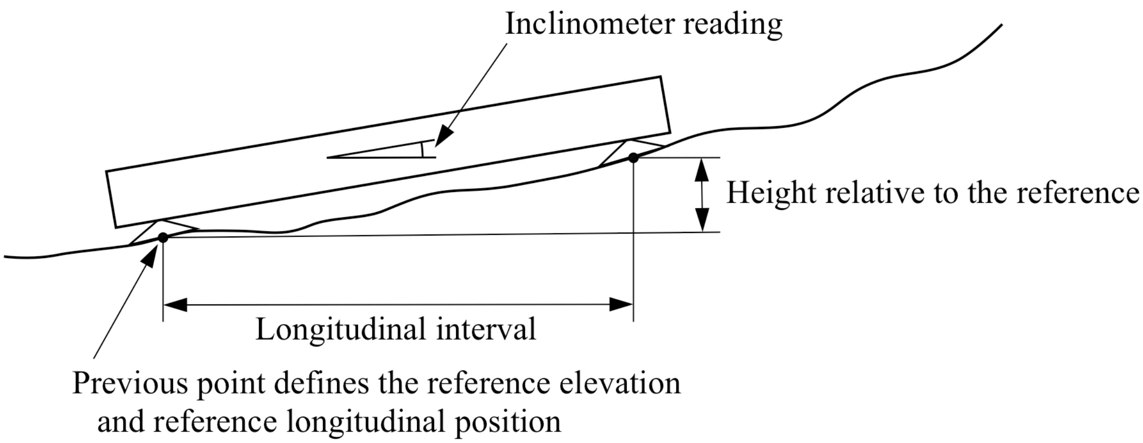
Inclinometer-Based Profilers

Inclinometer-based profilers measure profiles suited for roughness analysis more rapidly than the rod and level. They typically include a battery-powered, on-board computer for automatically recording data from the sensors and performing the calculations needed to produce a profile.

Classical inclinometer-based profilers collect readings statically.

Static inclinometer-based profilers contain a precision inclinometer that measures the slope between two supports, which are usually about a foot apart. To progress along the path being profiled, the device is either pivoted around the leading support or repositioned mechanically so that the trailing support for the current reading rests in the position of the leading support from the previous reading. Once the device detects that the sensor has stabilized, it records a reading and prompts the operator to move on to the next position.

With this design, the reference elevation at each position is the elevation value at the trailing support, which is calculated from the previous reading. The height relative to the reference at each resting position is based on the inclination of the device relative to gravity, together with the spacing between its supports. The spacing between supports is the longitudinal interval for the measured profile, and the longitudinal distance is determined by multiplying the number of readings with the spacing between the supports.



Modern inclinometer-based profilers roll over the path being profiled.

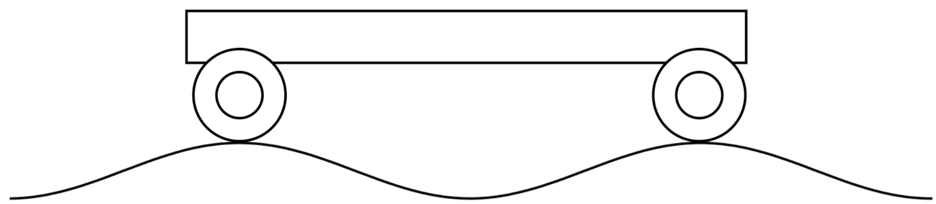
Rolling inclinometer-based profilers are supported by wheels, which are often approximately 10 inches apart, and are pushed along the path by the operator. Rolling devices typically collect inclination readings and calculate elevation at a longitudinal interval much shorter than the spacing between the supporting wheels.

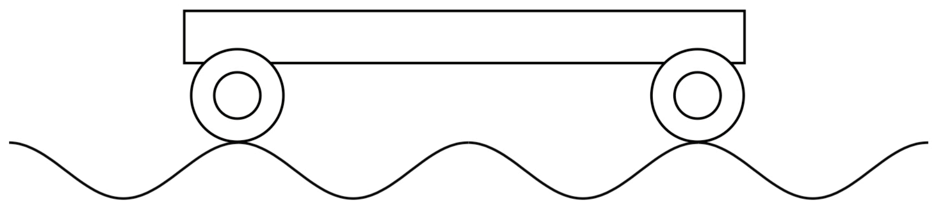
Rolling inclinometer-based devices trigger new readings at the desired spacing by monitoring wheel rotation. Like the static method, each new elevation value in the profile uses the previous elevation as a reference. The height relative to the reference depends on the longitudinal sample spacing and the inclination reading at the position where the system is triggered.

Inclinometers measure slope incorrectly when they are accelerating because they interpret the acceleration as if it were part of Earth's gravity. It is important to roll inclinometer-based devices forward at a constant speed. Excessive horizontal acceleration or rapid changes in tilt will contaminate the readings used to calculate the profile.

Inclinometer-based devices reduce or miss short-term fluctuations from the true profile.

Readings from inclinometer-based devices capture the average slope between the supports. As a result, inclinometer-based devices do not fully capture the influence of features in the profile that fluctuate over distances close to, or shorter than, the support spacing. For example, consider profile features that have their highest and lowest points the same distance apart as the device's support spacing. These may barely register or not register and all, because the body of the device stays level as it passes over them.





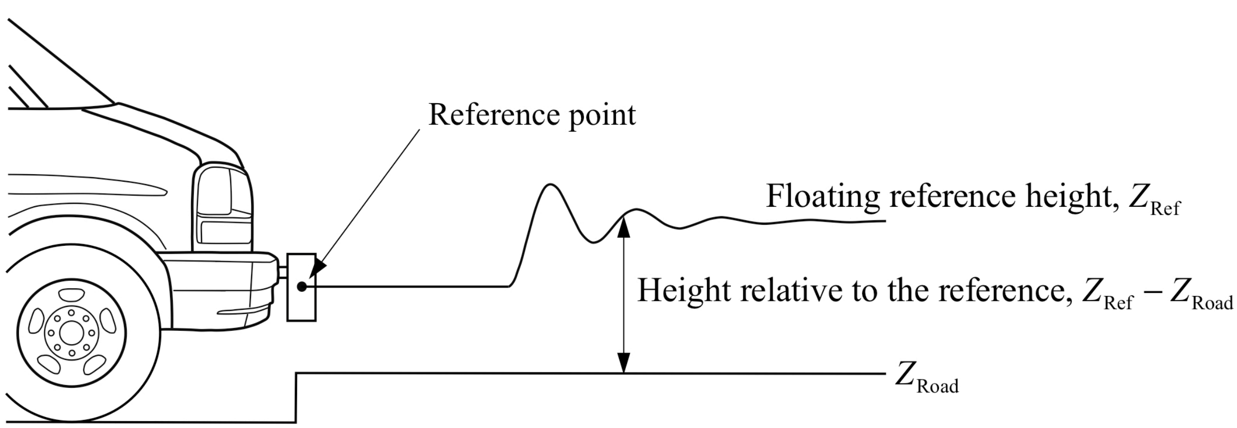
Some inclinometer-based profilers use additional sensors to help fill in the gaps.

Inertial Profiler

The development of the inertial profiler by the GM Research Laboratory in the 1960s made high-speed profile measurement possible. The efficiency of inertial profilers relative to other profile-measurement techniques and their ability to capture components of the true profile linked to roughness led to widespread use of inertial profilers for pavement management. Inertial profilers combine the same three ingredients as both the static rod and level and inclinometer-based profilers.

Inertial profilers get reference elevation from an accelerometer.

A vertically oriented accelerometer, mounted to a reference point on the profiler above the height sensor, provides a record of its motion. Data processing algorithms convert the record of acceleration to a record of height of the reference point. Since the body of the profiler host vehicle pitches, bounces, and rolls in response to the road profile, the height of the reference point fluctuates. The record of height is called a floating reference height or, more formally, an accelerometer-established inertial reference.



A height sensor provides height of the road surface relative to the floating reference.

A height sensor measures the vertical distance to the road surface directly underneath the accelerometer. This height is usually measured with a non-contacting sensor, such as a laser.

The schematic above shows fluctuation of the floating reference height starting at the location where the front wheels of the profiler host vehicle encounter the upward step in the profile. Simultaneously, the height sensor measures the fluctuating difference between the floating reference height and height of the road surface. Combining the two readings recovers the height of the road surface:

$$Z_{\text{Road}} = Z_{\text{Ref}} - (Z_{\text{Ref}} - Z_{\text{Road}})$$

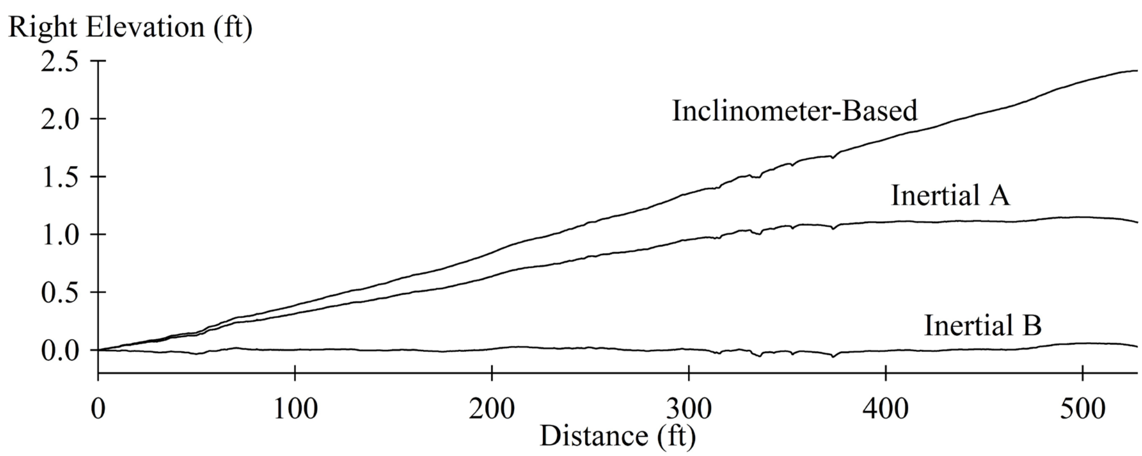
An inertial profiler must be moving to function.

Inertial profilers often measure longitudinal distance by monitoring the rotation of one or more of the host vehicle wheels. The most common sensor used for this is a rotational encoder. Some profilers use readings from the speedometer or the transmission instead; others incorporate Global Positioning System (GPS) sensors; and still others use sensors that interpret incremental changes in surface images.

Typical inertial profilers must travel at a minimum speed to function properly, because electrical noise and other errors will overwhelm the legitimate components of the sensor readings at low speeds. The minimum speed needed for accurate measurements depends on the road surface alignment, the host vehicle, the sensor specifications, and the algorithms used to calculate the profile from the sensor readings. The maximum speed is limited by safety and the need to position the profiler over the intended wheel track. Profiler manufacturers usually recommend a speed range.

Profiles from different devices may look very different.

The inertial reference from a profiler is useful, but it not as easy to visualize as the reference used in the rod and level or an inclinometer-based device. The figure below provides an example of profiles measured by three different devices on the same test section. The agreement between the profiles obtained with an inertial profiler and one obtained with an inclinometer-based profiler is good in some respects, but not in others.



The inertial profilers measured the test section at 25 mi/hr, and each profiler may not have perfectly followed the same path as the others. Still, this does not explain the completely different appearances among the profiles. The inclinometer-based profile shows a net positive grade equivalent to about 0.45 ft vertical per 100 ft horizontal. One of the inertial profilers shows an overall rise of about 1.15 ft over the length of the test section, and the other shows a profile with a trend that is relatively flat.

The inclinometer-based device measured the grade and very long undulations more accurately than the inertial profilers. However, none of these devices necessarily measures those aspects of the true profile perfectly. Plots of elevation versus distance from these three devices do not agree, even though the measurements are from the same true profile. Moreover, different plots may be obtained for repeated measures of the same true profile from the same device, just by choosing different settings before each test.

Accurate summary numbers can be obtained from profiles measured by inertial profilers.

Because the inertial profilers did not accurately measure the grade or reproduce the plot from the inclinometer-based device, one may initially think they are not useful, or that they are not sensing the true profile. Yet, even if the profile measured by an inertial profiler does not produce a plot that looks like the true profile, it may provide accurate summary numbers that are calculated from the profile. For example, the three profiles in the example plots produced a summary number called the International Roughness Index (IRI) within 0.8 percent of a painstakingly measured reference value. (This book describes the IRI in detail.) In this case, all three devices accurately measured the aspects of the true profile that contribute to the IRI.

The original design has been updated with new sensors and computers.

Early inertial profilers sensed the height of the vehicle reference point relative to the road surface using an instrumented follower wheel. The design worked, but the follower wheels were fragile and required testing at speeds low enough to avoid losing contact with the road surface. All inertial profilers that are sold today use non-contacting sensors instead of follower wheels.

Early systems performed profile calculations electronically and required that the vehicle operate at a constant speed. Modern inertial profilers correct for minor variations in speed and perform the calculations numerically with on-board computers.

Some inertial profilers are mounted on lightweight vehicles.

Many inertial profilers use highway vehicles, such as cars, vans, pickups, or light trucks as host vehicles, so they can operate within the flow of traffic on in-service roadways. These are called high-speed profilers.

Some inertial profilers that are used to measure newly placed pavement surfaces are mounted to all-terrain vehicles or golf carts. These are called lightweight profilers. Lightweight profilers use the same set of sensors as high-speed profilers and obtain profiles using the same measurement principle. Lightweight profilers are commonly used on pavement construction sites for two reasons. First, they offer a safe alternative to high-speed profilers to operate within limited space on a busy construction site, because they are much lighter and travel at a lower speed. Second, they are able to travel over concrete pavements that have not developed enough strength to support a high-speed profiler.

Other Profiler Types

The rod and level, inclinometer-based devices, and inertial profilers have all been widely used to measure profiles, but other measurement concepts are possible. Augmented versions of inclinometer-based profilers and inertial profilers have been proposed and have been shown to capture more features of the true profile or capture the profile accurately over a wider range of operating conditions. Some inertial profilers are designed for convenient movement between host vehicles. Measurement of the road surface in three dimensions is now available using an enhanced version of the inertial profiler measurement concept. Some high-speed profilers exist that do not incorporate an inertial reference at all.

What Can You Do With Profiles?

There are at least four common applications of profile measurements.

1. To monitor the functional condition of road networks for pavement management.
2. To evaluate the ride quality of newly constructed or repaired road sections.
3. To diagnose the condition of specific road sections and determine appropriate remedies.
4. To study the condition of specific road sections for research.

The technical requirements for the four categories cover quite a range. Monitoring a road network may require measurements of thousands of miles of profile per year. Some states measure the profile of more than 10,000 lane miles annually. At the other extreme, a research experiment might involve frequent measurement of the profile on a single test section, with the intent of identifying subtle changes or the onset of deterioration.

Most applications of profiles include the calculation of a roughness index.

Measuring the profile is only half of the job. Interpreting the profile is the other half. The most common way to interpret a profile is to reduce it to a roughness index. To obtain information of any type from a measured profile, there are two basic requirements:

1. The profiler must capture the relevant information from the true profile.
2. Computer software must exist to extract the desired information from the measured profile, such as a roughness index.

It is possible to process a profile several times, using different analyses to extract various kinds of information. However, it can be a challenge to select the roughness index or analysis that is the most useful. Although the technology to measure profile has existed for several decades, the research community routinely proposes new analysis options. This book describes some standard analyses that are commonly applied to profile measurements.

A profile-based analysis should be relevant.

The analysis applied to profile data should be targeted at an application. The fact that a roughness index can be obtained repeatedly and accurately does not make it useful. A relevant measure has been linked by research to a property of the road considered by engineers or the public to be important.

This book describes the IRI and Ride Number (RN) in detail. Both are linked to roughness. IRI has demonstrated strong compatibility with equipment used to develop pavement management systems and broadly represents vehicle dynamic response to road roughness. RN is linked by statistical correlation to public opinion of pavement ride quality.

What Is a Valid Profiler?

The word profile has appeared in descriptions of the output from road-measurement equipment for several decades. To some users, any device that produces a wiggly line might be called a profiler. However, in this book, we take a more restrictive view that a profiler must produce a wiggly line with an established relationship to the true profile.

A true profile contains more information than we can use.

A true profile includes a great deal of information. It tells whether the road is going up or down a hill. It gives roughness information. It has texture information. Most profiler applications do not require simultaneous measurements of large-scale geometric features, such as hills and valleys, and microscopic details about the surface texture. Doing so might require expensive sensors, create an undue data storage burden, and unnecessarily complicate the data processing. Instead, we often measure only part of the information in the true profile.

A valid profiler provides the same summary numbers that would be obtained from the true profile.

Valid profilers do not need to measure all of the information in the true profile. Instead, they must measure the information needed for the intended purpose. Of course, no measurement system is perfect, and error exists. Error levels are traded off against the cost of the system and the effort needed to use it. To be considered valid for obtaining a roughness index, a profiler must measure roughness index values that are neither high nor low, on the average, compared with the values that would be obtained from the true profiles. To do this with confidence over a wide range of conditions, valid profilers must capture the components from the true profile without distortion that affect the roughness index.

Roughness index values from two or more valid profilers are directly comparable, with no conversion required.

One true profile exists for a particular line or path over the road surface at any given time. All valid profilers are, by definition, capable of measuring profiles that produce the same roughness index values over the same path at the same time. Profiling is a rare technology that allows engineers to directly measure the same roughness index values using equipment with different measurement concepts, made with different proprietary designs, or made by competing manufacturers. Under the right conditions, there are inertial profilers and inclinometer-based profilers from several commercial manufacturers that are capable of reproducing the same IRI value for the same line on the road surface.

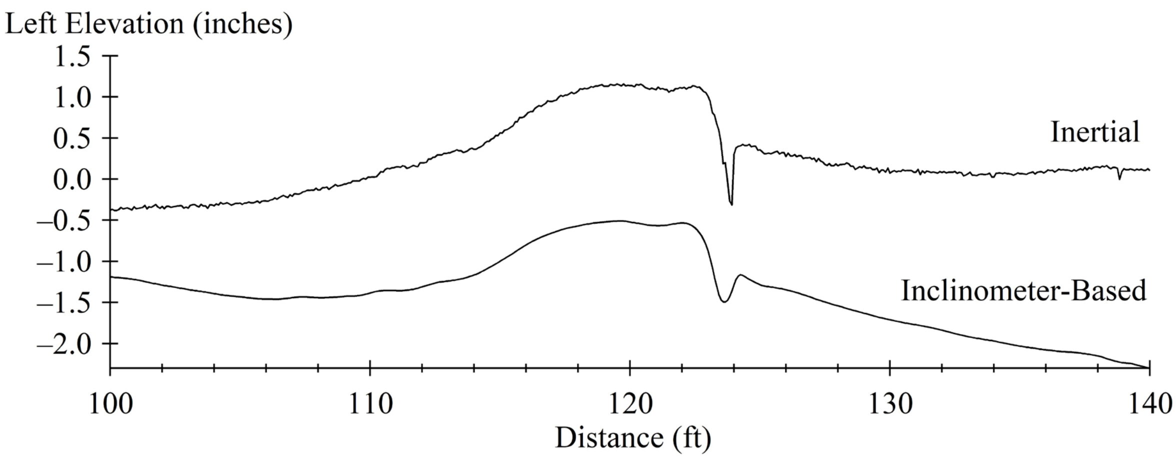
Statistics from valid profilers are stable over time.

The concept of a true profile is one that is simple and depends only on geometry. The definition will be the same in 100 years as it is today. Thus, an analysis or roughness index defined based on true profile provides information with meaning that does not change, so long as a valid profiler measured it. A requirement for effective management of a pavement network is the ability to compare data from multiple condition surveys over many years. Valid profilers offer a way to measure the roughness of roads on a stable, comparable scale over time. This was an important factor in the widespread adoption of profilers for pavement network monitoring.

No single device is the best reference for every use.

The validity of a profiler depends on its intended use. Every profiler has a limited range of applications for which it is valid. For example, an inclinometer-based profiler may be valid for determining IRI, and also for determining changes in road grade. However, it cannot sense cracks or observe profile features that are much narrower than the contact area under each supporting wheel or foot.

The figure below shows a small segment of a profile measured by an inclinometer-based device and an inertial profiler. The profile includes a transition from asphalt to concrete pavement with a wide crack at the interface. The inertial profiler's height sensor projects an area of light onto the surface that is longitudinally narrow relative to the supporting wheel used by the inclinometer-based profiler. The shape and depth of the crack is better defined in the measurement by the inertial profiler. In contrast, the inclinometer-based profiler correctly measured the grade surrounding the rough transition, whereas the inertial profiler did not.



<u>When Is a Profiler Not Valid?</u>

What if the statistics, such as values of a roughness index, measured by a profiler are systematically biased relative to an accepted reference? Say they tend to be 20 percent high in a comparison on multiple test sections. One might be tempted to define a "calibration constant" and reduce all the roughness index values measured by that profiler by 20 percent. Don't do it!

A profiler is not valid if its measures are systematically biased.

If a statistic measured by a profiler is systematically biased, then the device is not a valid profiler for that statistic. The bias indicates that the profiler is incorrectly measuring some part of the true profile needed for the statistic, and it should not be used to obtain that statistic. Although the results might differ by 20 percent over one set of test sections, there is no way to estimate the errors that may occur under other conditions, such as other types of pavements, other sources of roughness, or different operational conditions. Unless the source of the error is known, one cannot know if the error is consistent over time.

A profiler that is not valid for one statistic might still be valid for another. For example, many profilers used in the past were valid for measuring the IRI, but did not capture enough details within the profile for a valid measurement of RN.

A profiler is not valid if the random error for an individual measurement is too high.

What exactly does it mean to say the error is "too high"? That depends on the application. For monitoring the overall roughness of a large road network, some random error relative to the true value of roughness on each road segment is tolerable if it averages out over many segments. For evaluating the status of a specific road segment for construction quality assurance, the same level of random error might be unacceptable.

Valid measurement of the profile ensures valid measurement of summary numbers.

A portion of this book addresses profiler error sources and profiler verification, emphasizing the need for valid measurement of the relevant components within the true profile. Examining the profile, rather than just the roughness index values, is a more reliable way to detect and diagnose measurement errors.

What Is Signal Processing?

Civil engineers use information measured by profilers to assess the condition of road networks or of specific pavement sections. However, the technology used to measure profiles did not evolve from pavement test methods within civil engineering. Mechanical and electrical engineers invented the inertial profiler. Many of the methods used to calculate the profile from sensor readings and to extract information from the profile have been well established in electrical and mechanical engineering. These methods are not a standard part of the civil engineer's repertoire.

To understand how profiles can be used, users of profile measurements must understand terms such as analog signal, digital signal, signal processing, filter, and frequency response.

Analog signals are continuous.

Many profiler sensors produce a voltage or current that continuously varies in proportion to the quantities they are trying to measure, such as acceleration or the relative height of the road surface. The outputs of these sensors are examples of analog signals. The term analog implies that the signal is analogous to some physical quantity that fluctuates continuously. For example, a strip chart made with an ink pen (e.g., an old seismograph) is an analog representation. Analog signals are processed using combinations of circuit elements, such as resistors, capacitors, inductors, amplifiers, etc. Early inertial profilers processed and combined the sensor signals to create a voltage output, which was stored on magnetic tape. The analog output signal fluctuated continuously with a well-defined relationship to variations in the true profile.

A digital signal is a series of numbers.

Modern profilers convert analog signals to sequences of numbers that are stored in computer memory at regular intervals. Each of these sequences is a digital signal. Usually, the conversion to a digital signal includes scaling of the numbers into units that correspond to the engineering quantity they represent, such as acceleration, distance, height, or elevation.

Depending on their design, profilers may convert analog signals to digital signals at different stages of the measurement process. In the past, profilers combined analog sensor signals to create an analog signal for the profile and then converted the profile to a digital signal for storage and analysis. In most profilers, the outputs from analog sensors are captured and digitized before they are processed further and combined to calculate the profile. Some profilers use sensors that produce digital signals directly as output using their own internal circuits and processors.

Signal processing is the mathematical analysis and transformation of signals.

Signals are processed mainly for two reasons:

1. To improve the quality of a measurement by eliminating unwanted "noise" from the data.
2. To extract information of interest from the signal.

The calculation of a profile from sensor signals is a form of signal processing. The analysis of a profile to calculate a roughness index or to find the sources of roughness also requires signal processing.

What Are Sample Interval and Recording Interval?

The true profile is continuous. It is a slice of the road surface taken along an imaginary line. As such, a true profile is an analog signal that varies with distance along the imaginary line.

Nearly all profilers in use today are digital.

Nearly all profilers in use today are digital, because they measure and store profiles as a finite sequence of numbers. Static inclinometer-based devices and the rod and level produce an elevation measure with each static setup. These sequences of numbers are inherently digital: analog versions of these devices do not exist.

Inertial profilers have computers connected to the sensors. At some interval of time or distance, the computer "samples" the readings of the individual accelerometers and height sensors. For systems triggered using a clock, the sampling rate is expressed in terms of samples per second, which corresponds to a time-based sample interval in seconds. For systems that trigger digitization based on travel distance, the sampling rate is expressed in terms of samples per distance (e.g., samples/ft), which corresponds to a distance-based sample interval.

Recording interval is longitudinal distance between stored elevation values.

For a static inclinometer-based device, the recording interval is the distance between its two supporting feet. For a rod and level, the recording interval is the distance between positions where the rod is placed on the ground.

Inertial profilers and modern inclinometer-based profilers may internally compute profile samples at a very short (or fast) interval but store the profile at a recording interval longer than the sample interval. Be careful. The literature pertaining to road profile measurement and analysis often uses the terms sample interval and recording interval interchangeably.

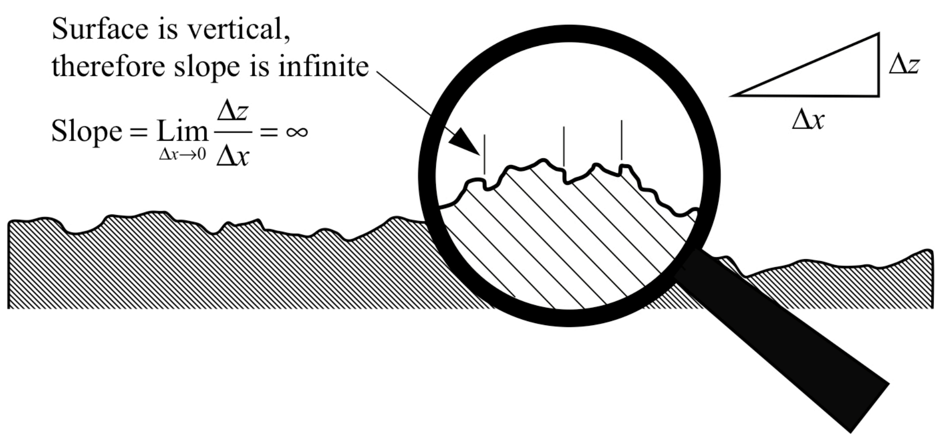
Recording interval determines the number of stored elevations per mile.

If recording interval Δx has units of feet, then a profile includes $5280/\Delta x$ recorded data values per mile. A small recording interval means that more storage is needed to record a profile. It also means that analysis of the profile will take longer, because there are more numbers to process. From the viewpoint of efficiency, we do not want to use a recording interval that is smaller than is necessary. Recording profiles with an unnecessarily short recording interval means we will have to process more numbers, allocate more computer storage, and use more network bandwidth to transfer the data than is required for the intended purpose.

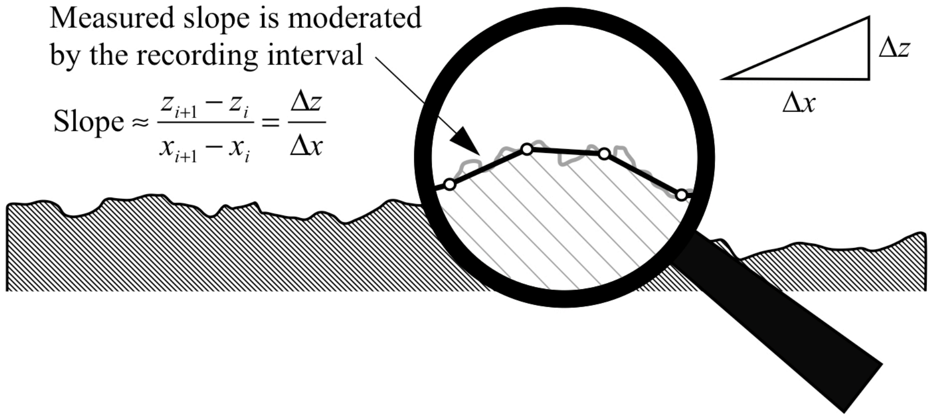
Recording interval limits the information contained in a profile.

After a measurement is made, all we know about the profile is the stored data. We have no information about the true profile between the recorded values of elevation. Ideally, the recording interval is small enough to capture the profile characteristics of interest.

Many statistics used in the past for representing road roughness have units of slope, such as inches/mile. With a profiler, we might suppose that a good measure would summarize fluctuations in “true slope.” However, there is a theoretical problem: in some locations the true slope of a real pavement is infinite. This is because, if we look closely at the texture level, we will find places where the profile is vertical.

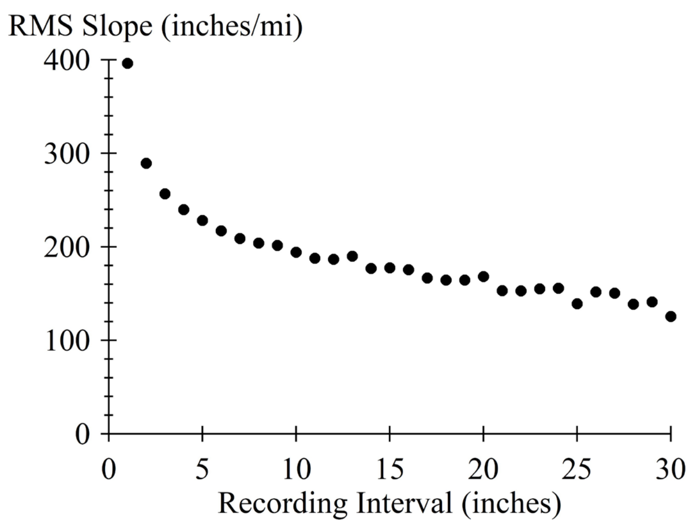


Now consider how the profile is represented with a set of sampled elevations. Instead of a continuous elevation z as a function of distance x, we have elevation values z_i at distances x_i, where i is the sample number. The slope between any two points is the rise between consecutive points divided by the longitudinal distance between consecutive points. Recall that the longitudinal distance between points is the recording interval, Δx.



For typical profiles, the root-mean-square (RMS) value of slope increases the smaller we make Δx. If users of profile measurements were to try to report "true RMS slope" as a roughness index, it would be meaningless unless the sample interval was standardized. If one profile measurement produces an RMS slope of 190 inches/mi and a different profiler on a different road measures a profile with an RMS slope of 253 inches/mi, we cannot say which road is rougher. The higher RMS slope could be due to the use of a shorter recording interval.

The next figure demonstrates the effect for the test section examined in the "What Is a Profiler?" discussion. It shows the RMS slope for a profile sampled at a very short interval and recorded at an interval of about 1 inch. The figure shows RMS slope for the same measurement at several simulated recording intervals. The larger recording intervals were obtained by throwing out points from the stored profile. The figure shows that the RMS slope changes systematically as the recording interval changes. As expected, the RMS slope diminishes with increasing recording interval, because steep slopes associated with small asperities are excluded or greatly reduced.



What Is Filtering?

Filtering is a common process to clean contaminants out of liquids such as water or oil, or to remove impurities from air. For instance, the filter on a faucet traps particles in the water, allowing the water to pass through. In electronics, components or circuits that modify a voltage continuously are called analog filters. A common application is to filter out unwanted voltage fluctuations from a power supply, providing a “clean” power source. Electronic signals are filtered to remove unwanted noise and to extract information of interest.

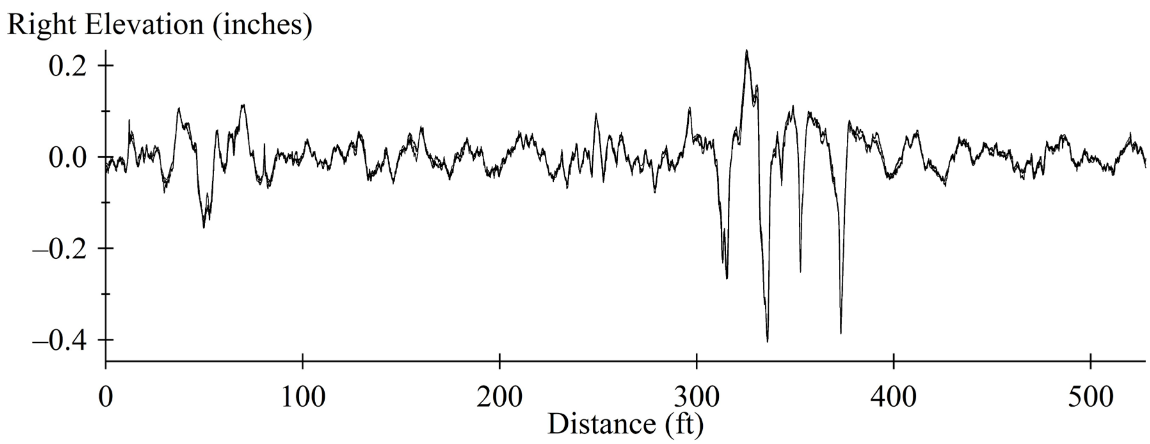
Digital filters, like their analog counterparts, can be used to remove noise or extract information of interest. However, digital filters operate on sequences of numbers, rather than continuous signals.

A digital filter is a calculation procedure that transforms a sequence of numbers (a digital signal) into a new sequence of numbers.

To make practical use of a profile measurement, it is essential to filter the sequence of numbers that makes up the profile. As a user of profile measurements, you do not have to understand the details of the transform, because the computer performs the calculations automatically. However, it is important to understand the importance of filtering, the role of filtering as a part of profile measurement and interpretation, and the behavior of the filters applied to a profile.

Filtering helps make different types of profile features more visible.

For example, consider the three measured profiles shown in the “What Is a Profiler?” section. Details of the roughness within the profile are barely visible without applying additional filtering after the profiles are measured. The figure below shows the same three profiles after they have been filtered to remove the road grade and very long undulations.



Notice the dips in the range from 300 ft to 400 ft. These are much less visible in the plot of raw profiles, because the vertical scale covers a range that is many times larger than the depth of the dips. When the grade and long undulations are removed mathematically using filtering, the vertical scale is much smaller, and the dips are easier to see. With depths of up to 0.4 inches, these dips will disturb passing vehicles and get the attention of anyone driving over them.

The raw profile plots demonstrate that different profilers will measure some components of the profile very differently; and the profile plots with filtering show that the same set of profilers can measure features of interest with the same size, shape, and placement. In the figure, the three filtered traces are barely distinguishable from each other.

Filtering is particularly important when viewing data from inertial profilers. This is because the most visible features of the unfiltered measurement—the underlying grade and overall curvature—are the least accurate parts of the data.

Filtering profiles is a fundamental part of the measurement process.

Every inertial profiler has at least one filter built into it. Typically, profilers filter the sensor signals to remove very rapid fluctuations in the readings associated with electrical noise. Filtering is used to convert the data originating from the accelerometer and the height sensor into the same units. Additional filtering is added to prevent slowly fluctuating sensor errors from causing a large drift in the calculated profile.

Some common analyses involve multiple filters—the output from one filter becomes the input to the next. Conceptually, this is just like putting several electrical filters (circuits) together or putting several water filters (wire meshes) on your faucet.

There is not a single filter that is used for all applications of profile measurement.

A filter is just a name for a mathematical transform that modifies a sequence of numbers. There are an unlimited number of filters that can be imagined and programmed. Several standard filters exist, and some of them are routinely applied to road profiles. The pages that follow include descriptions of some of them.

What Is a Moving Average?

A moving average is a simple filter used often in profile analysis, particularly when creating graphical views of profiles. For example, the plots shown in the previous section were processed with a moving average filter. Many other filter types exist. The moving average filter is covered here because it is more intuitive than most filters and examining the way it works helps demonstrate what filters do.

A moving average filter replaces each point in a profile with the average of several adjacent points.

When using an odd number of points, a moving average transforms a profile (p) using the following summation:

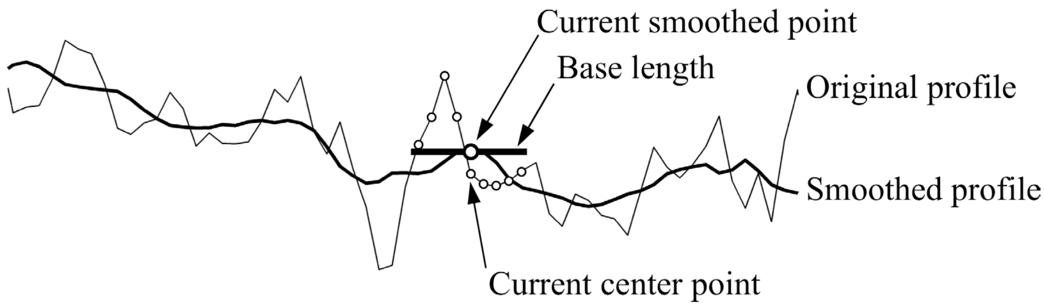
$$p_{LP}(i) = \frac{1}{2M+1} \sum_{j=i-M}^{j=i+M} p(j)$$

Where p_{LP} is the smoothed profile (also called a low-pass filtered profile), i is the point number, and the number of points in the average (N) is equal to $2M+1$. The base length (B) for the moving average is the number of points times the recording interval:

$$B = N\Delta x$$

Usually, the analyst selects a desired base length, which determines the number of points used in the summation. (The formula for a moving average can be adapted to accommodate an even number of points.) The actual base length can only be an integer multiple of the recording interval.

The smoothing effect of a moving average filter is demonstrated in the close-up view of a profile below. The figure shows the profile before and after a nine-point moving average. Each point is replaced by the average of nine points that surround it. The base length is nine times the recording interval. The moving average filter removes fluctuations shorter than the base length and retains trends that change over a distance longer than the base length.



A moving average filter can also be used to remove the smoothed profile.

In many cases, we are not interested in looking at a highly smoothed profile, which simply shows whether the road is going up, down, or staying level. Instead, we are interested in the deviations from the grade and long trends that degrade vehicle ride quality and annoy the traveling public.

An anti-smoothing version of the moving average filter subtracts the smoothed profile from the original profile:

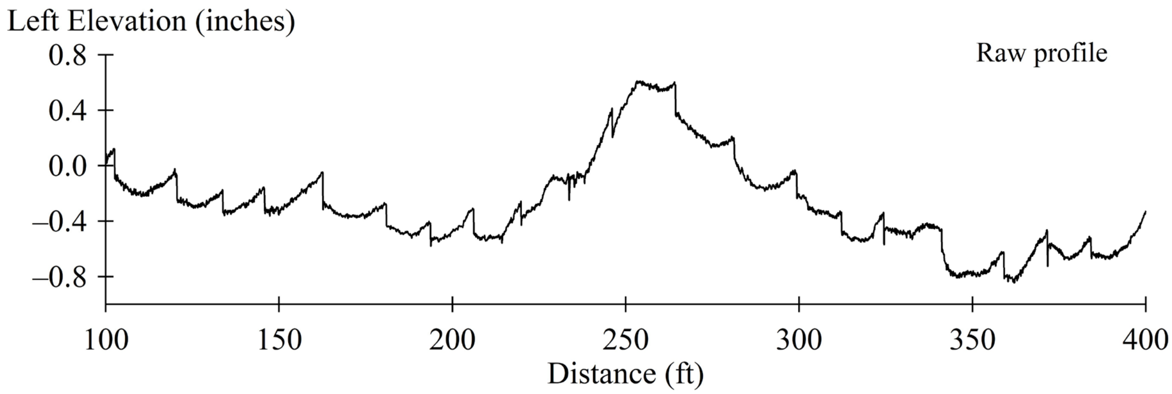
$$p_{HP}(i) = p(i) - p_{LP}(i) = p(i) - \frac{1}{2M+1}\sum_{j=i-M}^{j=i+M} p(j)$$

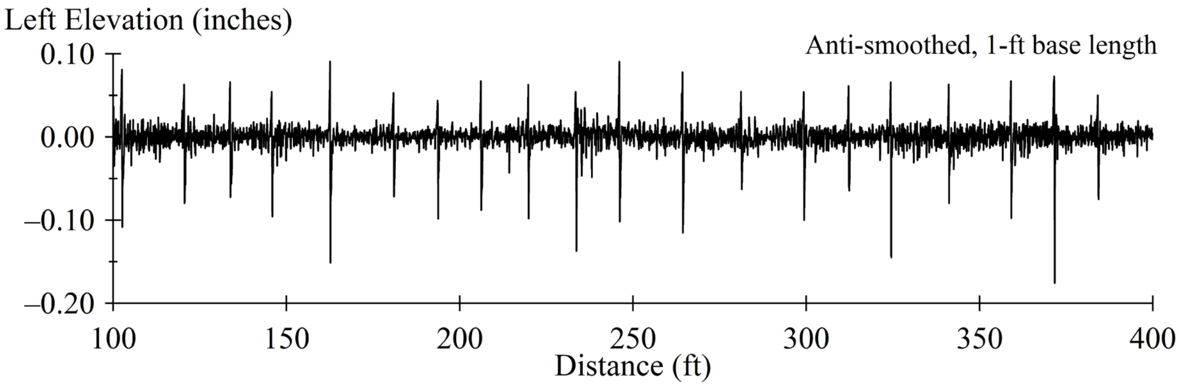
Where p_{HP} is the anti-smoothed profile (also called a high-pass filtered profile). The relationship between the number of points used in the summation and the base length is the same as in the smoothing filter.

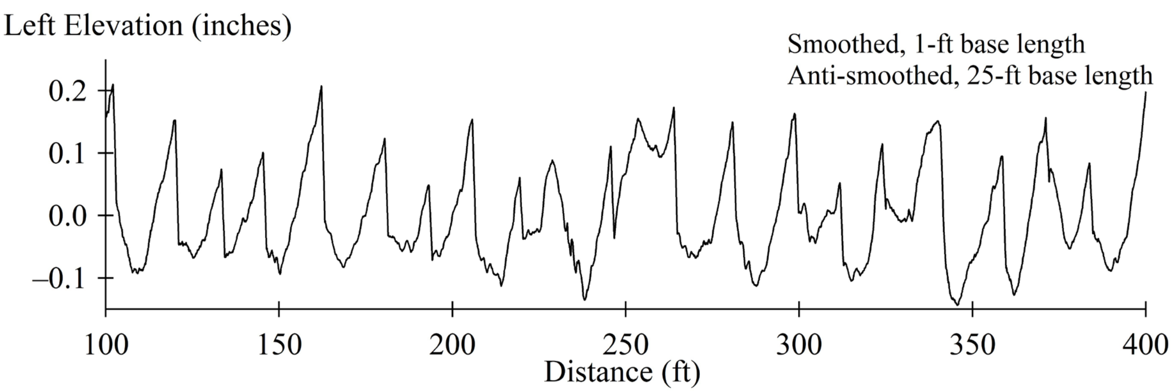
Both the smoothing (low-pass) and the anti-smoothing (high-pass) forms of the filter are useful. Both versions can be used on the same profile in succession, although it only makes sense to apply both filters if the base length is longer for the anti-smoothing version than the base length for the smoothing version.

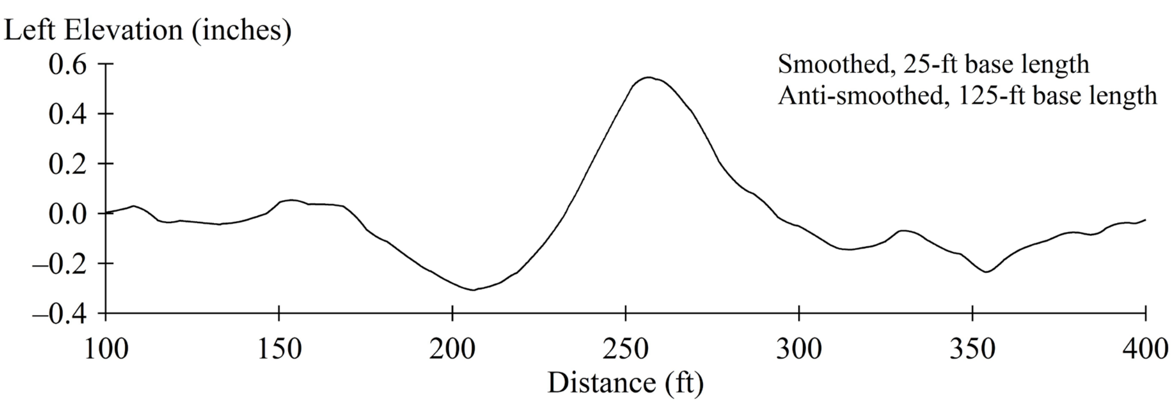
The best choice of base length depends on the use to be made of the data.

For example, the following four plots show the profile of a jointed concrete pavement section with faulting. The first plot shows a segment of profile exactly as it was measured. The other three plots show the profile after it was filtered three different ways. Each plot shows a different kind of information.









The second plot was filtered with the anti-smoothing version of the moving average using a base length of 1 ft. Shorthand for this version is a "1-ft anti-smoothing filter." To generate this plot, the profile was smoothed with a 1-ft moving average and subtracted from the original. The plot shows only the very short-duration deviations in the profile. Upward and downward spikes, which are caused by a combination of downward steps at each fault and narrow dips between slabs, appear at the location of each joint. A close-up view of the plot also reveals the irregular joint spacing pattern.

The third plot shows the profile after processing with a 1-ft smoothing filter and 25-ft anti-smoothing filter. All of the deviations shown in the first plot are eliminated in the second. Although the smoothing filter removed the surface harshness and reduced the abruptness of the faults, the upward slope and upward curvature within each slab are clearly visible.

The fourth plot shows the profile after processing with a 25-ft smoothing filter and 125-ft anti-smoothing filter. In this case, all of the deviations shown in the first two plots are eliminated through the 25-ft smoothing. The 125-ft anti-smoothing filter removed the trends and the longest undulations. The remaining deviations illustrate longer-duration undulations in the road, without the faulting or slab shapes.

A moving average filter is computationally efficient.

The moving average filter is an intuitive way to smooth a profile that is easy to understand and program. It is also efficient computationally. After the first point in the filtered profile is calculated, there is no longer a need to apply the entire summation to get the average. Instead, the influence of the current leading point is added, and the influence of the old trailing point is removed:

$$p_{LP}(i) = p_{LP}(i-1) + \frac{1}{N}\left[p(i+M) - p(i-M-1) \right]$$

Even if the average includes hundreds of points, it is only necessary to account for the effect of two values within the profile: one entering the averaging interval, and one leaving the interval.

What Are Sine Waves?

Sine and cosine waves are both called sinusoids. To understand profile analyses, it is essential to be familiar with sinusoids.

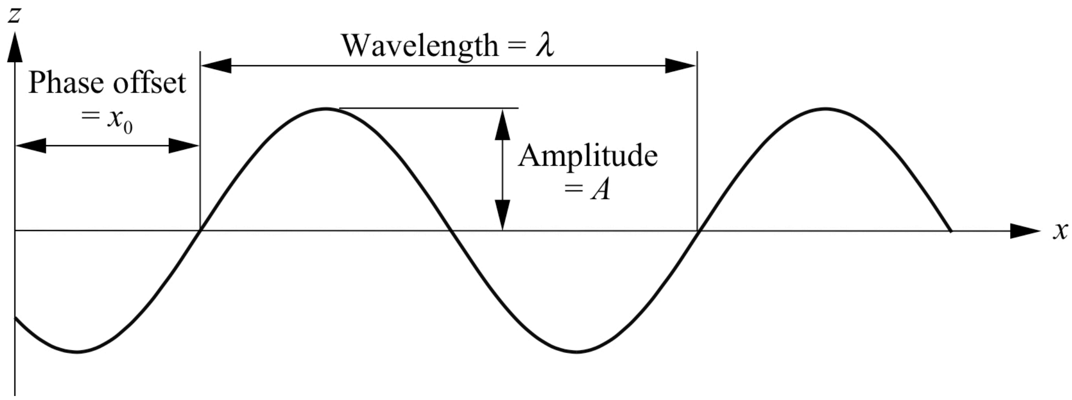
The discussions that follow use sine waves to understand how filters work. Sine waves provide the foundation for discussing wavelengths and frequencies, which in turn provides a systematic way to address topics such as how filters work, how vehicle ride is related to roughness, how various roughness measures are defined, and how measurement errors are caused.

A sinusoid is defined by wavelength, amplitude, and phase.

The equation for a sinusoid (z) as a function of x is:

$$z(x) = A\sin\left(\frac{2\pi}{\lambda}(x - x_0)\right)$$

The amplitude (A) is the maximum vertical deviation from the baseline. A sinusoid has a shape that repeats itself indefinitely. The wavelength (λ) is the distance it takes for each cycle to repeat. The phase offset (x_0) is the distance to the start of the first full cycle.



Frequency is a common way to characterize a sinusoid.

An alternative to defining the length of a cycle is to define how many cycles occur in a unit of length. When a sinusoid is defined as a function of length, it is characterized by spatial frequency, or wave number (ν), which is the reciprocal of wavelength:

$$\nu = \frac{1}{\lambda}$$

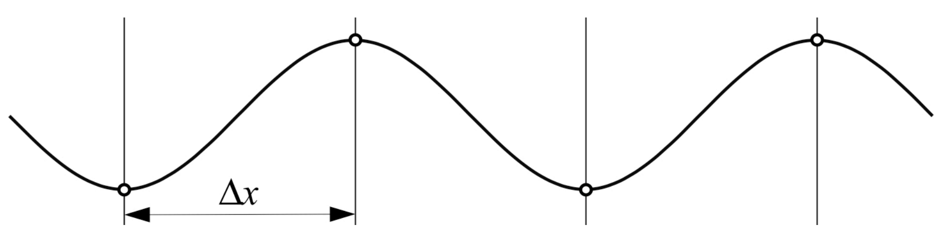
In applications related to profiling, wave number usually has units of cycle/ft.

In many signal processing applications, sinusoids are defined as functions of time, rather than distance, and the convention is to define the sinusoid with the temporal frequency of cycles per second, called Hertz (Hz).

It takes at least two samples per cycle to observe a sinusoid.

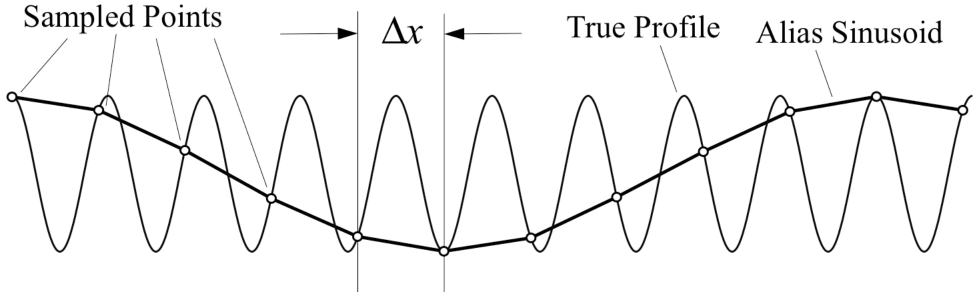
In early work on sampling of analog signals, Nyquist and others advised that digitizing a signal required sampling at a frequency of at least twice the highest frequency that appears in its content. In the context of spatial signals, like road profiles, Nyquist's advice means it is necessary to identify the shortest wavelength where no content appears at that wavelength or below and set the sample interval to one half of that value or less.

Engineers have sometimes incorrectly inferred that content at any wavelength can be accurately recorded at an interval that is half of that wavelength, as shown below. In the figure, any shift in the alignment of the sample points with the peaks and troughs changes the measured amplitude. Worse yet, additional content at shorter wavelengths can cause erroneous artifacts to appear in the recorded signal.



Sampling at an insufficiently short interval can lead to aliasing.

Aliasing is illustrated with a sinusoid in the figure below.



The dots indicate locations where the sinusoid is sampled. Recall that in order to see a sinusoid, the sample interval must be half the wavelength or smaller. The above example shows what happens when the sample interval is too large. Notice that when we connect the sampled values with straight lines, the samples seem to define a sinusoid with a much longer wavelength. This effect is called aliasing. The sinusoid with the long wavelength is an alias of the true sinusoid with the shorter wavelength.

Suppose the profile is being processed with an analysis that has zero response to the sinusoid in the true profile. The problem is that the analysis applied to the profile might respond to the aliased sinusoid, which has a longer wavelength. Given that the aliased sinusoid does not exist in the true profile, this is a source of error.

To avoid aliasing in measured profiles, filter the sensor outputs to remove content at wavelengths shorter than 2 times the sample interval. In systems that sample at a constant time step, filter the sensor outputs to remove content at frequencies above the "Nyquist frequency," which corresponds to ½ of the sampling rate.

What Is Frequency Response?

Familiarity with the concept of frequency response is very useful to users of profile measurements because frequency response plots describe the fundamental behavior of most profile analyses. Filters, roughness index algorithms, sensors, and vehicles are all systems that can be thought of as having an input and an output. Frequency response is a highly useful way to describe the relationship between the input and output of these systems.

A linear system is one in which the output is proportional to the input.

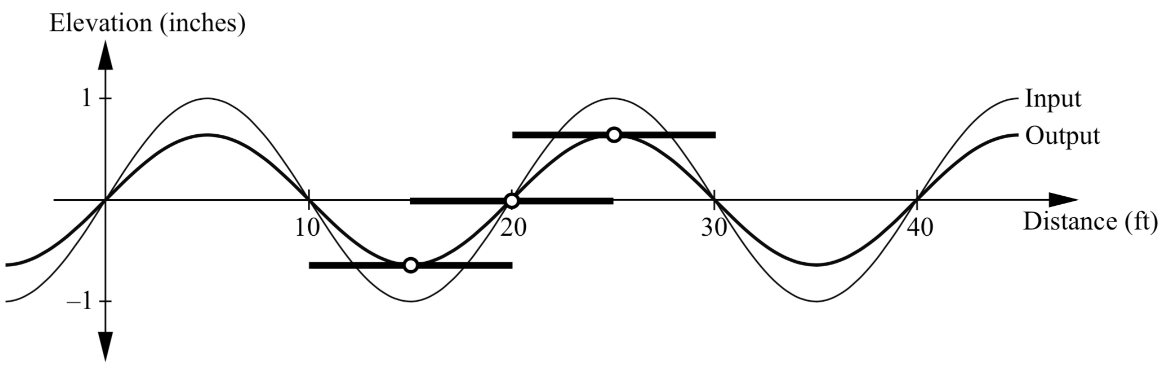
If you have the response of a linear device to a bump in the road, then if the bump is scaled up by some factor, the output of the device maintains its previous form, but it is scaled up by the same factor as the input. For a profile analysis, the test is applied to sets of numbers that are inputs and outputs of the analysis. If the input numbers are changed by a scale factor, the output numbers are changed by the same scale factor if the analysis is linear.

All of the filters described in this book are linear. Some of the methods used to accumulate a summary of the outputs after they are filtered involve nonlinear mathematical functions such as squaring or taking absolute values. However, the filtering process, which is done first, is linear.

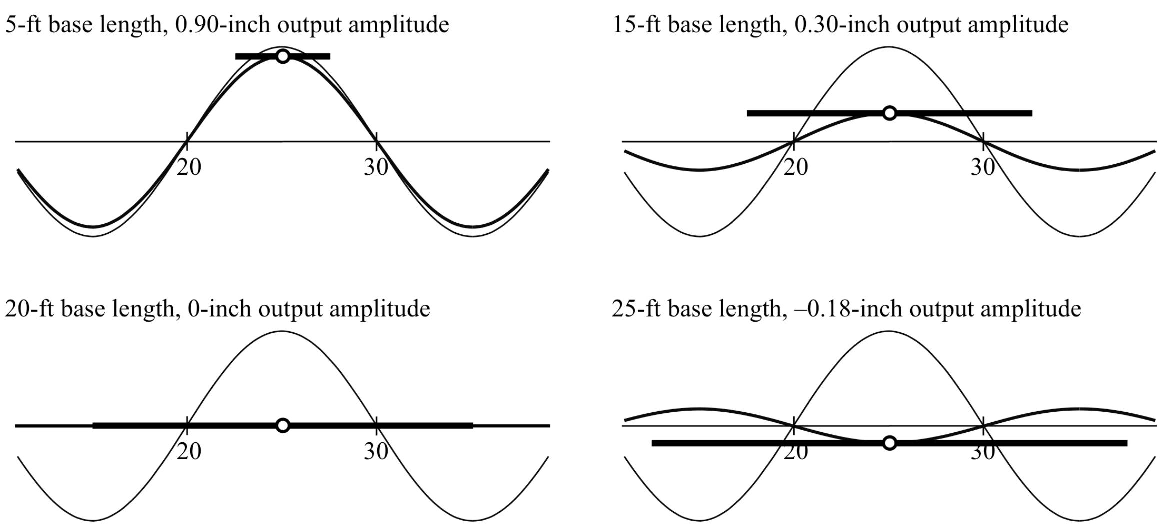
A sinusoidal input to a linear system causes a sinusoidal output with the same frequency.

For a sinusoid expressed as a function of distance, the output sinusoid from a linear system will have the same wavelength as the input sinusoid. In general, the amplitude and phase offset are different for the input and output. However, the wavelength is the same for a linear system.

For example, consider the moving average filter. The next figure shows a sinusoid with a 20-ft wavelength as affected by a moving average smoothing filter with a 10-ft base length. In the output of this filter, the extreme and zero crossing locations are aligned with the extreme and zero crossing locations of the input. That is, the shift in phase offset is zero. The figure shows the range covered by the moving average, which is the base length, at a few critical points. Consider the average of the input signal within the base length at each critical point and imagine the effect of changing the base length. The 10-ft moving average reduced the amplitude of this sinusoid from 1 inch to 0.64 inches.



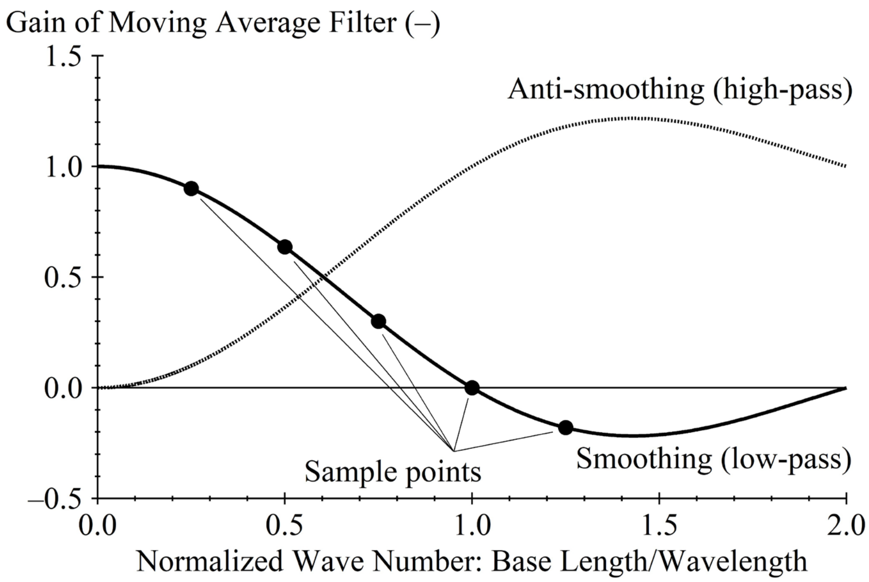
The graphic below shows a close-up view of the effect of other base lengths on the same sinusoid. Using a base length progressively shorter than 10 ft reduces the amplitude less and less, and the moving average with a 5-ft base length only reduces the amplitude to 0.90 inches. Using a base length progressively longer than 10 ft reduces the amplitude more and more. A moving average with a 15-ft base length reduces the amplitude to 0.30 inches. A moving average with a base length of 20 ft completely eliminates the sinusoid, because the average covers exactly one cycle. The 25-ft moving average greatly reduces the amplitude of the sinusoid, but the sign is reversed.



A frequency response plot shows the ratio of output to input for a sinusoid.

Since the output of a linear system is a sinusoid with the same wavelength as the input, the output can be completely defined by the amplitude and phase offset. An amplitude frequency response plot shows the ratio of the output amplitude to the input amplitude for a range of input wave numbers. The amplitude ratio is called the gain. A phase frequency response plot shows the phase offset of the output sinusoid relative to the input for a range of wave numbers. Together, the amplitude and phase frequency response plots fully describe the relationship between the input and output of a linear system. Let's examine the amplitude response of the moving average filter.

The following figure shows a plot of the gain of the moving average smoothing filter as a function of wave number. The plots include markers at points that were included in the examples above. Application of the smoothing filter to a sinusoid with a 20-ft wavelength using a 10-ft base length corresponds to a normalized wave number of 0.5. At this frequency, the plot shows a gain for the smoothing filter of 0.64, which is the ratio of the output amplitude (0.64 inches) to the input amplitude (1 inch).



The figure also shows the gain for the moving average anti-smoothing filter. Recall that the anti-smoothing version of a moving average involves subtracting a smoothed profile from the original. That means the frequency response gain for the anti-smoothing version is one minus the gain of the smoothing version. At any given normalized wave number, the two gains add up to one.

A moving average smoothing filter is called a low-pass filter.

Like any other filter, some things pass through unchanged, some things are reduced, and some might be removed completely. The examples showed that the smoothing filter reduces the amplitude of sinusoids with wavelengths approaching the base length and above the base length. That is, it attenuates or rejects high wave numbers. The smoothing filter attenuates sinusoids with longer wavelengths much less or leaves them intact. That is, it passes low wave numbers. For this reason, it is called a low-pass filter.

The moving average anti-smoothing filter is called a high-pass filter.

Filters can also amplify signals. Note that the anti-smoothing version of the moving average amplifies sinusoids when the ratio of base length to wavelength is between 1 and 2. The anti-smoothing filter passes short wavelengths and rejects long wavelengths. Since short wavelengths correspond to high wave numbers, anti-smoothing is a high-pass filter.

The table below summarizes the behavior of each filter.

Filter	Smoothing	Anti-smoothing
Type	low pass	high pass
Passes	low wave numbers (i.e., low frequencies) long wavelengths	high wave numbers (i.e., high frequencies) short wavelengths
Rejects	high wave numbers short wavelengths	low wave numbers long wavelengths

The gain of filters often depends on the ratio between wavelength and a filter parameter. Therefore, they are often plotted using normalized frequency or wave number, as was done in the moving average frequency response plot. Normalization helps the user select a filter type and the proper filter setting once the goal of the filtering is defined.

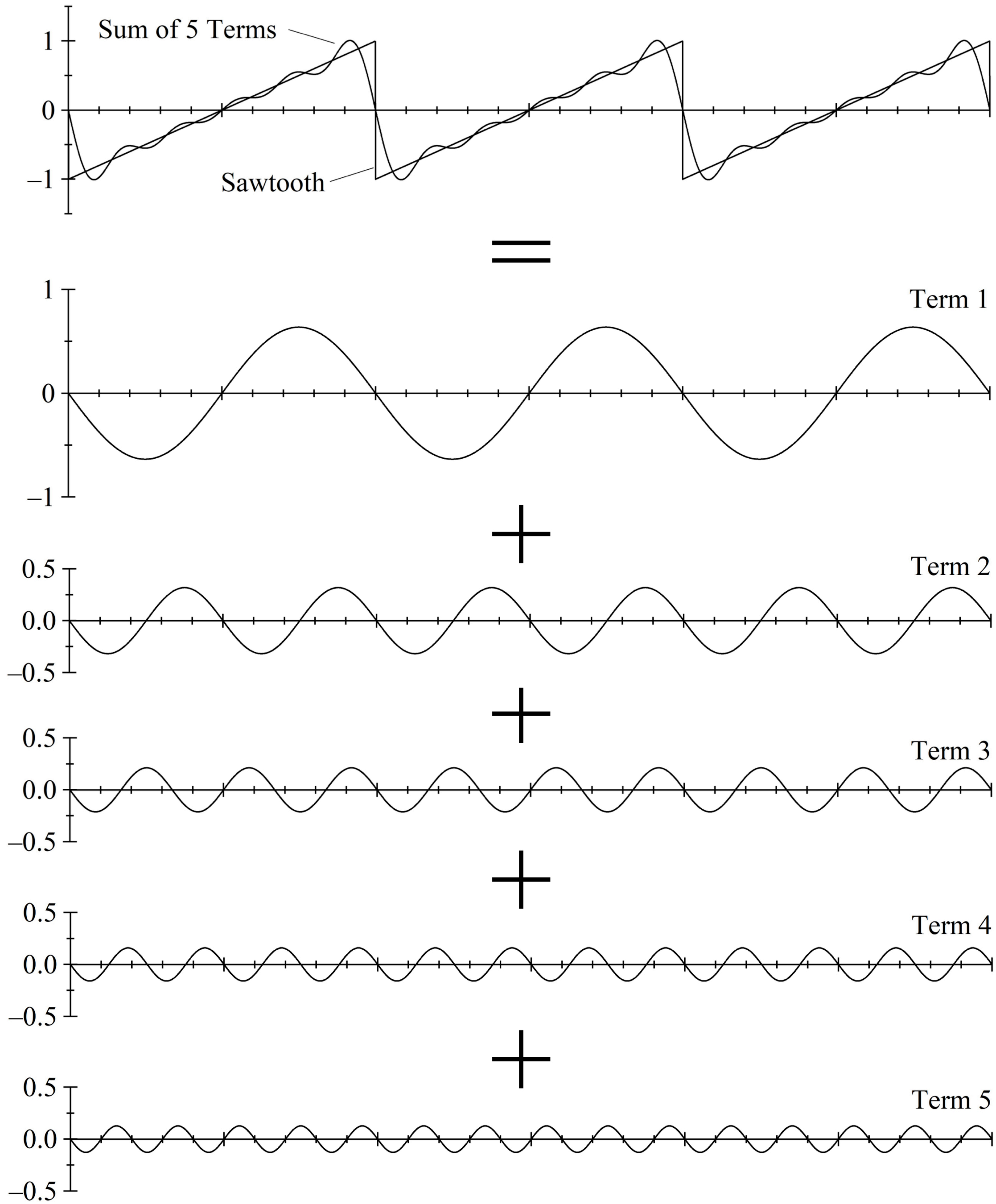
The moving average examined in this section is a very simple example of a finite impulse response (FIR) filter, which confines the response of finite disturbances to a finite distance. Several other FIR filters exist, and much more detail is available in the literature for a deeper exploration of FIR filters. Useful search terms include Bode plot, FIR filter, windowing, Gaussian filter, and Hanning window.

What Is a Fourier Transform?

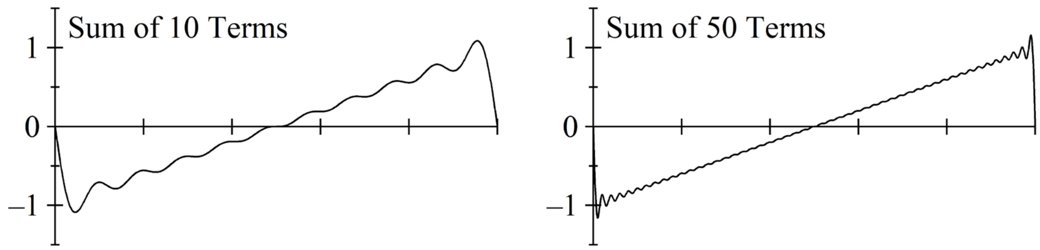
A typical road profile has no direct resemblance to a pure sinusoid. Instead, a typical road profile encompasses a spectrum of sinusoidal wavelengths. The power spectral density (PSD) function is a statistical representation of the importance of various wavelengths. PSD functions are created using a Fourier transform. This section introduces the Fourier transform to provide background for the discussion of PSD functions. Understanding the Fourier transform is useful when interpreting PSD plots, but it is not mandatory. The section following this one provides information about PSD functions, and the section after that addresses the interpretation of profile PSD plots.

Periodic signals can be decomposed into a series of sinusoids.

An arbitrarily shaped signal can be constructed mathematically from a series of sinusoids with different wavelengths, amplitudes, and phase offsets. For example, consider a repeating pattern like the sawtooth shown in the figure below. The Fourier series approximates the repeating shape using the sum of several sinusoids. The Fourier series reconstructs the shape of the signal using a sum of sinusoids with the same frequency, twice the frequency, three times the frequency, etc. For a spatial signal like a profile, the frequencies correspond to wavelengths that equal the total length, 1/2 of the length, 1/3 of the length, etc. The figure below shows the first five terms and their approximation of the sawtooth.

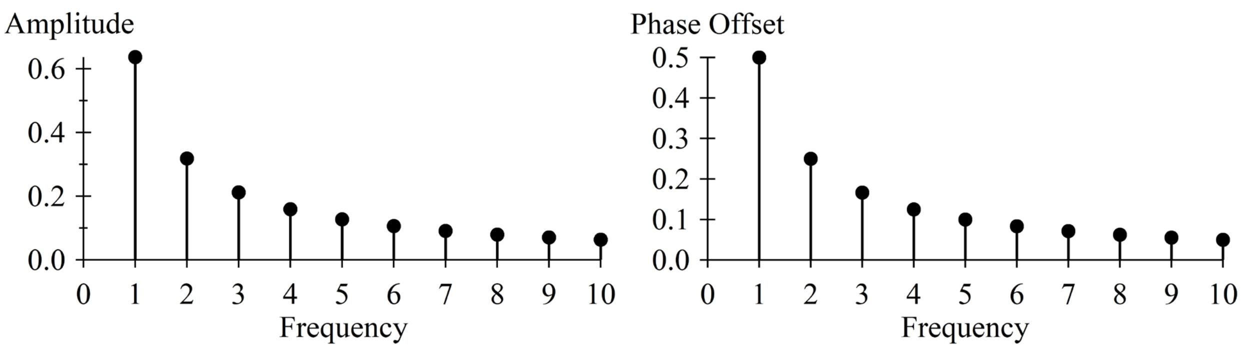


Each term in the Fourier series represents an independent contribution to the overall shape at a given frequency; and no term in the series can be replaced by terms at other frequencies. The figures below show approximations of the sawtooth with more and more of the frequency content included. Instructional materials about Fourier transforms often include examples of this kind. However, an interesting property of analog signals with step changes, like the sawtooth, is that a sum of sinusoids never perfectly replicates the discontinuities. For the sawtooth, even using an infinite number of terms overshoots the high and low points of each tooth. (To learn more, search on "Gibbs phenomenon.")



Plots of amplitude and phase provide an alternative representation of a signal.

For an analog signal, the Fourier series includes an infinite number of terms. The figures below show the amplitude and phase offset of the first 10 terms in the Fourier series for the sawtooth. In this case, amplitude and phase offset relate systematically to frequency. If each tooth has duration of 1 unit (e.g., 1 ft, 1 sec) the frequency (f) is equal to the number of the term in the series (n). The amplitude of each term is $2/(n\pi)$, and the phase offset is $1/(2n)$.



In the figure above, phase offset is expressed in the same units as the duration of each cycle. That is, phase offset represents the shift of each sinusoid along the x-axis. For the sawtooth, the phase offset of each term in the Fourier series is half the duration of the sinusoid. In many practical applications, phase is not presented as an offset. Instead, it is presented as a leading angular shift within the 360-degree cycle for a sinusoid. That changes our definition of a sinusoid as follows:

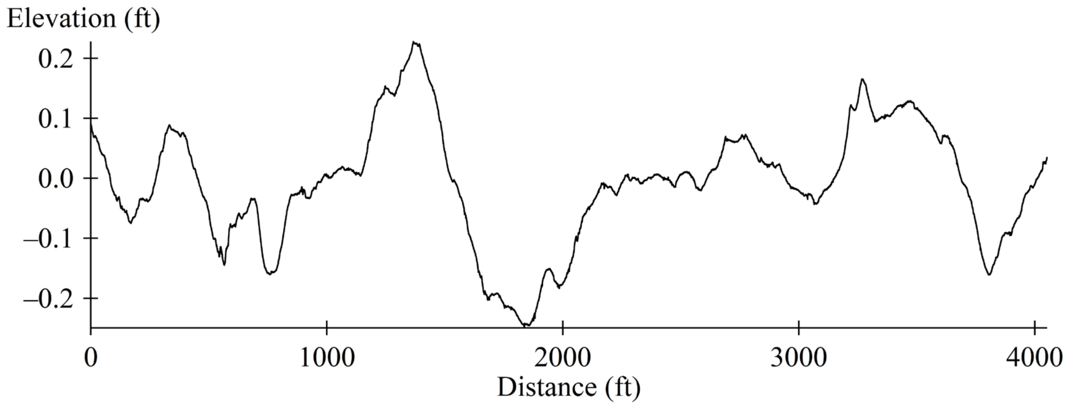
$$z(x) = A \bullet \sin\left(\frac{2\pi}{\lambda}(x - x_0)\right) = A \bullet \sin\left(\frac{2\pi}{\lambda}x + \phi_0\right)$$

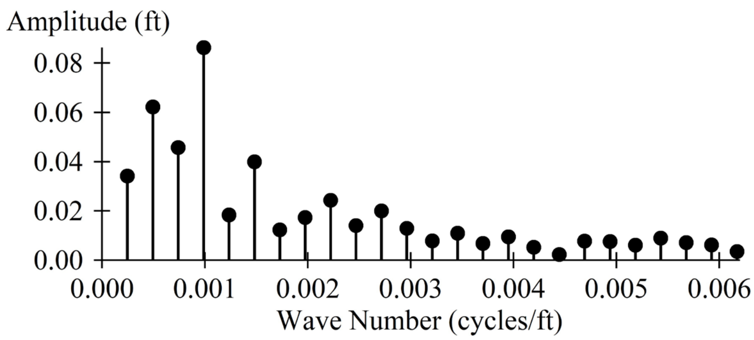
For the sinusoids in the Fourier series that make up the sawtooth, the phase angle (ϕ_0) of every term is 180 degrees (π radians), which means a leading offset of one-half cycle. (See the sine waves in the figure above.)

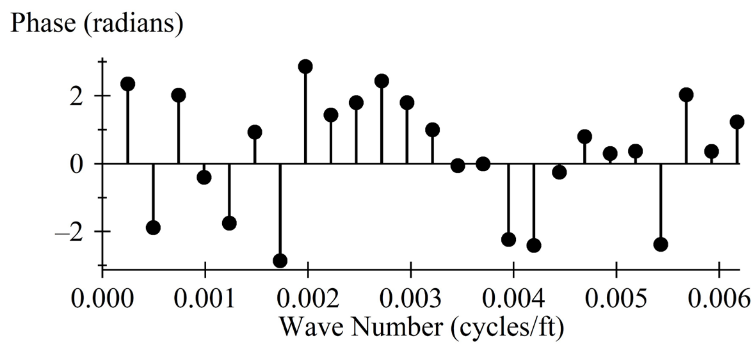
A measured profile can be decomposed into a series of sinusoids.

A discrete Fourier transform provides an alternative representation of a digital signal. The discrete Fourier transform converts the sampled points in a measured profile to a set of amplitudes and phase angles. For a profile with length L, the discrete Fourier series includes terms with wave number values equal to $1/L$, $2/L$, $3/L$, etc. That corresponds to wavelengths of L, *L*/2, *L*/3, etc., respectively. The shortest wavelength in the series is equal to twice the recording interval. Unlike the sawtooth, the phase angles for typical profiles do not follow a systematic trend. However, some broad trends appear in the relationship between amplitude and wavelength.

Consider the measured profile shown below. The profile includes 32,768 points recorded at an interval of 0.1237 ft, which covers a total length of 4053.0 ft. The Fourier transform for this profile includes 16,384 terms, which correspond to wavelengths of 4053.0, 2026.5, 1351.0, etc. down to 0.2474 ft. Each term includes a value of amplitude and phase angle. The figures that follow show the amplitude and phase angle of the first 25 terms of the Fourier transform.

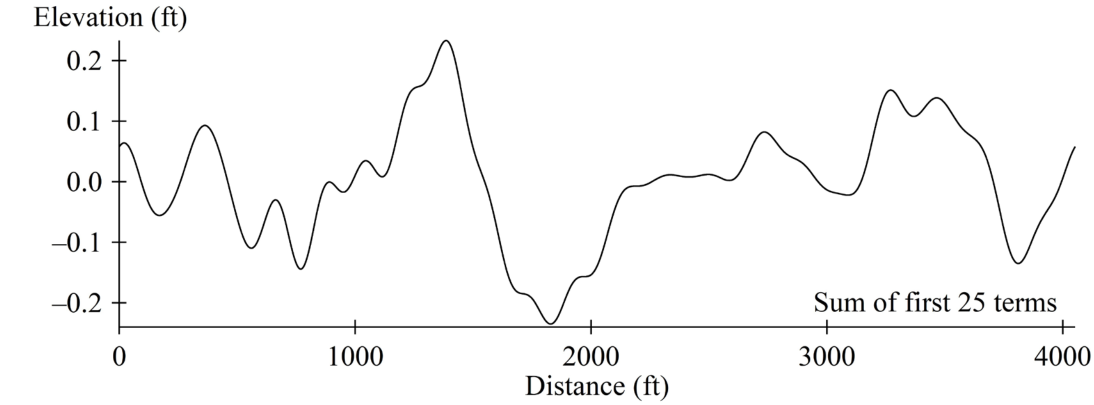




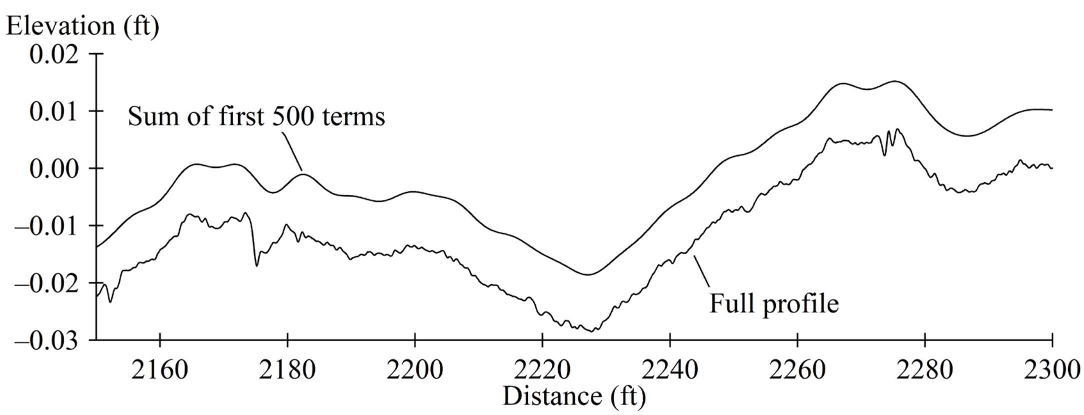


Each term in the Fourier series represents a small contribution to the overall spectrum. The first term corresponds to a wave number of 0.000247 cycles/ft, which is a wavelength equal to the total length of the profile. The first 25 terms only capture components of the original profile with wavelengths greater than or equal to one twenty-fifth of the total length, which is 162.1 ft. The figure below shows a reconstruction of the partial profile using only the first 25 terms in the

Fourier series. The long undulations from the original profile are present, but not the variations associated with shorter wavelengths. Reconstruction of the profile in this manner is equivalent to applying a low-pass filter with perfect retention of content at wavelengths of 162.1 ft and above, and perfect rejection of content at wavelengths shorter than 162.1 ft.



The next figure compares a close-up view of the profile to a reconstruction of it using the first 500 terms in the Fourier series. The first 500 terms include the content at wavelengths of 8.11 ft and above. Note that the reconstructed profile still excludes many of the short-wavelength asperities present in the full profile. The full set of 16,384 terms in the Fourier transform captures the entire profile. No information is lost. Rather, it is reorganized to represent components of the profile in terms of amplitude and phase angle versus wave number, rather than elevation versus distance.



Understanding discrete Fourier transforms involves many details.

This discussion omitted several important technical issues pertaining to the calculation of the Fourier transform. For example, the core algorithm used to apply the transform required the input to include a number of data points equal to 2 raised to an integer power. Often, signals are padded with zeros at the end to overcome this limitation. The Fourier transform also assumes that the input signal repeats itself, so differences in height at the two ends are treated like discontinuities. The sawtooth example demonstrated the artifacts that might result from that assumption. For more information, search on the term "discrete Fourier transform" using the sub-topics windowing, padding, and bleeding.

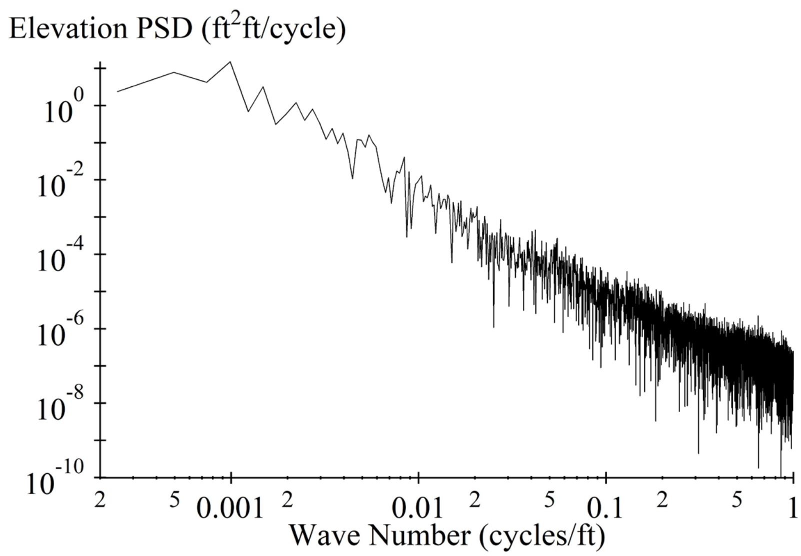
What Is Power Spectral Density?

The power spectral density (PSD) function is a statistical representation of the importance of various frequencies. It is based on the Fourier transform, with scaling that shows how the variance of a signal is distributed over a range of frequencies.

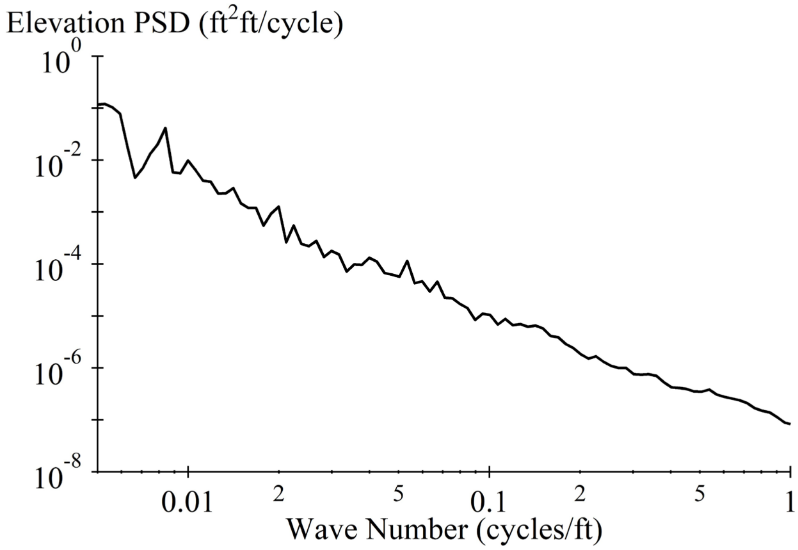
A PSD function shows how variance in a profile is distributed over wave number.

The "power" in the name comes from early applications in electronics. In a PSD of a road profile, the variance has units of elevation squared, which has nothing to do with power. PSD functions calculated from time-based signals provide variance as a function of frequency in cycles/sec. For profiles, frequency is expressed in terms of wave number (e.g., cycles/ft).

The figure below shows a raw PSD function constructed for the profile examined in the discussion of Fourier transforms. This PSD function includes one point for each term in the Fourier transform calculated from the example profile. The point at each frequency is derived from the corresponding amplitude but rescaled to represent the variance of that sinusoid. In this case, each point in the plot represents a waveband that is 0.000247 cycles/ft wide. Multiplying the bandwidth for each point by the magnitude of the PSD function yields units of variance in ft^2. For example, the first point in the PSD plot below is 2.3616 ft^2ft/cycle. When that is multiplied by the bandwidth (0.000247 cycles/ft), the square root of the result is 0.0241 ft. This is the root mean square of the first sinusoid in the Fourier series, which is the amplitude of the first term (0.0341 ft) divided by the square root of 2.



The next figure shows the same PSD function with additional processing to consolidate the large number of points into a smoother plot, in which the trends are easier to recognize. This is a more common way to view PSD functions for signals like profiles. In this plot, the points are evenly spaced on a logarithmic scale. The specific spacing applied here results in 12 points each time the frequency doubles. The spacing is called 12 bands per octave. Each point in the plot includes the influence of several terms from the Fourier transform. The fundamental scaling of the PSD is maintained, and each point identifies the contribution to variance of the terms contained within the waveband that it represents. The smoothing effect associated with using wider bands in this manner makes the plot easier to interpret. However, information about phase angle is discarded when consolidating the individual terms from the more detailed plot.

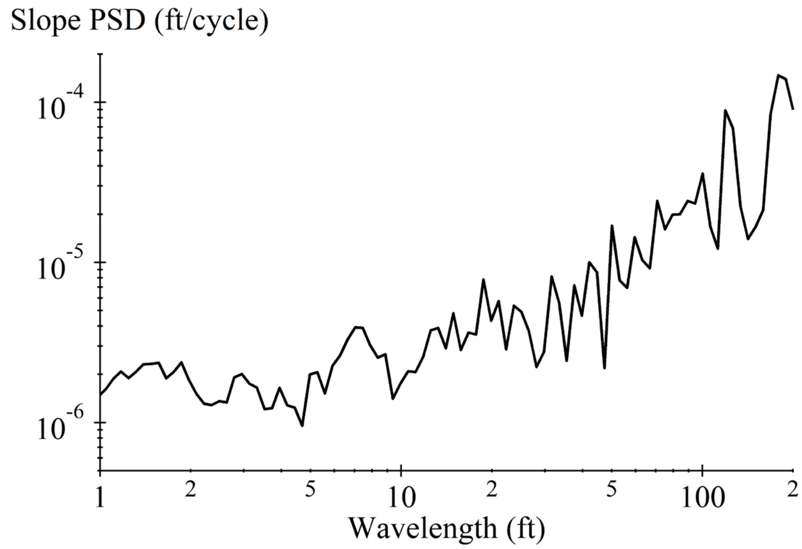


The specific relationship between amplitude and wave number depends on the profile. However, most profiles share a similar trend, in which the amplitude diminishes rapidly with increasing wave number. That is, the amplitude gets smaller as wavelength decreases.

The amplitude of profile slope is more uniform than the amplitude of elevation.

PSD functions can be computed for derivatives of profile elevation, such as slope and curvature (i.e., spatial acceleration). PSD functions of profile slope cover a much smaller vertical scale than PSD functions of elevation, because the basic spectrum of roughness over wave number is more uniform. Slope PSD functions provide an easier way to see the details of the frequency content.

The next figure shows a common alternative to PSD of elevation versus wave number for viewing the frequency content within a profile. First, the figure displays PSD of slope, rather than elevation, for the example profile used in the previous plots. The plot covers the same waveband as the elevation PSD, but the vertical range shown on the plot is reduced from six orders of magnitude to two. Second, the figure displays slope PSD versus wavelength, rather than wave number. Pavement engineers often find wavelength more practical or intuitive than wave number.



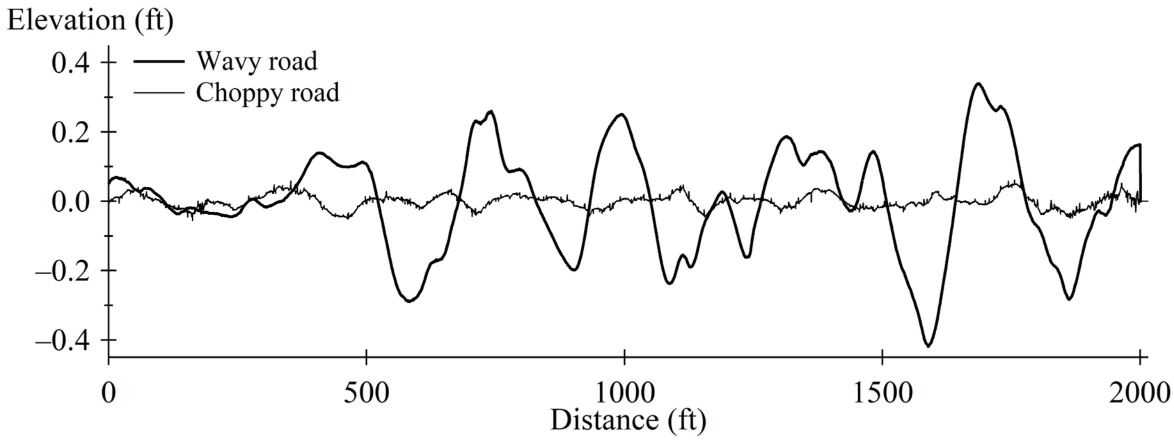
Be advised that PSD of slope has unusual units of slope squared per wave number. For unitless slope, such as ft/ft, the units reduce to the reciprocal of wave number (e.g., ft/cycle). Be further advised that slope PSD versus wavelength obscures a useful feature of PSD analysis: The variance of a signal for a given frequency range is the integral of the PSD function for that range. In a plot of PSD versus wave number using linear scaling on both axes, the relative contribution of each waveband is visible at a glance. Plots of PSD versus wavelength with logarithmic scaling are two steps removed from that interpretation.

What Does a Power Spectral Density Plot Reveal about a Profile?

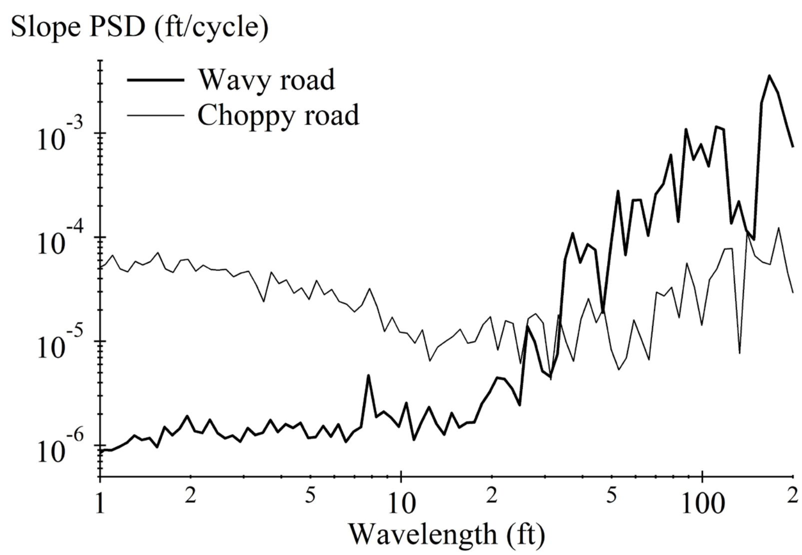
Viewing a measured profile or filtered profile plot helps isolate features of interest by their shape and location. The PSD plot obscures that information. What it provides instead is a view of the frequency content.

A profile PSD plot provides a view of the relative importance of each wavelength range.

Consider the two dissimilar road profiles shown in the plot below. One of the profiles was measured on a parkland road intended for very low speed travel. The design requirements did not allow for alteration of the long undulations in the terrain, but the paving process smoothed out the short undulations. The other profile was measured on a high-speed minor arterial roadway. This road's design profile was flat and level. However, the profile includes roughness caused by frost heaves, transverse cracking, and thick sealant at many of the cracks. The parkland road profile appears wavy relative to the choppy profile from the distressed minor arterial roadway. Their IRI values, which quantify their overall roughness, are very similar.



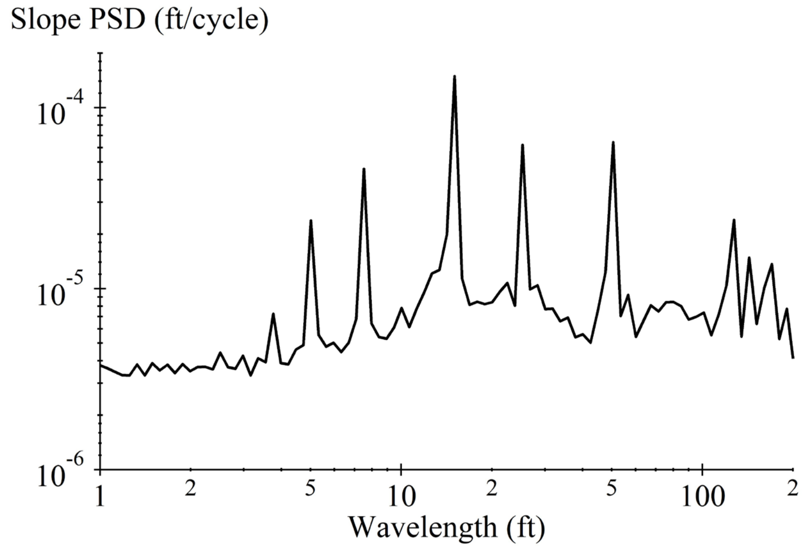
The next figure compares the slope PSD functions for the two profiles. Roughness in the wavy profile exceeds the roughness in the choppy profile for wavelengths greater than 30 ft. In contrast, the choppy profile includes more roughness than the wavy profile for wavelengths shorter than 30 ft. Each end of the spectrum shown in the figure affects vehicle response to the profile differently, as described in other sections of this book. The efficacy of various corrective strategies also depends on whether the dominant source of roughness is associated with long-wavelength content or short-wavelength content.



Interpret with caution cases when the slope PSD increases steeply as wavelength diminishes. The increase has several potential causes that are hard to distinguish using the slope PSD plot alone, including many small distresses, a small number of severe distresses, downward spikes in the profile at deep cracks or joints, and poor data sampling practices.

A profile PSD plot helps identify sources of repetitive roughness.

The following figure shows the slope PSD for a jointed concrete pavement. The slope PSD includes spikes at several wavelengths. Given the logarithmic scaling of the plot, it appears that the spikes account for a large share of the roughness in terms of profile slope. The spike at a 15-ft wavelength corresponds to the slab length. The slabs within this pavement section curled downward. The curling created a pattern in the profile with a concave downward shape that repeated every 15 ft. The spike at a 50-ft wavelength corresponds to the stringline stake spacing. Sag in the stringline imposed a pattern in the profile with a concave upward shape that repeated every 50 ft.

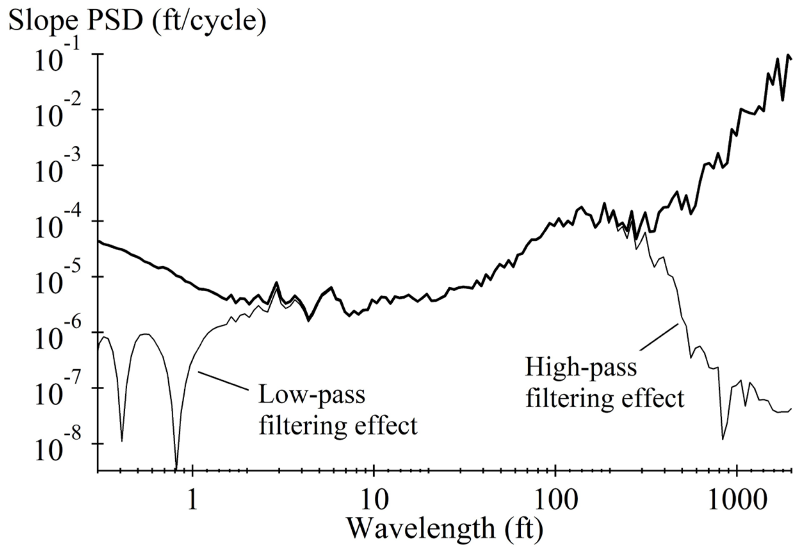


Spikes appear in the slope PSD plot at other wavelengths besides 15 ft and 50 ft. The spikes at 7.5 ft, 5 ft, and 3.75 ft are upper harmonics of the curl slab shape. This is called an upper harmonic, because each spike occurs at a higher frequency. Each spike occurs at a wave number that is a multiple of one over the slab length: 0.067 cycles/ft. Like the sawtooth in the Fourier transform example, the curled slab profiles include contributions from several sinusoids, each with a wavelength that divides evenly into the fundamental length. Similarly, the spike at 25 ft is an upper harmonic of the pattern caused by stringline sag.

In the example above, the concentrated roughness caused by curling and stringline sag accounted for a large share of the overall roughness. As such, an experienced analyst could identify these features without the PSD plot using filtered profile plots. Even so, each of the two major sources of roughness obscured each other somewhat in the profile plots. When repetitive roughness is less dominant, profile features from other wavebands also obscure them. The PSD plot provides an efficient way to identify the presence of repetitive features in a profile that cause roughness concentrated at a specific wavelength, and the relative importance of the upper harmonics. Note that the PSD plot did not indicate the direction of curl, or the concave upward nature of the stringline sag.

Evidence of filtering appears in PSD plots.

Filtering affects the range of wavelengths that appear in the profile. The plot below shows the slope PSD for a profile measured with minimal filtering. An additional trace appears on the plot for a version of the profile after the application of two post-processing filters: (1) a high-pass filter with a cut-off wavelength of 300 ft, and (2) a moving average smoothing filter with a base length of 0.82 ft.



The moving average smoothing (low-pass) filter reduced content at wavelengths of twice the base length or less and virtually eliminated content at wavelengths equal to both the base length and half of the base length.

The high-pass filter was a type of Butterworth filter, which is not described in this book. The "cut-off" wavelength characterizes the gain response of the Butterworth filter. In this context, the cut-off wavelength is the wavelength at which the PSD of the filtered profile is reduced to half of the value of the PSD of the unfiltered profile, as shown. The high-pass filter further reduced the content at wavelengths longer than the cut-off. Inertial profilers often apply a high-pass filter, because they have difficulty measuring profile features with very long wavelengths accurately.

The PSD plot provides a way to detect the presence of filtering in the profile. To detect a high-pass filter, the profile must cover a length that is many times the cut-off wavelength. Similarly, detection of a low-pass filter requires a sufficiently short recording interval. The example above provides a clean view of the effects of filtering. In practical examples, the unfiltered profile may not be available for comparison, because not every profiler retains the unfiltered signals.

What Is Vehicle Ride?

The ride quality experienced by the public strongly affects their satisfaction with the roadway.

Ride quality is measured using accelerations in the vehicle body.

Ride comfort is subjective, and it depends on several factors. Automotive engineers measure accelerations at interfaces between the vehicle and the driver or passenger for objective estimation of ride quality. This includes the seat/buttock interface, the seat/back interface, the floor/foot interface, and the steering wheel/hand interface. Other factors, such as noise, affect overall ride comfort; it is often difficult to distinguish these factors from the influence of tactile vibrations. Even so, the strong link between occupant accelerations and perceptions of ride quality is well established.

Human response to vibration is an important aspect of ride quality.

A large body of literature exists that addresses human response to vibration. The literature includes studies in several contexts related to transportation (cars, trucks, ships, trains, aircraft, spacecraft, tanks, etc.). For seated passengers, experimental studies usually include a combination of objective motion measurements and subjective evaluation of comfort on laboratory shaker tables, on laboratory ride simulators, or in a test vehicle.

A broad assessment of the literature shows that the human body has a minimal tolerance to vertical vibration in the range from 3 Hz to 8 Hz. Most seated humans will experience resonant vibrations of organs in their abdominal cavity somewhere in this range. That is, their guts will shake at odds with the rest of their body. The feeling is quite uncomfortable.

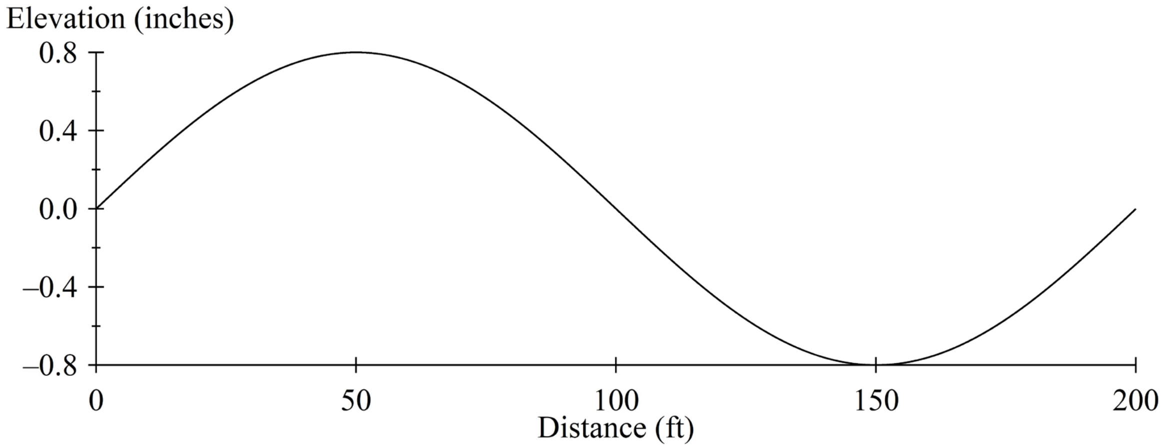
Minimum tolerance for horizontal acceleration occurs at 1 Hz. Lateral acceleration may be caused by vehicle roll when variations in the left and right profile do not align. Most high-speed roads do not produce much roll excitation unless they are in poor repair. Roll input from the roadway can be significant on low-speed roadways. Longitudinal acceleration can result from pitch of vehicles when high and low points on the profile have the same spacing as a vehicle's wheelbase. Pitch and roll impose more pronounced effects on riders of high vehicles such as vans, utility vehicles, and trucks. Ride engineers normally tune the front and rear suspensions to minimize pitch; however, it is not always possible on trucks. Hence, truck drivers experience more pitch induced longitudinal vibrations from road roughness than occupants of passenger cars.

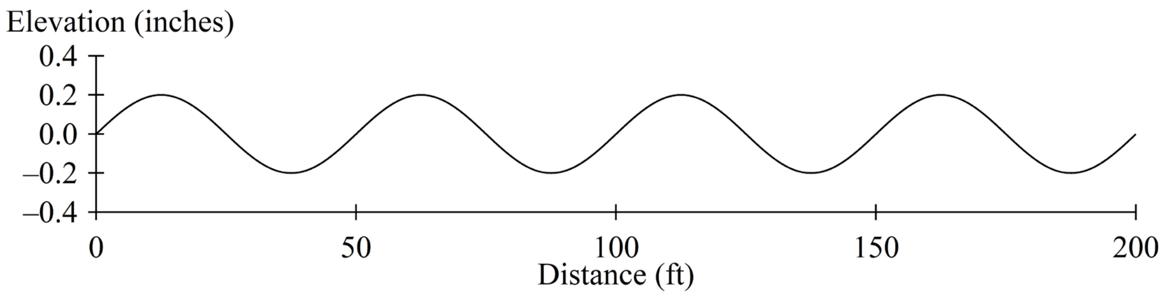
Acceleration level is only one aspect of ride comfort.

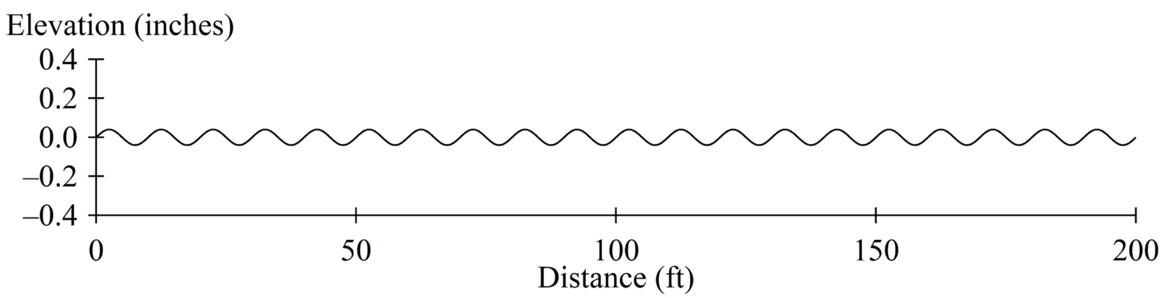
The research community does not fully agree on the use of acceleration to measure comfort or the relative importance of each frequency range. Experimental results from different practical contexts are often difficult to reconcile. Further, some experiments examine the limit of perception, others examine comfort, and still others examine health and safety. Many studies have endorsed the use of "jerk" as an alternative to acceleration. Jerk is the time rate of change of acceleration. The importance of jerk hints at an important aspect of ride quality: In some situations, a single, short-lived, intense vibration event may influence a person's perception of ride quality to a greater extent than the nominal level of consistent background vibration.

How Is Vertical Acceleration Related to Profile?

Given the strong relationship between vertical acceleration and vehicle ride comfort, it is important to understand the relationship between profile elevation and acceleration. Consider an example using the three sinusoids shown below. To represent them as acceleration inputs to a passing vehicle, we need to take their derivatives twice and consider the vehicle speed.







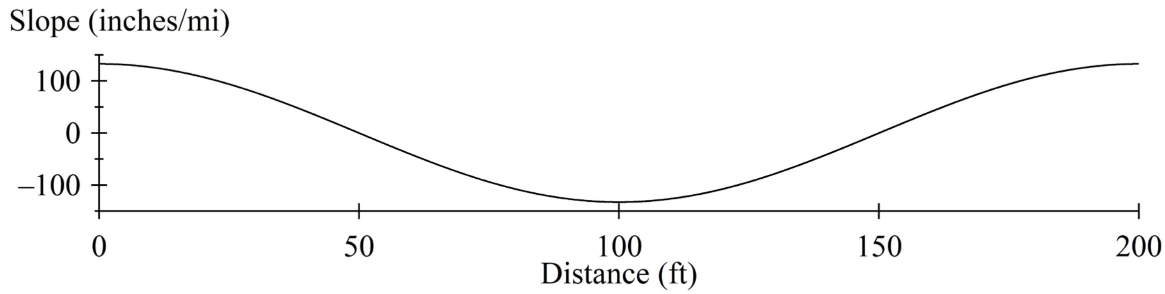
The derivative of a sinusoid is a sinusoid with the same wavelength.

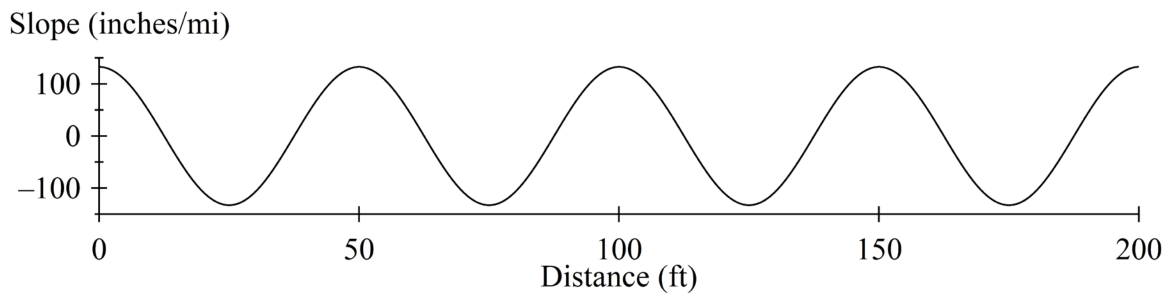
To convert the elevation sinusoid to slope, take the derivative with respect to longitudinal distance. The amplitude of the slope sinusoids is:

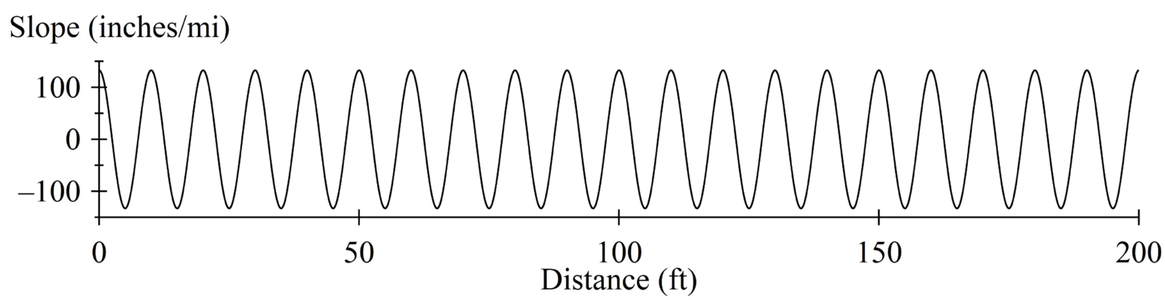
$$\text{Slope amplitude} = \frac{2\pi A}{\lambda}$$

The units of slope depend on the units of elevation amplitude (A) and wavelength (λ). When A and λ have the same units (e.g., both ft), multiply by 12•5,280 to get inches/mi.

The spatial derivative of each example sinusoid is a profile of slope, as shown below. Each slope profile has the same wavelength as the corresponding elevation profile. Using the formula above, all three profiles have the same slope amplitude. The dimensionless value is 0.002094, which is equal to about 132.7 inches/mi. Converting from elevation profile to slope profile escalates the magnitude for shorter wavelengths more than longer wavelengths, because the reversals occur more frequently.







Taking the derivative of each slope profile produces sinusoids that represent spatial acceleration. This further escalates the magnitude of sinusoids with shorter wavelengths relative to sinusoids with longer wavelengths.

Travel speed affects the way vehicles experience the road profile.

A vehicle moving over the road experiences a sinusoid as a function of time. Frequency and cycle duration have definitions in both time and space. Note that so many engineering applications use temporal frequency that it is usually just called "frequency" for short, and its units are named after a famous physicist.

Domain	**Cycle Duration**	**Sample Units**	**Frequency**	**Sample Units**
Distance	Wavelength, λ	ft	Wave number, $\nu = 1/\lambda$	cycles/ft
Time	Period, T	sec	Frequency, $f = 1/T$	cycles/sec = Hertz (Hz)

For travel over a road profile, a particular wave number can be converted to frequency by multiplying by speed (V). A particular wavelength can be converted to a time period by dividing by speed:

$$f = \nu V,\ T = \lambda / V$$

To understand how a vehicle experiences a sinusoidal profile, it is often useful to calculate the frequency that corresponds to its wavelength:

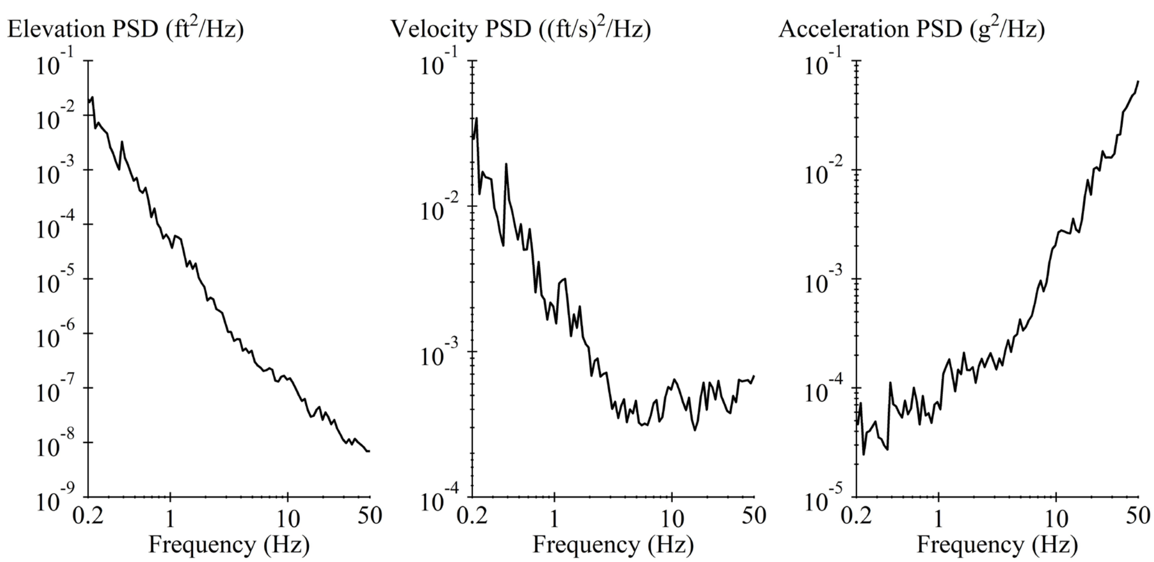
$$f = V / \lambda$$

The table below lists some characteristics of the three example sinusoids for a speed of 100 ft/sec, which is approximately 68.2 mi/hr. Note that vertical acceleration amplitude is proportional to elevation amplitude divided by the square of wavelength. For the sinusoid with a 10-ft wavelength, the acceleration amplitude is more than 40 percent of the acceleration due to gravity.

Wavelength	**Amplitude**	**Slope Amplitude**	**Frequency**	**Acceleration Amplitude**
200 ft	0.8 inch	132.7 inches/mi	0.5 Hz	0.02 g
50 ft	0.2 inch	132.7 inches/mi	2.0 Hz	0.08 g
10 ft	0.04 inch	132.7 inches/mi	10.0 Hz	0.41 g

Profile PSD functions can be shown as acceleration.

The calculations applied to the three example sinusoids can be extended to the broad spectrum of wavelengths in typical roads. The figure below shows PSD functions calculated for a 6.14-mile-long profile that includes the 4053-ft-long segment discussed in the "What Is a Fourier Transform?" and "What Is a Power Spectral Density?" sections of this book. The left side of the figure below shows the road as an elevation PSD input to a vehicle when its travel speed is 50 mi/hr. Since the speed is known, the PSD function can be plotted versus frequency, rather than wavelength or wave number. Taking the derivative once produces the spectrum of vertical velocity, and differentiating again produces the spectrum of vertical acceleration. Note that acceleration input is greatest at high frequencies, which correspond to short wavelengths.



The assumed travel speed acts to scale the PSD functions. However, the profile PSD functions retain the same shape, whether they are calculated as a function of wave number or frequency in Hz. The elevation PSD with units of length^2/Hz has the same shape as the elevation PSD as a function of wave number. Only the units are changed, to involve time rather than distance. The velocity PSD has the same shape as the slope PSD. Although it has not been shown, the acceleration PSD corresponds to a spatial acceleration PSD.

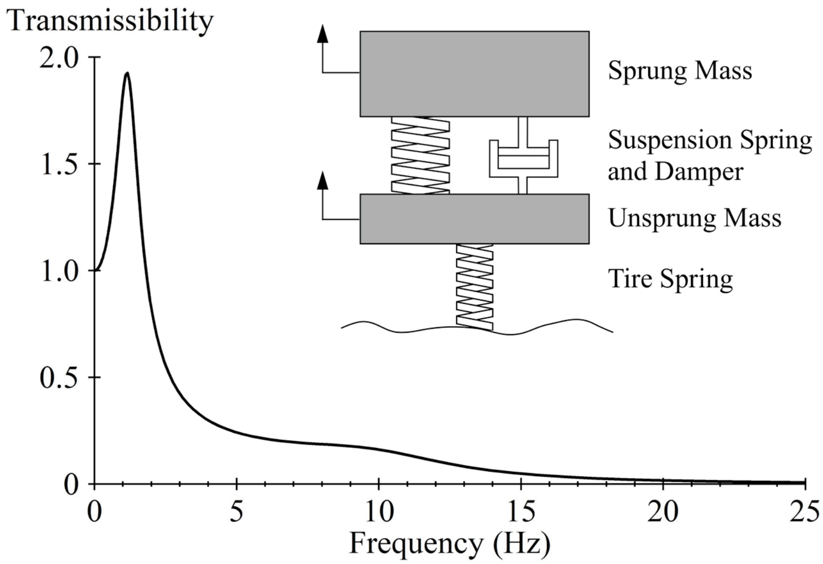
How Does Ride Relate to the Road?

Although road roughness is a dominant cause of vibration on a motor vehicle, the public is able to separate the role of the car from that of the road. Past experiments have demonstrated that the vehicle type that people are riding in (e.g., luxury or compact) influences their judgment of their overall ride experience. However, when asked to rate the road, they tend to look beyond the vehicle and rate the roughness of the road comparably regardless of the type of vehicle they are riding in.

Automobile suspensions isolate the rider from the severe acceleration inputs of the road.

A quarter-car model provides a very simple, but useful, representation of important aspects of the ways suspensions and tires isolate the vehicle body from road roughness. The quarter-car model provides an idealized representation of vertical vehicle dynamics at a single wheel position. The sprung mass represents the inertia of the vehicle body supported by the suspension. The unsprung mass represents the inertia of the hardware between the suspension and the road, such as the tire, the wheel, the brakes, etc. The quarter-car model also includes elements that transmit forces to the two masses through stiffness and damping in the suspension and stiffness of the tire. In this model, the two masses only vibrate vertically, and the dynamics are forced by the profile through the tire.

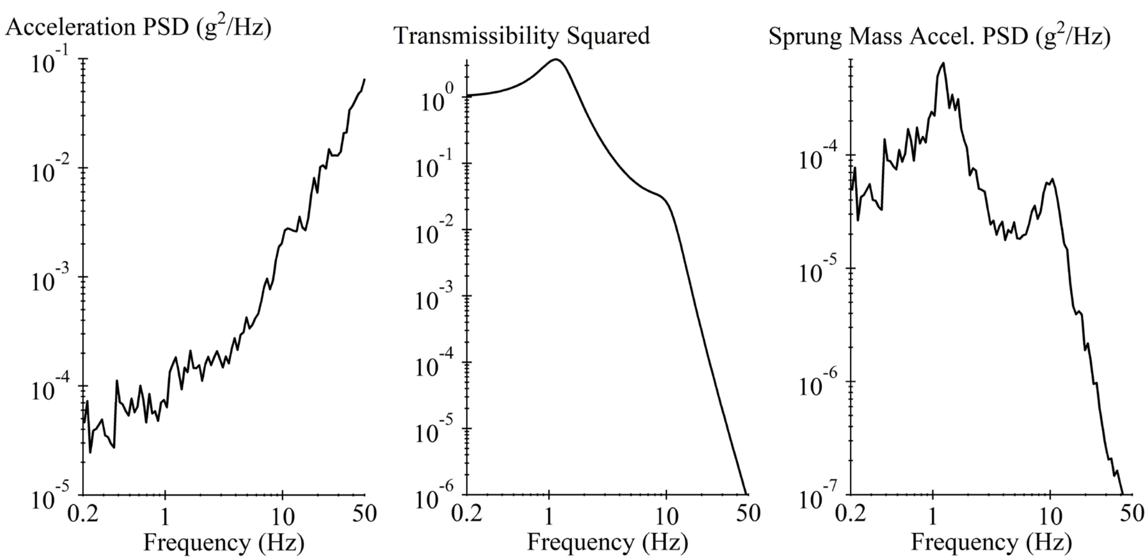
The next figure shows the transmissibility from the road to the sprung mass for the quarter-car model. Transmissibility works like the gain plots introduced in the discussion of filtering, and it defines the ratio of output amplitude to input amplitude for a sinusoid with a given frequency. The road-to-body transmissibility for the quarter-car model depends on the values of the input parameters used to characterize the vehicle, such as the masses, the stiffness values, and the suspension damping. The figure shows the transmissibility for a specific set of parameter values.



At very low frequency the body moves up and down in tandem with the road profile, and the suspension and tires deflect very little if at all. Within the range from 1 Hz to 3 Hz, the body of a typical passenger vehicle resonates on the suspension and amplifies the input from the road profile by a factor of 1.5 to 3.0. At higher frequencies the suspension absorbs the road inputs and isolates the body from the road. Within the range from about 10 Hz to 15 Hz, the wheel resonates and vibrates between the body and the road with movements larger than the input provided by the profile. Wheel vibration diminishes the isolation at the body somewhat in this frequency range, but only a fraction of that motion transmits through the suspension.

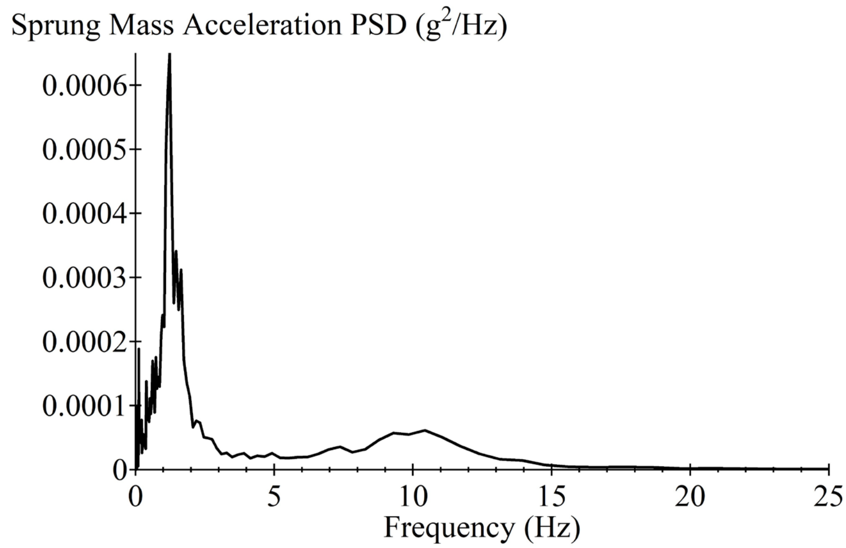
A vehicle behaves as a mechanical filter with a frequency response function.

Consider the transmissibility function described above for the quarter-car model. If it is applied to the PSD function of road profile acceleration, the output is vertical acceleration of the vehicle body. The figure below demonstrates this calculation. The output PSD at each frequency is the input PSD multiplied by the square of the transmissibility.



The dynamic properties of the idealized vehicle model cause it to amplify the input at about 1 Hz. The vehicle progressively attenuates inputs above 1 Hz, with the exception of an inflection near 10 Hz for unsprung mass resonance. A significant level of response appears in the output at 10 Hz, despite the relatively low transmissibility, because the acceleration input is so large. Notice, however, that the isolation of the suspension helps avoid peak responses in the range from 3 Hz to 8 Hz, because in this range the human body is sensitive to vertical movement.

The example above leads to an important result that may help explain the attention given to sinusoids, linear transforms, PSDs, and gain functions. The plot on the right side provides a way to examine the importance of each frequency range to acceleration on the vehicle body. The integral of the output PSD is equal to the variance of body acceleration, and the integral over any frequency range within the plot is equal to the contribution from that range. Consider the output PSD below, which is shown using linear scaling. In this format, contributions to variance are proportional to the area under the curve. The standard deviation, which is the square root of variance, over the entire frequency range 0.0294 g. The range from 0.5 Hz to 25 Hz accounts for 96 percent of the total, with 56 percent in the range from 0.5 Hz to 5 Hz and 40 percent in the range from 5 Hz to 25 Hz.



Are the dynamics of vehicle ride quality this simple?

No. The quarter-car model includes two very important modes of vibration common to most road vehicles: body bounce and axle hop. However, other motions, such as pitch and roll, also contribute to the gross vehicle dynamic response to road roughness. Vehicles also exhibit resonant responses to road roughness associated with frame flexure, tire structural vibration, and installation of large masses on flexible mounts (such as an engine or battery).

What Is Road Roughness?

By far, the main application for road profilers is to measure road roughness.

There is not a single, standard definition of road roughness.

Here is the American Society of Testing and Materials (ASTM) E867-23 definition of traveled surface roughness:

> *"The deviations of a surface from a true planar surface with characteristic dimensions that affect vehicle dynamics, ride quality, dynamic loads, and drainage, for example, longitudinal profile, transverse profile, and cross slope."*

This definition covers the factors that contribute to road roughness. However, it does not provide a quantitative definition or standard scale for roughness. It is also very broad, including qualities such as drainage and ride quality that are generally unrelated to each other.

The International Organization for Standardization (ISO) 13473 defines an international counterpart to roughness called "unevenness" as follows:

> *"deviation of a pavement surface from a true planar surface with the characteristic dimensions along the surface of 0,5 m to 50 m, corresponding to wavelengths with one-third-octave bands including the range 0,63 m to 50 m of centre wavelengths."*

ISO associates the quantitative definition for longitudinal unevenness with vehicle vibrations that affect ride comfort and road holding.

For the traveling public, the concept of roughness is simple:

> *"I know it when I feel it."*

Smoothness is a lack of roughness.

Some engineers prefer to consider smoothness as a more optimistic view of the road condition. Because smoothness is a lack of roughness, you would have to first quantify roughness and then transform the number to an alternative scale. The most common index in use in the U.S. is a scale that increases as roughness increases.

Some engineers think of roughness as the output of a specific device.

Roughness has been of interest as long as there have been public roads. Users of old methods for measuring roughness often think of the output as standard. The next section describes some of the instruments used in the past. Unfortunately, unless a roughness measure is based on profile, it cannot be reproduced.

Roughness involves variation in surface elevation that induces vibrations in traversing vehicles.

There are many kinds of vibration, ranging from sickening heaves due to long-wavelength roughness, to the rapid teeth-jarring impacts and irritating noise caused by short-wavelength roughness. The public considers any road feature that causes unwanted vehicle acceleration as roughness.

Roughness is defined over an interval of profile.

It is meaningless to talk about the roughness of a point. Instead, one must consider roughness as a summary of deviations that occur over an interval between two points.

There are many types of roughness.

Road users can identify various types of roughness by the form of unwanted vehicle vibrations that they cause. It is reasonable to compute more than one roughness index from a profile, if each index provides independent information about the state of the road. Not all types of roughness are unique. Many of the roughness indices that have been calculated from profiles are so strongly correlated that one is statistically sufficient. It is of dubious utility to compute two indices that tell essentially the same thing.

Different types of roughness are associated with different wavelengths.

For example, vehicle manufacturers are concerned with different aspects of ride quality, ranging from the heaving motions that mainly involve suspension movement, to audible noise involving acoustics of the body. These phenomena correspond to widely different wavelength ranges.

Roughness analyses can be compared on the basis of how they process a sinusoid. Most of the analyses filter out very long wavelengths and very short wavelengths.

Roughness is not as easily identified as a single dimensional property.

Length, weight, and other measures involve static physical properties of objects. Roughness involves at least two dimensions in a complex way: it involves variation of a profile height along its length. One cannot experience the roughness of a single profile point—it must be experienced over some length.

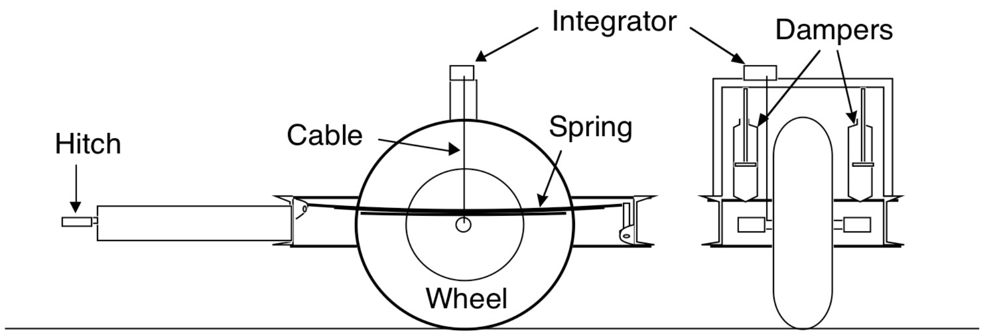
Roughness is analogous to a sound level for noise. Although air pressure is a static property, noise level involves changes in air pressure over time. Measurement of noise has been standardized to the extent that we can buy inexpensive sound meters that combine the measurement of air pressure over time with a mathematical analysis (implemented electronically) to process the variations and produce a single output level. There are several standard analysis methods to choose from that are selectable by a switch on the sound meter. The standards allow people who care about noise to communicate using measures with confidence that they will have the same meaning.

What Are Response-Type Systems?

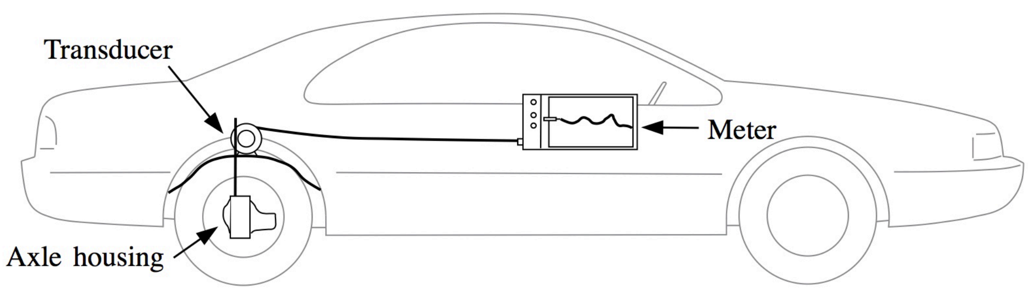
This book addresses profile measurement and interpretation. However, a review of legacy systems used to measure road roughness provides technical insight into the way road roughness measurement has evolved. In the past, many engineers understood the concept of roughness in the context of the behavior of older systems.

As early as the 1920s, highway engineers installed devices in cars to record suspension stroke as a measure of roughness. These were called road meters and had several generic names, including: response-type road roughness measuring systems (RTRRMS), response-type systems, and road meter systems. In these systems, the vehicle is a passenger car, a van, a light truck, or a special trailer. A road meter is a transducer that accumulates suspension motions. Road meters included commercial systems such as the Mays Ride Meter, the PCA meter, and the Cox meter, as well as various agency-built systems.

Nearly all of the road meter designs followed the concept of the Bureau of Public Roads (BPR) Roughometer, and they accumulated deflections of the vehicle suspension as it traveled down the road. The BPR Roughometer was a single-wheel trailer with a one-way clutch mechanism that accumulated the suspension stroke in one direction. (The total stroke is twice that value.) See the drawing below.



Road meters were more commonly used in passenger cars, as shown in the next figure. The output expressed roughness as “inches” of accumulated suspension stroke, divided by the distance traveled. The measure was usually reported with engineering units such as inches/mi, although in some cases arbitrary units for suspension movement were used based on the instrumentation hardware, e.g., “counts/mi.” This measure of vehicle response is similar in its frequency content to the accelerations on the vehicle body, so it is correlated to ride vibration.



Measures from response-type systems are subject to every variable that influences the dynamic response of the host vehicle.

Even when the host vehicle type is standardized, differences remain between individual vehicles that one might think are identical. To further compound the problem, the response properties of the individual vehicles change with time, because of component wear, sensitivity to temperature, etc. The fact that response-type systems depend on the dynamics of the host vehicle has two unwanted effects:

1. Response-type roughness measuring methods are not stable with time. Measurements made today with response-type systems cannot be compared on the same scale with confidence to those made in past road surveys.

2. Response-type roughness measurements are not transportable. Road meter measurements made by one system are seldom reproducible by another.

These problems existed in part because road meters were typically inventions devised to be inexpensive, rugged, and easy to use. A rigorous understanding of how they function together with a vehicle did not exist until 1980, when the variables were studied in a research project funded by the National Cooperative Highway Research Program (NCHRP). See:

Gillespie, T.D., Sayers, M.W., and Segel, L., "Calibration of Response-Type Road Roughness Measuring Systems." *National Cooperative Highway Research Program Report 228* (1980) 81 p.

NCHRP Report 228 describes how response-type systems work.

A second source of difficulty involving response-type systems in prior generations had been the lack of a standard roughness scale. A standard, reproducible, time-stable roughness scale helped overcome some of the problems inherent in a response-type system. Using the standard roughness scale, the output of each response-type system was rescaled by calibration. The lack of a standard scale was at first not seen as a serious problem by many of the users of RTRRMSs. Roughness data for a city, county, or state could have arbitrary units, so long as the internal historical database remained consistent. However, even the repeatability of the instruments within a single jurisdiction was a problem.

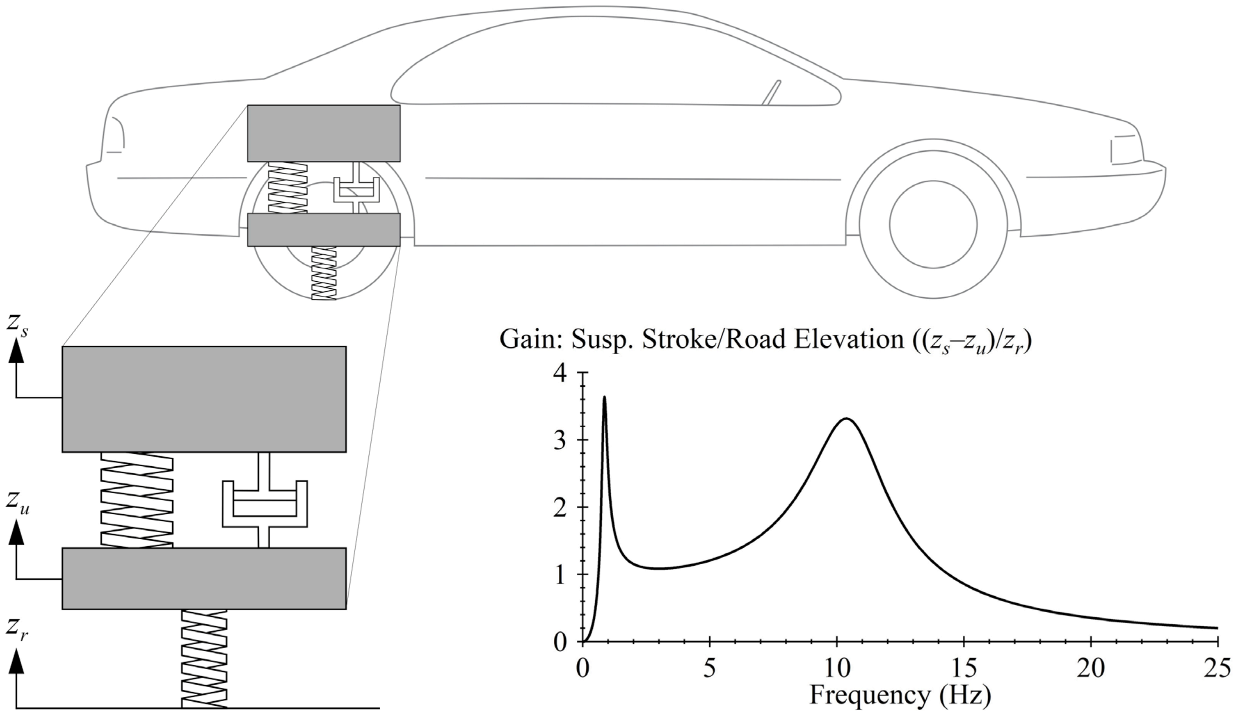
The "inches/mile" measures from response-type systems are useful.

Although there are problems involving reproducibility and portability of data collected with response-type systems, one reason that they had been so popular for five decades preceding the proliferation of inertial profilers is that they work. Engineers viewed the measurements produced by response-type systems as matching their experience for determining pavement ride quality in a meaningful way. If they provided results that were reproducible between vehicles, at different travel speeds, and over time, it is possible that there would not be as much interest in profiling methods today.

Response-type systems filter the profile through vehicle dynamics.

Automobiles and other vehicles designed for the highway use suspensions and pneumatic tires to isolate the drivers and cargo from the high-amplitude accelerations imposed by roughness of the road surface. The dynamics of the vehicle amplify the input at some frequencies and attenuate it at others. The discussion of ride quality in a previous section of this book demonstrated the filtering effect analytically using body acceleration.

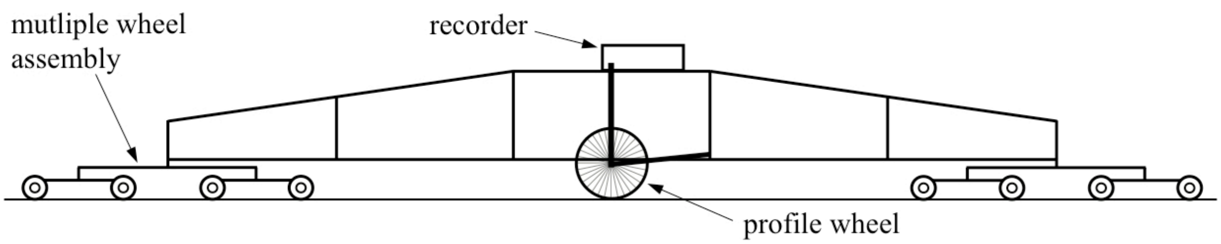
The frequency response for road meter output resembles the response for vertical acceleration of the vehicle body. In a road meter, the response is characterized by relative vertical velocity between the body and the axle. The figure below shows the gain for road meter response to road profile for the quarter-car model using estimates of vehicle properties from a specific Mays Ride Meter used in the 1970s. Body resonance (1 Hz to 3 Hz) and the axle resonance (roughly 10 Hz to 15 Hz) motions affect both passenger acceleration and road meter output, with some isolation near 3 Hz to 5 Hz and for frequencies above 15 Hz.



The response of the vehicle, and its "filtering action" on road profiles, depends on time-based dynamics. The frequency response of a car is approximately independent of speed, when the frequency is defined in units of Hz (e.g., cycle/sec). Mapping the response to study the sensitivity to wavelengths in the road profile requires knowledge of the travel speed.

What Are Straightedge-Based Devices?

The drawing below, which resembles a drawing that appeared in the 1978 version of California Test 526, shows the geometry of a straightedge-based device called the California profilograph.



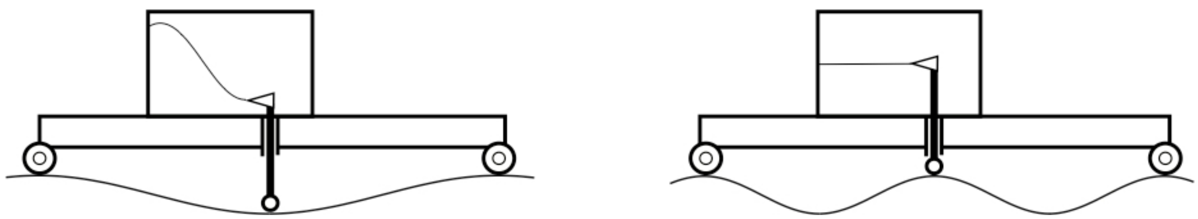
Straightedge-based devices are hand-propelled systems that measure pavement roughness. They usually consist of a body made of a long beam or truss supported by a set of wheels at both ends. A measurement wheel follows the contours of the pavement and is free to move vertically relative to the body of the device. The measurement wheel typically appears in the longitudinal center between the supporting wheel sets, as shown in the figure above. These devices capture a record of the vertical movement of the measurement wheel relative to the body of the device versus distance travelled along a pavement section.

The first published example of using a straightedge-based device to measure pavement roughness appeared in 1907. Straightedge-based devices, including several versions of the profilograph, were routinely used for quality control and quality assurance of new pavements for many decades. Pavement engineers used the output of profilographs to measure the overall roughness of pavement sections and to identify large deviations in profile for correction. Simpler devices, such as the rolling straightedge, were used to locate bumps and dips on the pavement.

Straightedge-based devices filter the profile through their geometry.

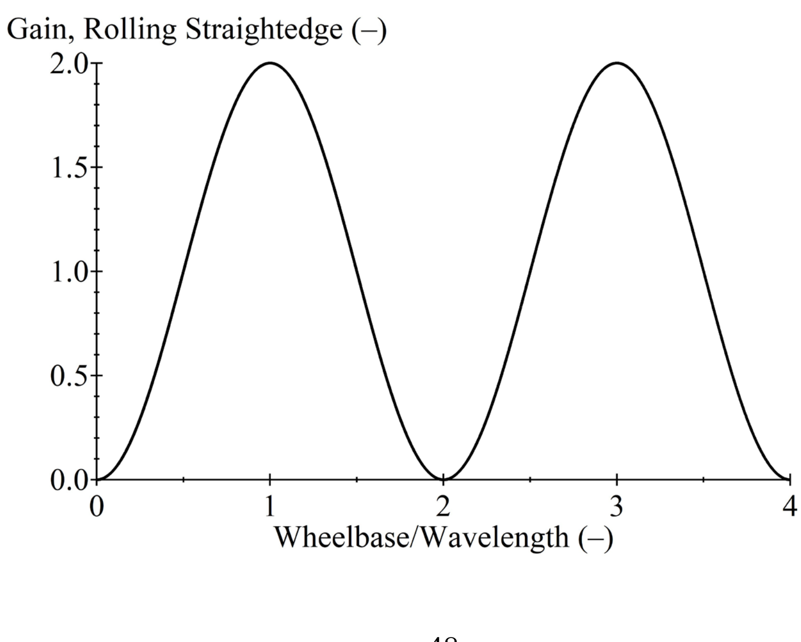
Most straightedge-based devices behave like linear mechanical filters. Given a sinusoid as an input, they produce a sinusoid with the same wavelength, which is then processed by the on-board instrumentation to produce a summary index.

The schematics below provide examples of rolling straightedges traveling over sinusoidal profiles. Rolling straightedges measure roughness using the same principle as a profilograph, but they are typically supported by wheels in only one position at each end and are not as long as profilographs. Each wheel position usually includes two wheels side-by-side.



The left schematic shows the rolling straightedge on a sinusoid with a wavelength equal to the wheelbase. In the position shown the device records a downward deviation equal to twice the amplitude of the sine wave. When the rolling straightedge moves forward by half the wheelbase, it detects an upward deviation equal to twice the amplitude—this is a gain of two. The right schematic shows a case where the wheelbase of the rolling straightedge is twice the wavelength of the sinusoid. In this situation, the measuring wheel and the supporting wheels move up and down together as the device rolls along, and the measured trace is flat—this is a gain of zero.

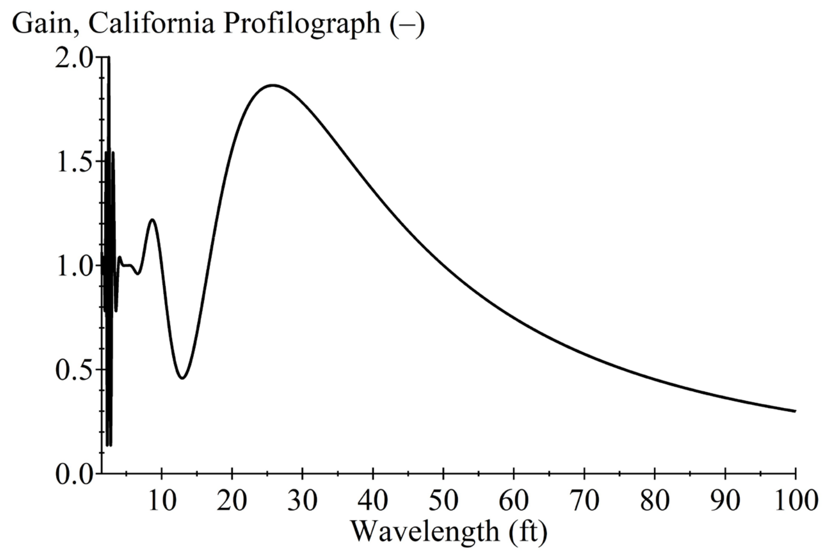
The figure below shows the response gain for a rolling straightedge. The plot shows spatial frequency in terms of wheelbase. Typical wheelbases for modern rolling straightedges are 10 ft and 12 ft, although many varieties exist.



Straightedge-based devices behave as band-pass or high-pass filters.

Long-wavelength undulations dominate the elevation values in a typical true profile. Most straightedge-based devices remove most of the long-wavelength content but are sensitive to content over at least some of the wavelength range that affects vehicle ride quality.

The gain response for the rolling straightedge at long wavelengths corresponds to low values on the horizontal axis of the plot above. The rolling straightedge reduces any content with a wavelength greater than four times the wheelbase by at least 70 percent. The plot below shows the gain response for a California profilograph. The plot shows responses as a function of wavelength. Like the rolling straightedge, the profilograph rejects content at wavelengths that are much longer than its wheelbase.



The profilograph is a physical system that is similar in concept to an anti-smoothing moving average. The supporting wheels establish the "average," because they determine the reference height of the device at the position of the measurement wheel. The measurement wheel records changes in the height of the road surface relative to the reference.

The gain response above corresponds to the California profilograph shown at the start of this section. The wheelbase of this device, which is the longitudinal spacing between the centers of the supporting wheel sets, is 25 ft. Note that, like the rolling straightedge, the profilograph amplifies content at wavelengths near the wheelbase. The details of the gain response at wavelengths below the wheelbase depend on the arrangement of the supporting wheels. The wheel sets include wheels on the left and right side of the device, which are not all visible on the schematic. The gain function was derived analytically and does not account for side-to-side variations in the profile, or the enveloping action of short asperities in the profile by the supporting wheels and the measurement wheel. For wavelengths close to and below the contact length of the measurement wheel, the actual gain response is reduced.

A great variety of roughness measurement devices preceded the introduction of inertial profilers.

The literature describes many straightedge-based devices. The devices include several versions of the profilograph with other layouts and alternative supporting wheel geometry.

RTRRMSs also evolved through several iterations, and some of them used direct measurement of vertical acceleration on the host vehicle body. The following reference provides an excellent inventory of historical road roughness measurement devices.

Hveem, F.N., "Devices for Recording and Evaluating Pavement Roughness." *Highway Research Board Bulletin 264* (1960) pp. 1-26.

Straightedge-based devices and response-type systems exemplify a contrast to valid measurement of profile. A valid profiler measures the features of interest from the true profile without distortion caused by device geometry or by its dynamic response properties. Valid profile measurement includes the requirement that gain for measured profile versus the true profile is very close to 1 over the wavelength range of interest for the intended purpose.

What Is a Profile-Based Roughness Index?

A profile is a signal with a sequence of relative elevation values recorded at a constant distance interval. Typical longitudinal profiles contain tens of thousands of elevation values per mile. Each pass by a profiler may produce a profile measurement for one, two, three, or many slices of the road surface. In a short time, profilers accumulate millions of numbers to store and analyze. A profile-based roughness index (i.e., a roughness index) reduces those millions of numbers into a few summary numbers that (hopefully) provide useful information.

A profile-based roughness index is a summary number calculated from the many numbers that make up a profile.

Details of the calculation determine the significance and meaning of the roughness index. The number might relate to the output of a mathematical vehicle model, identify locations for corrective action, or express an abstract concept of roughness. Or it might not be linked to anything at all.

The calculated index is only valid if the profile data is valid.

If the profiler is not functioning correctly, or is not suited for the index of interest, it is not possible to get the same numerical value that would be obtained from the true profile. Not every profiler can measure every possible roughness index. The accuracy of a calculated index is limited by errors in the measured profile.

There is a true value for any given index.

The true value of a profile-based roughness index or any profile-based statistic is the value that would be obtained by calculating it from true profile.

A profile-based roughness index is portable and reproducible.

If an index can be calculated from the true profile, then any valid profiler can measure the property it represents. Thus, a profile-based roughness index is portable between different profiler types, so long as they are valid for that index. A profiler is valid for measurement of a given index if it measures the aspects of the true profile that affect that index.

Some profile analyses are not as portable as others. For example, if an analysis requires a specific sampling or recording interval, a profiler is valid for its calculation only if the interval matches.

A profile-based roughness index is stable with time.

Because the concept of a true profile has the same meaning from year to year, it follows that a mathematical transformation of the true profile is also stable with time.

Most profile-based roughness indices in use are calculated with a basic four-step method.

The mathematical transforms used to compute almost any roughness index from profile can be organized into four steps. Details of the calculations done in each step define the index:

1. Select the profiles needed to capture the surface conditions of interest.
   Most profile-based roughness indices are calculated from a single profile. If a profiler measures the profile of the left and right wheel tracks in a lane, then a separate index value can be calculated for each. However, some indices require the use of a profile from the left and right wheel track.
2. Filter the profile.
   Most profile-based roughness indices that are now in use involve at least one filter to remove content in wavelength ranges that are not of interest and emphasize content in wavelengths that are of interest. Some analyses involve several filters applied in sequence.
3. Accumulate a summary number.
   The sequence of transformed numbers produced in the filtering stage must be reduced to a single value. Common methods for doing this include accumulation of the absolute values of the numbers or accumulation of the squared values of the numbers. The result is a single summary number.
4. Scale the summary number.
   The last step is to convert the summary number to an appropriate scale, which almost always involves dividing by the number of profile points or the length of the profile to normalize the roughness by the length covered. For example, many historical roughness indices have had units of inches/mile. A scale factor may be used to obtain standard engineering units. A transformation equation may be used to convert from a scale in engineering units to an arbitrary scale.

Many indices can be calculated from a single profile.

An advantage of using profilers to determine roughness, besides the portability, is flexibility. One can obtain several indices from the same profile. Each index may potentially describe a different characteristic of the profile. For example, this book presents background information about the IRI, which represents a broad set of vehicle response to road roughness, and RN, which provides an estimate of user perceived ride quality.

What Is the International Roughness Index?

Almost every automated road profiling system includes software to calculate a statistic called the IRI. Through the RPUG, the Long-Term Pavement Performance (LTPP) Study, and the

FHWA Highway Performance Monitoring System (HPMS), profiler users have shared experiences measuring and interpreting the IRI. IRI measurements from different states are largely comparable. Even IRI measurements from different countries can be compared directly.

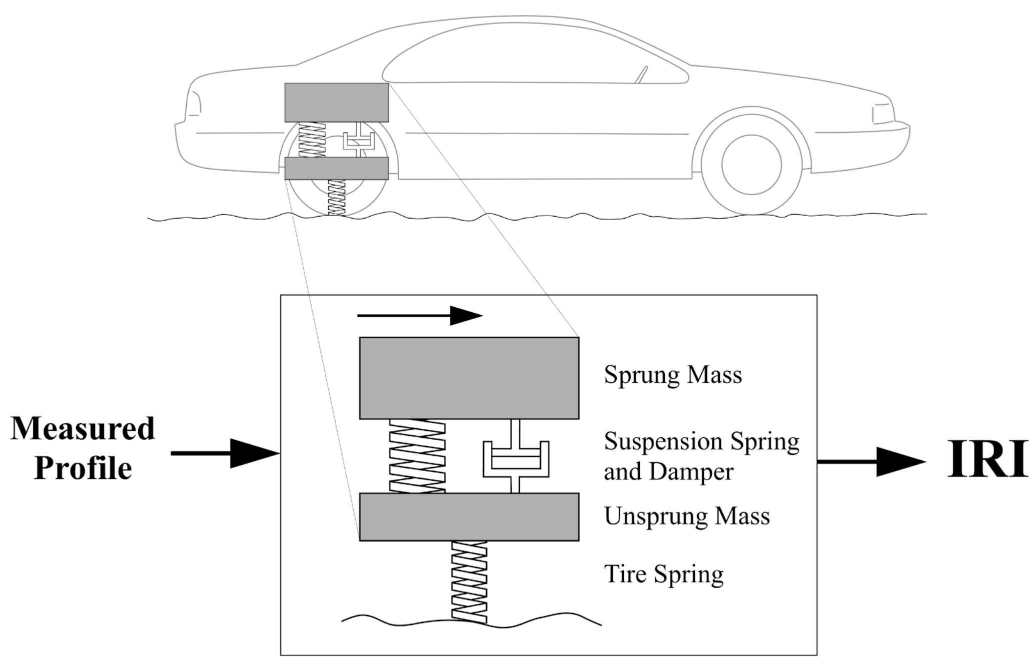
Background

The IRI is a continuation of the inches/mile roughness statistic used by response-type systems.

Although the inches/mile measure from response-type systems had been popular since the 1940s, it was not possible to obtain the same values from different vehicles, or even from the same vehicle over time. The IRI was originally developed in an NCHRP research project as a correlation and calibration standard for casting roughness measurements from response-type systems onto a consistent scale.

The IRI is a computer-based "virtual response-type system."

To calibrate the response-type systems, an ideal system was defined for the computer based on a quarter-car model. The quarter-car model accepted a measured profile as input and produced an index that was a mathematical function of the longitudinal profile.



To correlate well with response-type systems, the quarter-car model was customized to simulate a perfect road meter. The model accumulates suspension motion and divides by the distance traveled to give an index with units of slope (e.g., inches/mi). The quarter-car simulation in the IRI calculation algorithm uses a specific set of parameters for vehicle properties that were selected to obtain maximum correlation to the output of response-type systems. With this parameter set, the simulation is called the Golden Car. The name was intended to convey the status of the computer-based representation as a stable calibration reference, as an analogy to a gold bar of standard weight that might be stored in a vault by a standardization body.

Research by NCHRP and The World Bank established the relevance of the IRI.

The NCHRP research that led to the development of the IRI built on the 50 years of experience accumulated by the states and others using inches/mile roughness indices. See:

Gillespie, T.D., Sayers, M.W., and Segel, L., "Calibration of Response-Type Road Roughness Measuring Systems." *National Cooperative Highway Research Program Report 228* (1980) 81 p.

Subsequently, research supported by the World Bank examined the trade-offs between the cost of investing in road quality and costs to users on poor quality roads. In the World Bank research, the use of the Golden Car as a calibration reference was found to be the most suitable for defining an international roughness scale to help guide investment decisions in various countries. The World Bank finalized the IRI algorithm and published procedures for calculating it from measured profile. See:

Sayers, M.W., Gillespie, T.D., and Patterson, W.D.O., "Guidelines for Conducting and Calibrating Road Roughness Measurements." *World Bank Technical Paper Number 46* (1986) 87 p.

The IRI is reproducible, portable, and stable with time.

The IRI was the first widely used profile-based roughness index where the analysis method was intended to work with different types of profilers. The IRI is defined as a property of the true profile; therefore, it can be measured with any valid profiler. The analysis equations were developed and tested to minimize the effects of some profile measurement parameters like recording interval.

The IRI is a general pavement condition indicator.

The IRI summarizes the roughness qualities that affect vehicle response and is most appropriate when a roughness index is desired that relates to overall vehicle operating cost, overall ride quality, overall surface condition, and dynamic wheel loads. In this context, dynamic wheel loads include fluctuations in vertical forces applied to the road from heavy trucks, which potentially affects pavement wear, and fluctuations in vertical forces at the interface between the tire and the road surface on all vehicles, which is detrimental to braking and cornering.

Properties of the IRI Analysis

The IRI scale starts at zero for a perfectly smooth road and increases in proportion to roughness.

If all elevation values in a measured profile are increased by some percentage, the IRI increases by the same percentage. There is no theoretical upper limit to the IRI scale. However, as the IRI increases beyond very high levels, the maximum traffic speed for avoiding an uncomfortable ride experience or wear of vehicles and their cargo decreases. For example, pavements with IRI values above 500 inches/mi are nearly impassable except at reduced speeds.

In 2017, the U.S. Department of Transportation (DOT) published a regulation that defined thresholds for pavement condition ratings based on the IRI within the Transportation Performance Management (TPM) system. The table below defines the IRI range for various performance categories used by the FHWA before and after the TPM system was finalized. In both eras, the poor category was considered “less than acceptable.”

Rating	TPM IRI Range (inches/mi)	Pre-TPM IRI Range Interstate (inches/mi)	Pre-TPM IRI Range non-Interstate (inches/mi)
Very Good	–	< 60	< 60
Good	< 95	60-95	60-95
Fair	95-170	95-120	95-170
Mediocre	–	120-170	170-220
Poor	> 170	> 170	> 220

The IRI describes road roughness that causes vehicle vibrations.

The quarter-car model used in the IRI algorithm is just what its name implies: a model of one corner (a quarter) of a car. The model includes one tire, represented with a vertical spring, the mass of the hardware below the suspension, a suspension spring and a damper, and the mass of the body supported by the suspension for that tire.

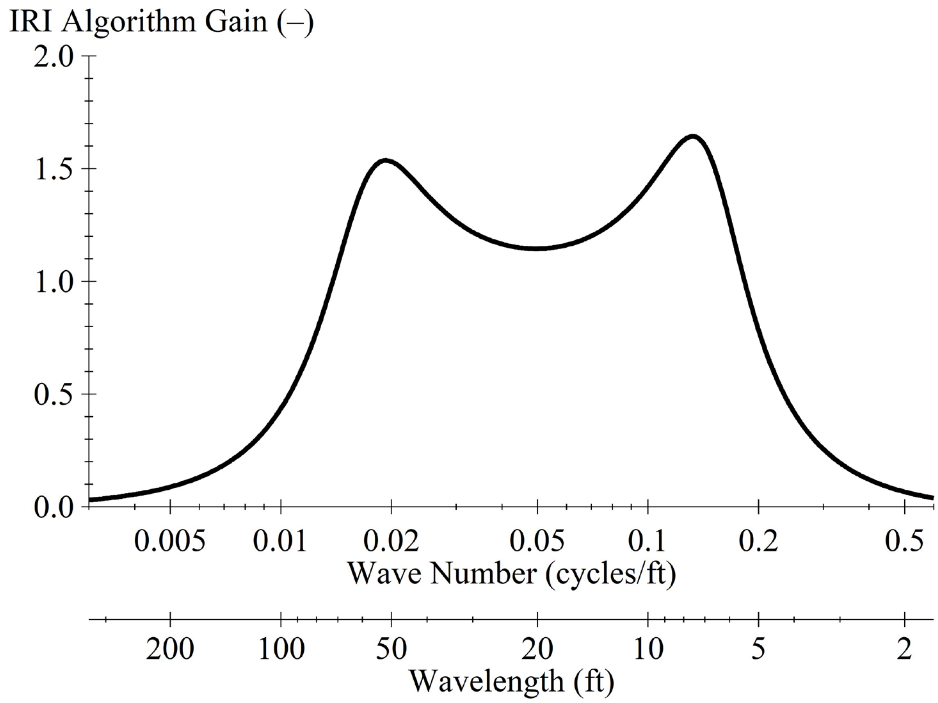
The response of the IRI to sinusoids is intentionally very similar to measured physical responses of highway vehicles. It was developed to match the responses of passenger cars, but subsequent research has shown good correlation with light trucks and heavy trucks. The IRI has become recognized as a general-purpose roughness index that is strongly correlated to most kinds of vehicle response that are of interest. Specifically, IRI is very highly correlated to three vehicle response variables:

1. suspension stroke (for durability and energy loss)
2. vertical passenger acceleration (for ride quality)
3. tire load (for vehicle controllability and safety)

The IRI is not related to all vehicle response variables. For example, it does not correlate well with either axle acceleration or fluctuations in vertical passenger position.

IRI is influenced most by wavelengths ranging from 3.3 ft to 126 ft.

The figure below shows the wave-number response (i.e., spatial frequency response) of the IRI algorithm in terms of sensitivity to slope input. The figure includes a wavelength axis for ease of interpretation. For any sinusoid within the input profile, the amplitude of the corresponding sinusoid in the output of the IRI algorithm is the amplitude of the input’s slope multiplied by the gain shown in the figure. The figure shows what features within the profile affect the output of the IRI algorithm, given their wavelength. For example, the IRI algorithm greatly reduces the influence of aspects of the profile with wavelengths less than 2.32 ft and more than 189.25 ft, because the gain is 0.1 or less. The gain is at 0.25 or above in the wavelength range from 3.30 ft to 125.76 ft.



The parameters used in the IRI calculation were selected to maximize relevance to as many highway vehicles as possible. To do this, the Golden-Car model uses a high value for the suspension's damping coefficient. The higher damping level flattens out the peaks in the frequency response relative to a typical vehicle and creates a more uniform gain response across the wavelength range that affects typical vehicles. Flattening the frequency response prevented the model from tuning in to roughness at specific wavelengths that affect some vehicles much more than others. In contrast, the frequency response shown in the discussion of response-type systems includes very high response isolated at two frequencies that correspond to body resonance and axle resonance in a particular vehicle but fails to apply adequate weighting to frequencies that strongly affect other vehicles.

Definition of the IRI

The above descriptions of the IRI background and properties are intended to describe what the IRI computer software is intended to simulate, and how you can interpret the IRI scale. However, the IRI is rigorously defined as a specific mathematical transform of a true profile. The specific steps taken in the computer program to compute IRI are listed below.

The IRI is calculated for a single profile.

If a profiler measures several profiles simultaneously, then one can get the IRI for each profile. The IRI standard does not specify how to locate the line on a road that defines the profile. Any line on the road has an associated IRI value. The standard does not specify how to combine IRI values for different profiles taken along the same segment. They can be averaged, but the result is not IRI—it is the average of more than one IRI value.

The profile is filtered with a moving average with a 250-mm (9.85-inch) base length.

The moving average is a low-pass filter that smoothens the profile. The moving average is included in the IRI algorithm to simulate the ability of a tire to envelop small-scale asperities within the contact patch. The moving average also reduces the sensitivity of the IRI algorithm to variations in the sampling and recording intervals used by different profilers. Note that the wave-number response plot for the IRI algorithm provided above includes the combined effect of the moving average and the Golden-Car model.

The 250-mm (9.85-inch) moving average filter should be omitted for profiles obtained with some systems.

The computer program used to calculate the IRI does not apply the filter unless the profile's recording interval is shorter than 6.56 inches. If the recording interval is less than 6.56 inches, the moving average should be omitted if the profile has already been filtered by an anti-aliasing filter that attenuates wavelengths shorter than 1.5 ft. For example, K.J. Law Profilometers from a past era detected elevation values at intervals of 0.98 inches, applied a moving average filter with a base length of 11.81 inches, and recorded the profile at 5.91-inch intervals. In this situation, the effect of the filtering applied prior to recording the profile was very similar to the effect of the 250-mm (9.85-inch) moving average from the IRI algorithm. Consequently, the 250-mm (9.85-inch) moving average was not applied to calculate the IRI, because it was redundant.

The profile is further filtered with the Golden-Car simulation.

The parameters used in the Golden-Car simulation are specified as part of the IRI definition, and the simulated travel speed is specified as 80 km/hr (49.7 mi/hr). The Golden-Car parameters are:

$$\frac{k_s}{m_s} = 63.3\ (1/\text{sec}^2) \qquad \frac{k_t}{m_s} = 653\ (1/\text{sec}^2) \qquad \frac{c}{m_s} = 6\quad (1/\text{sec}) \qquad \frac{m_u}{m_s} = 0.15\ (-)$$

Where k_s is the suspension spring rate, m_s is the sprung mass, k_t is the tire spring rate, c is the suspension damping rate, and m_u is the unsprung mass. The output of the filter represents motion across the suspension of the simulated quarter car.

The filtered profile is accumulated by summing absolute values and then is divided by the profile length.

The resulting IRI statistic has units of slope. As a user, you can express the slope in any appropriate units. The most common choices are inches/mi (multiply slope by 63,360) and m/km (multiply slope by 1000).

Details of the IRI are handled in computer software.

The above summary illustrates how the IRI fits the earlier description of a profile-based roughness index. The analysis is applied to a single profile, the profile is filtered (twice), the filtered result is accumulated, and the accumulated value is cast onto a standard by dividing by the length of the profile.

ASTM E1926-08 provides a standard for calculating the IRI from profile. A reference for more information about the IRI calculation method is:

Sayers, M.W., "On the Calculation of International Roughness Index from Longitudinal Road Profile." *Transportation Research Record 1501* (1995) pp. 1-12.

What Are Mean Roughness Index and Half-Car Roughness Index?

Given two profiles, one for the left wheel track and one for the right wheel track, there are two ways users of profilers have processed them with the IRI algorithm to produce a single index for both profiles. These indices are called the Mean Roughness Index (MRI) and Half-car Roughness Index (HRI).

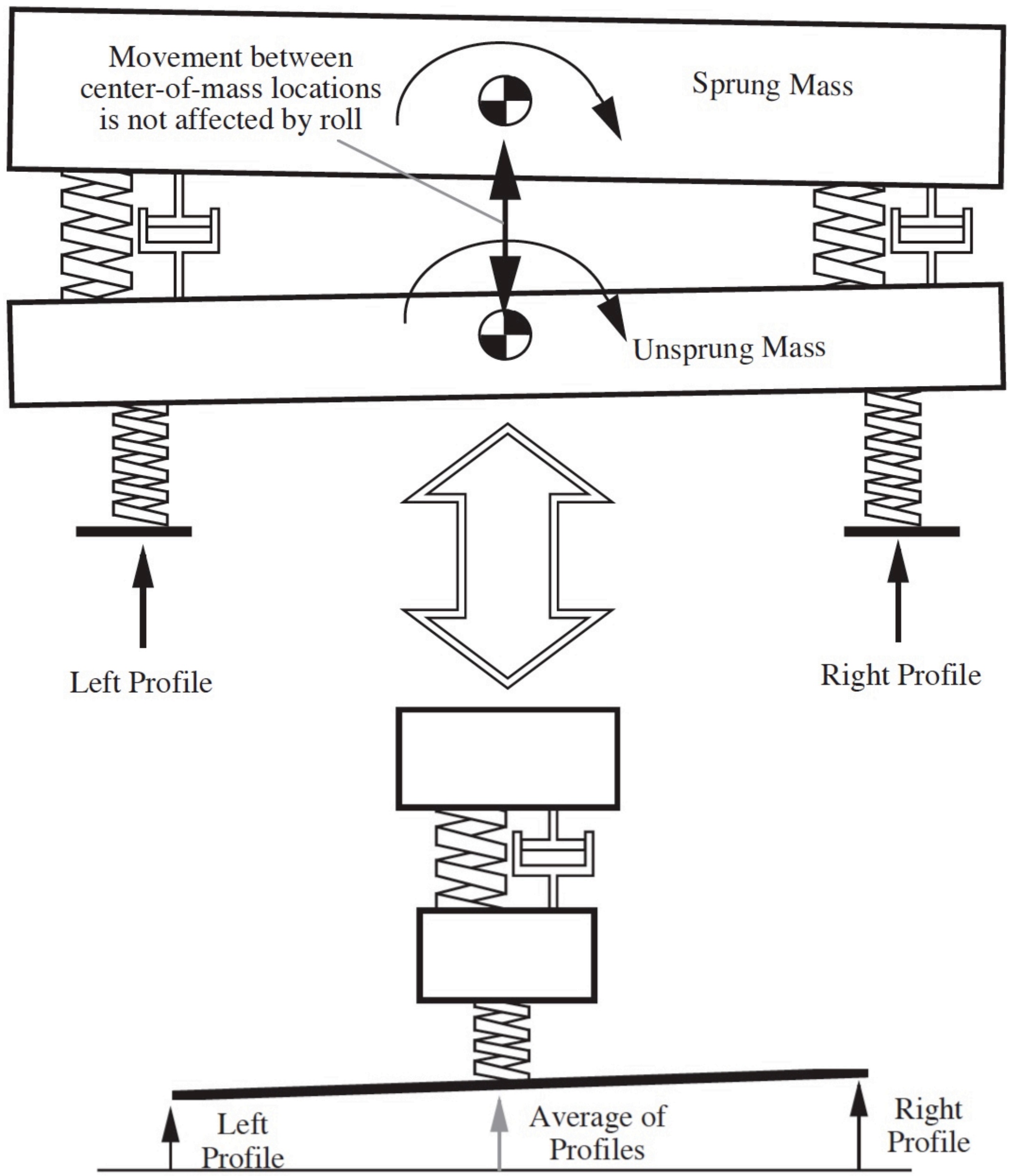
MRI is the average IRI from the left and right wheel track.

For profilers that measure two wheel tracks, using the MRI reduces the volume of summary index values by half. However, using the MRI alone conceals cases where much more roughness appears on one side of a lane than another.

HRI is the IRI algorithm applied to the average of two profiles.

To calculate the HRI, collapse the profiles from the left and right wheel tracks into a single profile, in which each point is the average of the corresponding elevation values from both sides. Then apply the IRI algorithm to the averaged profile.

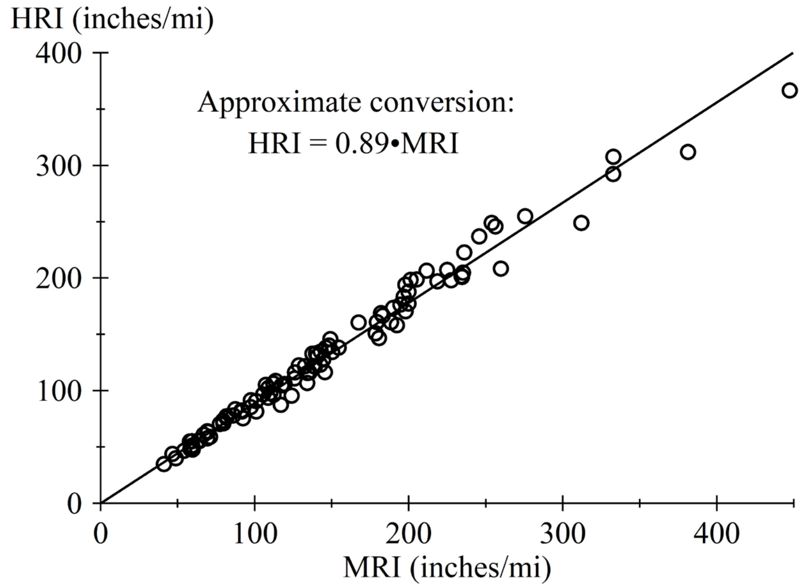
The diagram below illustrates the effect of averaging the two profiles before applying the IRI algorithm. The result is equivalent to simulating relative motion between the sprung mass and unsprung mass in the center of a solid axle. Some engineers used the HRI early in the proliferation of the IRI algorithm, in part because the half-car model represented the way road meters were installed in their host vehicles. As a result, using the HRI helped to maintain a strong correlation to their historical databases of roughness measurements.



1/2 Car = 1/4 Car using averaged
profile input

The HRI is only sensitive to aspects of roughness that are consistent in both wheel tracks. Consider a pair of sinusoidal profiles. If both sides receive the same sinusoid, in phase, then both sides of the vehicle bounce in response. It does not roll at all. However, if the sinusoids are out of phase, such that the left side goes up when the right side goes down and vice versa, the vehicle rolls but does not bounce. With a road meter installed at the center of an axle, it senses bounce but not roll. Real roads have roughness that causes a mixture of bounce and roll. Roughness that causes bounce contributes to the HRI, but roughness that causes roll is eliminated when the two profiles are averaged. Consequently, the HRI must be less than or equal to the MRI for the same two profiles.

Profile data taken in some research projects in the 1980s were analyzed both ways. The HRI values, calculated from the average of the left and right profiles, were compared to the MRI. The correlation between the HRI and MRI statistics was high, indicating that little or no additional information is provided by the HRI for most applications. However, the conversion equation shown in the figure is not valid for all conditions. The difference between HRI and MRI depends on the inconsistency between profiles from the left and right side of the lane.



The HRI requires two perfectly synchronized profiles as input. For devices that measure profiles in two wheel tracks simultaneously, the two are properly synchronized. However, for profilers that measure only one wheel track per pass, it would be difficult, and time consuming, to measure the left and right wheel tracks in separate passes and then align the two profiles as needed for the analysis. Practically speaking, the HRI analysis can only be used for systems that measure the profiles of two wheel tracks simultaneously.

For more information about the differences between HRI and IRI, see:

Sayers, M.W., "Two Quarter-Car Models for Defining Road Roughness: IRI and HRI." *Transportation Research Record 1215* (1989) pp. 165-172.

What Are Panel Ratings?

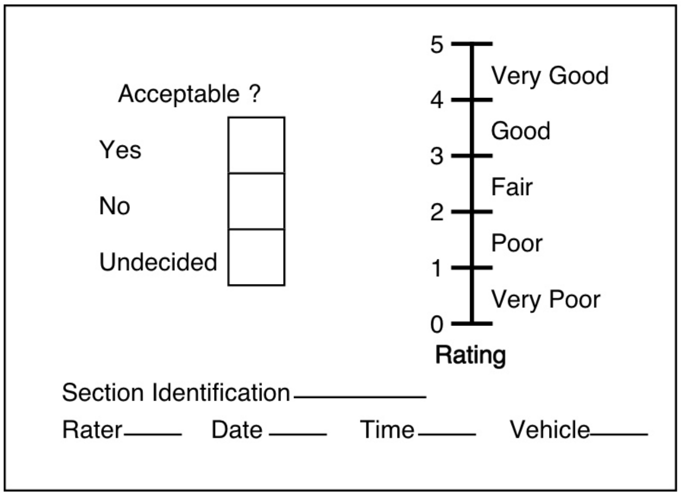
In 1960, F.N. Hveem noted, "Ever since roads and highways have been constructed, the people who use them have been keenly aware of the relative degrees of comfort or discomfort experienced in traveling." Long before high-speed profiling technology existed, engineers attempted to estimate the general opinion of the traveling public of specific roadways. The most direct method is to drive people over sections of road and ask them what they think.

Panel ratings are subjective.

Ratings from people reflect their opinions and are subjective. A person's expectations, prior experience, standards, beliefs, and mood affect the ratings they provide. In contrast, measures obtained from analysis of profile data are considered objective.

Subjective rating scales for roads usually go from 0 to 5.

The next figure shows a rating form in which a person rates a road on a scale of 0 to 5. A 0-to-5 scale was used for a large-scale road test conducted by American Association of State Highway Officials (AASHO) in the 1950s, in which roads were subjected to mixed traffic and researchers tracked the condition of the pavement. A panel of pavement experts evaluated the condition of the test pavements based on close inspection, the experience of driving over them, and the use of measures taken from several instruments in use at the time.



Ratings from the original AASHO test were called Present Serviceability Rating (PSR).

The ratings from the panel of experts were processed to assign a single number to each pavement that represented its serviceability, defined as "…the ability of a specific section of pavement to serve high-speed, high-volume, mixed (truck and automobile) traffic in its existing condition."

Carey Jr., W.N. and Irick, P.E., "The Pavement Serviceability-Performance Concept." *Highway Research Board Bulletin 250* (1960) pp. 40-58.

The summary number was called PSR. The researchers also asked people who were not engineers to rate the pavements. Similar results were obtained.

The current meaning of PSR is not standard. Some engineers consider PSR to be a one-of-a-kind measure that applied only for the tests done in the 1950s. Since the numbers were based on human opinion at the time, there is no way to confirm or deny the relation of ratings taken today with the original PSR scale. With this concept, the scale can no longer be used. Other engineers use the name PSR to refer to any study in which ratings are taken for roads on a 0-to-5 scale.

Predictions of PSR were called Present Serviceability Index (PSI).

In addition to the rankings obtained from the panel of raters in the original AASHO tests, several measures were taken of surface roughness and distress. Using the measurements, PSR could be estimated using an equation obtained from statistical analyses of the data. The estimate of the PSR was called the PSI. Typically, the PSI was computed by using an equation where the inputs to the equation to predict PSI were slope variance (a measure of roughness), rut depth, and area of cracking and patching on the pavement.

Statistical processing is used to calculate mean panel rating (MPR).

Past research has shown that opinions of a single person tend to be unreliable, relative to objective measures, and also relative to opinions of other persons. However, ratings averaged over a group of people, called a panel, offer consistent results when proper statistical methods are followed. Panel rating experiments are designed to estimate the opinion of "the public" from the small group composing the panel. A typical panel size is about 30 people, but reasonable results can be obtained with smaller groups.

After statistical processing, the results yield a single rating for the panel, typically called MPR. In most studies, ratings are re-scaled to remove bias and systematic errors (error of leniency, central tendency, etc.) before they are averaged. Thus, the MPR is not necessarily the mean value of the original ratings of the panel members.

Subjective ratings depend on questions and instructions.

Subjective panel ratings depend strongly on the instructions given to the members of the panel to define what physical property or quality is being judged. The instructions must "train" the rater. Yet, in a research program, the physical properties are not fully known—that may be the point of the research. The NCHRP sponsored two research projects in the 1980s to develop a methodology for obtaining valid ratings, resulting in the concept of ride number described in the next section. However, even today, procedures are not standard for obtaining panel ratings.

MPRs are not practical for network use.

There are two problems with using MPR data directly for evaluating the state of a road network:

1. The rating scale is not a measure of road condition that is stable with time. For example, roads considered "good" by a panel 50 years ago might be considered something else today.
2. It is expensive to obtain panel ratings due to the number of people required, and the need to transport them to the roads being rated.

What Is Ride Number?

Ride Number (RN) is a profile-based roughness index intended to estimate the opinion of the traveling public of the ride quality of roads on a scale like PSI.

Background

The direct collection of subjective opinions in the form of MPR is expensive and provides no guarantee of continuity from year to year.

RN is the result of NCHRP research in the 1980s.

The NCHRP sponsored two research projects in the 1980s that investigated the effect of road surface roughness on ride comfort, as described in:

Janoff, M.S., Nick, J.B., Davit, P.S., and Hayhoe, G.F., "Pavement Roughness and Rideability." *National Cooperative Highway Research Program Report 275* (1985) 69 p.

Janoff, M.S., “Pavement Roughness and Rideability Field Evaluation.” *National Cooperative Highway Research Program Report 308* (1988) 54 p.

One objective of the research projects was to link features in measured road profiles to subjective opinion about the ride quality of the road from members of the public. During two studies, spaced at about a 5-year interval, MPRs were determined experimentally on a 0-to-5 scale for test sites in several states. Longitudinal profiles were measured in the left and right wheel tracks of the road segments that were rated to support the development of an index that predicted MPR from profiles.

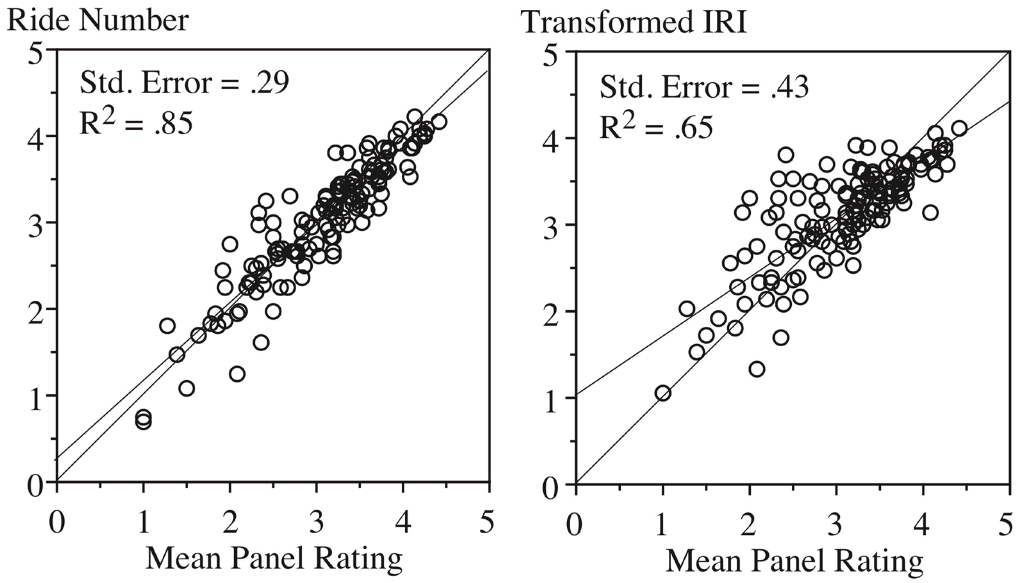
When the NCHRP studies were performed, the IRI was not well known. However, the researchers investigated a quarter-car analysis nearly identical to the IRI and found that the quarter-car index correlated significantly less well to panel ratings than a profile-based roughness index sensitive to short-wavelength content. (Subsequent studies have shown that higher correlation is obtained with IRI if an appropriate nonlinear transform is applied to it.)

Profile-based analyses were developed to predict MPR. A method was developed in which PSD functions were calculated for two profiles and reduced to provide a roughness index called PI (profile index). The PI values for the two profiles were then combined in a nonlinear transform to obtain an estimate of MPR. Although the abbreviation was the same, the PI used to relate to MPR had no explicit connection to PI measured by a profilograph (i.e., Profilograph Index).

RN is an estimate of MPR.

The mathematical procedure developed to calculate RN is described in NCHRP Report 275, but not in complete detail. Software for computing RN with the PSD method was never developed for widespread use.

In 1995, some of the data from the two NCHRP projects and a panel study conducted in Minnesota were analyzed again in a pooled-fund study initiated by the FHWA. The objective was to develop, test, and document a portable algorithm for calculating RN from measured profiles. The figure below shows the correlation of the refined version of RN to MPR. For comparison, the figure shows a correlation involving a transformation of IRI.



Properties of the RN Analysis

The RN calculation algorithm uses the same filtering method as the IRI, with some modifications to emphasize the specific content that relates to user satisfaction with rideability.

RN uses the 0-to-5 PSI scale.

The 0-to-5 scale for PSI was used because it is so familiar to the highway community. However, the methods used in the NCHRP research were different from the methods used in the older tests. The newer methods are based on a better understanding of psychological scaling than existed when the early tests were done.

RN is a nonlinear transform of a statistic called PI.

The index used in the ride number analysis is called PI, for "profile index." Like other profile-based roughness indices, PI generally ranges from 0 for a perfectly smooth profile to positive values proportional to a type of roughness. PI is transformed to a scale that ranges from 5 (perfectly smooth) to 0 (the maximum possible roughness). The experimental data validate the scale for values from 1 to 4.5.

The choice of scale creates a highly nonlinear relationship between profile variations and RN. If the RN is known for a profile, and the values of elevation are all doubled to increase roughness by a factor of 2, the RN will go down. However, the amount that RN decreases cannot be determined simply.

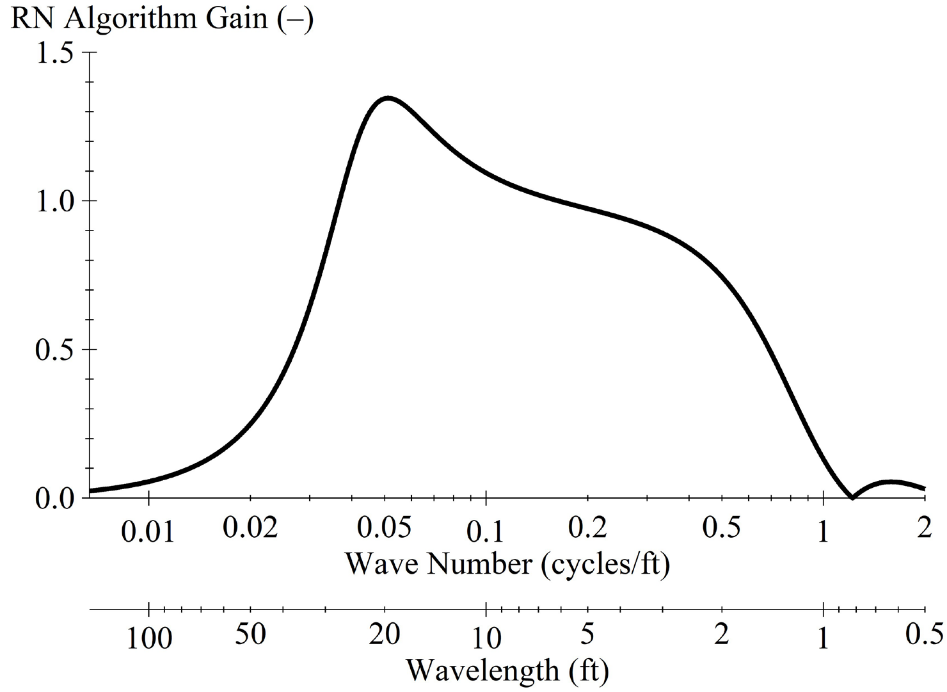
Nonlinearity limits some applications of RN.

Nonlinearity poses no problem for the collection of roughness information to describe the condition of a road network. For roughness collected on a per-mile basis (or any standard length), PI values are converted to the 0-to-5 scale and entered into a data base.

Some advanced capabilities of the IRI, such as the roughness profiles, are difficult to apply. The problem is that RN values for adjacent sections of profile cannot be averaged in the same way as IRI. For example, if one mile has an RN value of 3 and the next has an RN of 4, the RN for the two-mile segment is not 3.5—it is approximately 3.37.

PI and RN are sensitive to shorter wavelengths than the IRI.

The figure below shows the sensitivity of PI. As in the earlier section on IRI, this figure shows the response of the PI calculation algorithm for a slope sinusoid. If given a sinusoid as input, the PI filter produces a sinusoid as output. The amplitude of the output sinusoid is the amplitude of the input slope, multiplied by the gain shown in the figure.



The maximum sensitivity of the RN algorithm occurs at a wave number of 0.051 cycles/ft, which is a wavelength of 19.54 ft. The RN algorithm greatly reduces aspects of the profile with wavelengths less than 0.96 ft and more than 75.80 ft, because the gain is 0.1 or less. The gain is at 0.25 or above in the wavelength range from 1.13 ft to 49.95 ft. Recall that the gain of the IRI is 0.25 or above for wavelengths of up to 125.76 ft. As such, the IRI algorithm is much more sensitive to roughness in the wavelength range from 49.95 ft to 125.76 ft than the RN algorithm. In contrast, the RN algorithm is much more sensitive than the IRI algorithm to roughness in the wavelength range from 1.13 ft to 3.30 ft.

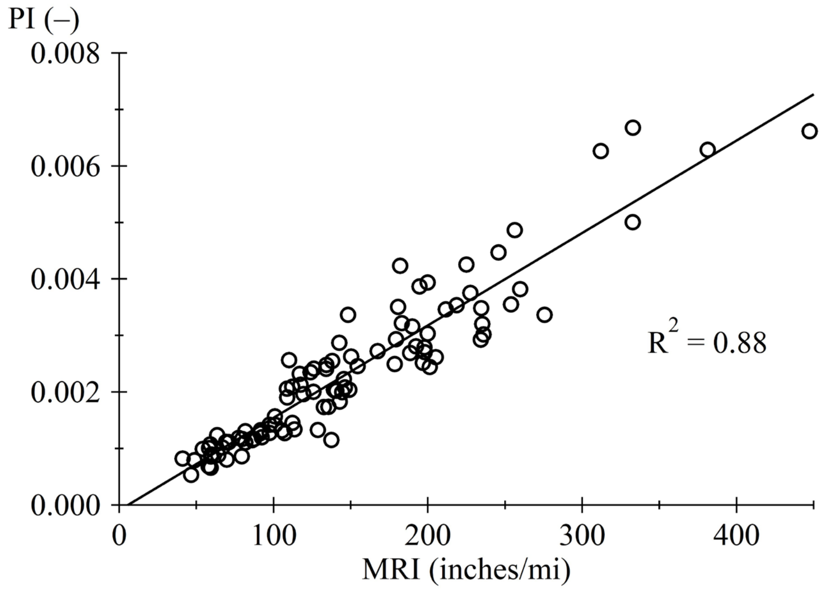
Valid measurement of RN is more difficult than IRI.

RN is portable, but not as much as the IRI. Reproducible measurement of RN requires a higher level of consistency in filtering, sampling, and recording practices, because it is sensitive to shorter wavelengths than the IRI.

RN is correlated to IRI but the two are not interchangeable.

The content of a road profile that affects RN is different than the content that affects the IRI. Each provides unique information about the roughness of the road, although there is correlation.

For example, the next figure shows the correlation between MRI and the PI statistic (defined below), which is used to determine RN. Both MRI and PI incorporate roughness from both wheel tracks. The correlation for an individual wheel track is lower.



Definition of RN

The descriptions of the RN background and properties above are intended to give an idea of how to interpret the RN scale. RN is rigorously defined as a specific mathematical transform of a true profile. The specific steps taken in the computer program to compute RN are listed below.

RN is calculated from one or two profiles.

RN is ideally calculated from the profiles in the left and right wheel tracks. Each profile is processed independently, and the results are combined in the last step. RN can also be calculated for a single profile if only one is available, but the quality of the relationship to MPR is reduced.

The profile is filtered with a moving average with a 250-mm (9.85-inch) base length.

The moving average is included in the RN algorithm to simulate the ability of a tire to envelop small-scale asperities within the contact patch. The moving average also reduces the sensitivity of the RN algorithm to variations in the sampling and recording intervals used by different profilers. Note that the wave-number response plot for the RN algorithm provided above includes the effect of the moving average.

The 250-mm (9.85-inch) moving average filter should be omitted for profiles obtained with some systems.

The computer program used to calculate the RN does not apply the filter unless the recording interval of the profile is shorter than 6.56 inches. If the recording interval is less than 6.56 inches, the moving average should be omitted if the profile has already been filtered by an anti-aliasing filter that attenuates wavelengths shorter than 1.5 ft.

The profile is further filtered with band-pass filter.

The filter in the RN algorithm uses the same equations as the quarter-car model in the IRI. However, different coefficients are used to obtain the sensitivity to wave number shown in the last figure. The parameters for the RN filter are:

$$\frac{k_s}{m_s} = 390\ (1/\text{sec}^2) \qquad \frac{k_t}{m_s} = 5120\ (1/\text{sec}^2) \qquad \frac{c}{m_s} = 17\quad (1/\text{sec}) \qquad \frac{m_u}{m_s} = 0.036\ (-)$$

The settings for the RN algorithm also specify 80 km/hr (49.7 mi/hr) as the simulated travel speed. These parameters do not describe the properties of a vehicle. Rather, they were tuned to maximize the ability of the RN to predict MPR in past research.

The filtered profile is accumulated to produce a PI value.

The filtered profile is reduced to yield an RMS value called PI, which has units of dimensionless slope (ft/ft, etc.).

PI is transformed to RN.

RN is defined as an exponential transform of PI according to the equation:

$$\text{RN} = 5e^{-160(\text{PI})}$$

If a single profile is processed, its PI is transformed directly. If two profiles for both the left and right wheel tracks are processed, values for the two are combined with the following equation, and then the transform is applied.

$$\text{PI} = \sqrt{\frac{\text{PI}_\text{L}^2 + \text{PI}_\text{R}^2}{2}}$$

Details of Ride Number are handled in computer software.

The above summary illustrates how the RN fits the earlier description of a generic profile-based roughness index. The analysis is applied to two profiles, the profile is filtered (twice), the filtered result is accumulated, and the accumulated value is cast onto the familiar PSI scale.

ASTM E1489-08 provides a standard for calculating the RN from measured profiles. References for more information about the development of the standard RN algorithm are:

Sayers, M.W. and Karamihas, S.M., "Estimation of Rideability by Analyzing Longitudinal Road Profile." *Transportation Research Record 1536* (1996) pp. 110-116.

Sayers, M.W. and Karamihas, S.M., "Interpretation of Road Roughness Profile Data." Federal Highway Administration Report FHWA RD-96-101 (1996) 166 p.

What Is the Effect of Length?

For monitoring the overall health of a pavement network, it is sufficient to determine roughness levels on a per-mile basis or some other manageable length. For evaluation of a specific pavement segment or for construction quality control, it is often suitable to report roughness over a shorter length. Reporting roughness using a short length helps identify specific profile features that contribute disproportionately to the overall roughness.

Roughness indices can be computed for various lengths of profile.

Consider the following table showing IRI values for one mile of profile. When the profile is partitioned into 0.1-mile-long segments, the lowest IRI value is 44.0 inches/mi and the highest IRI value is 130.3 inches/mi. When the IRI is calculated over the entire mile, the IRI value is 71.0 inches/mi, which is the average of the IRI values for the ten 0.1-mile-long segments.

Segment Start (mi)	Segment End (mi)	IRI (inches/mi)
0	0.1	51.0
0.1	0.2	48.7
0.2	0.3	44.0
0.3	0.4	118.2
0.4	0.5	130.3
0.5	0.6	59.9
0.6	0.7	83.3
0.7	0.8	57.4
0.8	0.9	57.1
0.9	1.0	60.0

Roughness indices are affected by the segment boundaries.

The following table shows IRI values for the same profile used in the previous example, with a starting point that is shifted 0.05 miles downstream. This 1-mile-long segment shares 95 percent of its length with the previous example, and the average IRI is similar: 70.9 inches/mi. For the 0.1-mile-long segments using the new boundaries, the lowest value is 41.2 inches/mi and the highest value is 197.5 inches/mi. The change in values for the individual segments owes to the shift in starting point. With the old starting point, the effect of the largest source of roughness between 0.35 miles and 0.45 miles was split between the two segments from 0.3 miles to 0.4 miles and 0.4 miles to 0.5 miles. With the new starting point, the roughest area appears within a single segment.

Segment Start (mi)	Segment End (mi)	IRI (inches/mi)
0.05	0.15	41.2
0.15	0.25	49.7
0.25	0.35	50.6
0.35	0.45	197.5
0.45	0.55	50.6
0.55	0.65	78.0
0.65	0.75	68.5
0.75	0.85	52.7
0.85	0.95	63.3
0.95	1.05	57.5

Segment length affects variation in roughness index values.

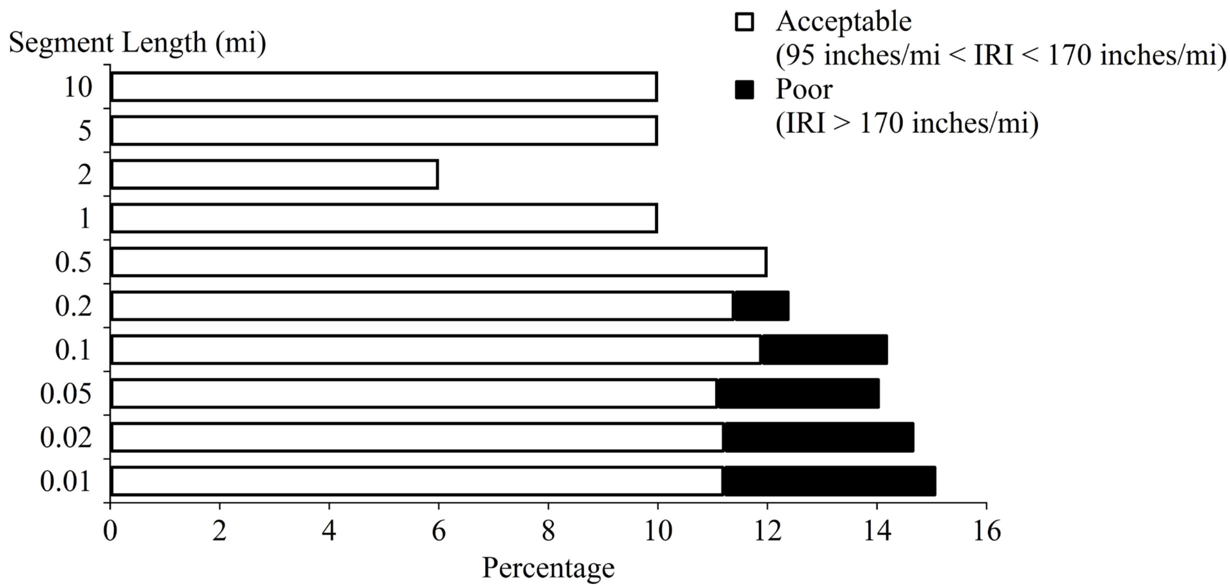
For the same overall length, more variation in roughness appears as the segment length decreases. The following table summarizes IRI values using various segment lengths for a 100-mile-long profile collected on an interstate highway. For each segment length, the table provides the percentage of segments within each of six roughness ranges.

Segment Length (mi)	Number of Segments	< 60 inches/mi	60-95 inches/mi	95-120 inches/mi	120-170 inches/mi	170-220 inches/mi	> 220 inches/mi
100	1	0	100	0	0	0	0
50	2	50	50	0	0	0	0
20	5	20	80	0	0	0	0
10	10	40	50	10	0	0	0
5	20	40	50	10	0	0	0
2	50	42	52	4	2	0	0
1	100	41	49	7	3	0	0
0.5	200	45	43	7.5	4.5	0	0
0.2	500	47.6	40	6.2	5.2	1	0
0.1	1,000	50.7	35.1	7.4	4.5	1.8	0.5
0.05	2,000	53.25	32.70	6.45	4.65	2.15	0.80
0.02	5,000	55.98	29.34	6.50	4.72	1.70	1.76
0.01	10,000	58.11	26.81	5.94	5.27	1.88	1.99

The IRI for the entire 100 miles is 68.3 inches/mi. So long as the overall length is split into segments of equal length, the average IRI is 68.3 inches/mi. The IRI for the first 50 miles is 76.7 inches/mi, and the IRI for the second 50 miles is 59.8 inches/mi. Half of the overall length appears within each of two roughness ranges. As the segment length gets shorter, the individual roughness values spread out over a greater range. Using a segment length of 0.1 miles produces 1,000 individual roughness values, which occupy all six of the roughness categories in the table.

Summarizing roughness statistics using categories requires the use of a uniform segment length.

In 2017, the U.S. DOT established a segment length of 0.1 miles in their definition of a "pavement section" for the TPM system. The use of a uniform segment length reduces inconsistencies in the percentages classified into each roughness category. The following figure provides an illustration using the 100-miles of profile from the previous table. When the segment length is 0.1 miles, 11.9 percent of the segments fall within the range from 95 inches/mi to 170 inches/mi, which corresponds to the "fair" classification. The percentage varies from 6 to 11.22 for the segment lengths shown and is zero when the segment length is 20 miles, 50 miles, or 100 miles. When the segment length is 0.1 miles, 2.3 percent of the segments have IRI values greater than 170 inches/mi, which corresponds to the "poor" classification. The proportion of the overall length in the poor category also changes with segment length, and no segment appears in the poor category when the segment length is 0.5 miles or longer.



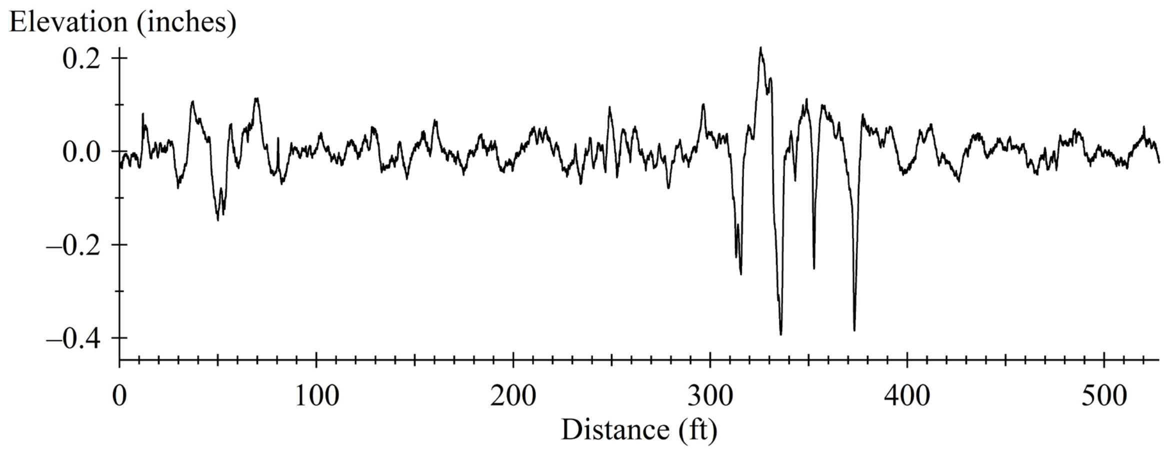
What Is a Roughness Profile?

A roughness profile provides a continuous report of roughness within a road section. Rather than a segment-by-segment summary of the roughness, it provides details about the way roughness varies along the profile. A roughness profile is sometimes called a continuous roughness report or a hot-spot plot, depending on the application.

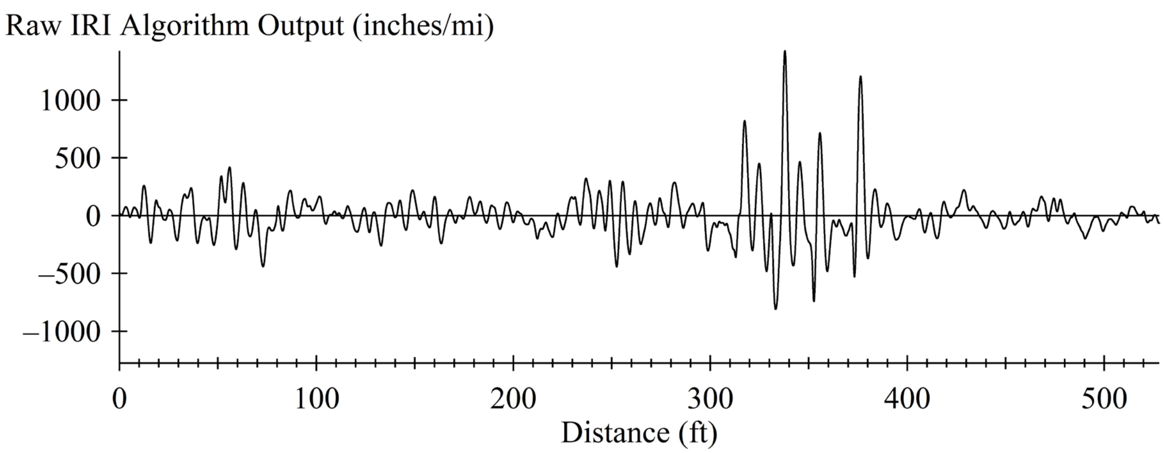
A roughness profile shows roughness versus distance.

Each value in a roughness profile is the average roughness over a specific length, called the base length (*L*). Any point in a roughness profile is the average roughness over a range that runs from a location *L*/2 prior to that point to a location *L*/2 past that point. The following plots illustrate the details of the way the IRI and an IRI-based roughness profile are calculated.

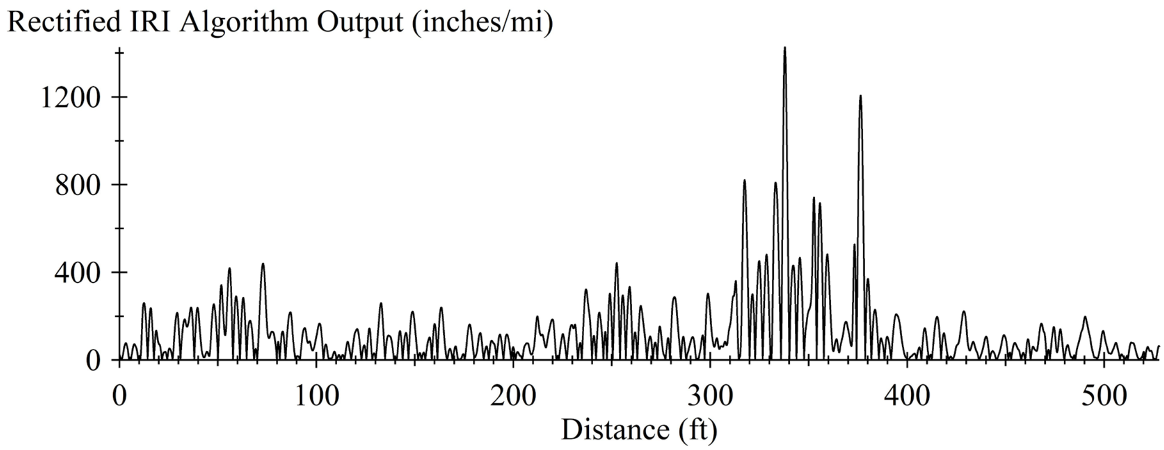
The first plot shows a 528-ft-long profile, which is high-pass filtered to help recognize features that affect roughness. Note the area from 300 ft to 400 ft, which seems rougher than the rest of the profile.



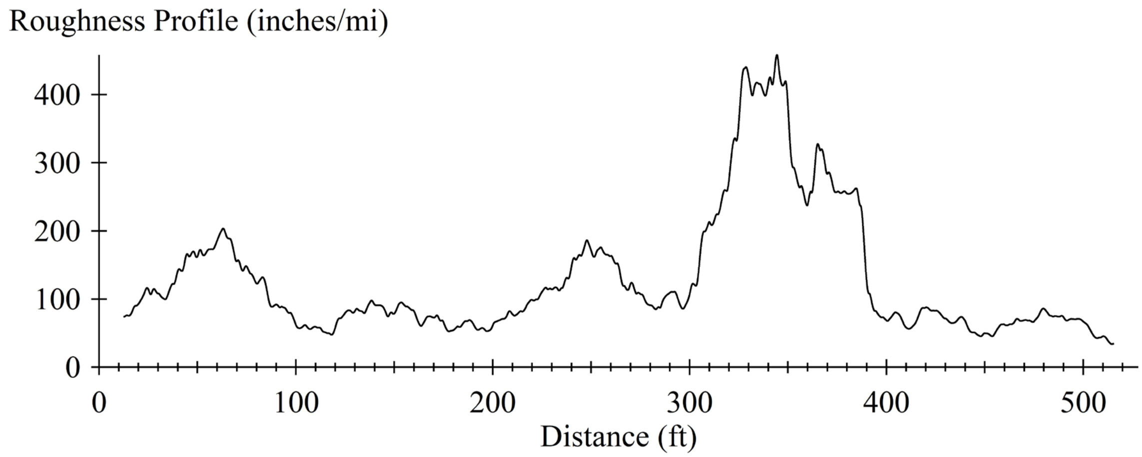
The second plot shows the raw output produced by the IRI algorithm using the profile as input. The plot fluctuates around zero, and has units of inches/mi. An interpretation of the IRI algorithm output is motion across the suspension of the Golden Car normalized by travel distance. (The IRI algorithm output has an equivalent mathematical interpretation: spatial velocity across the suspension.) The raw output of the IRI algorithm is not particularly intuitive. Fortunately, the raw output from the IRI algorithm is just an intermediate step toward getting the roughness profile, and few users of road profiles ever need to examine the plot.



The third plot is rectified output from the IRI algorithm. It shows the absolute value of every point in the second plot. The overall IRI of the profile is the average of the rectified IRI algorithm output over the entire length. For this example, the overall IRI is 123.8 inches/mi. Note that the area from 300 ft to 400 ft contributes more to the overall IRI than other ranges within the profile.



The final plot is the roughness profile. It was generated by applying a moving average to the rectified IRI algorithm output with a base length of 25 ft. As such, each point in the plot is the IRI of a 25-ft-long segment with its center at that location. For example, the peak value in the roughness profile of 458.5 inches/mi appears at 344.4 ft. This is the roughness of a segment with a starting point at 331.9 ft and an ending point at 356.9 ft. The roughness profile shows that the area between 300 ft and 400 ft is much rougher than the rest of the profile, which corresponds to the range within the profile that stood out in the high-pass filtered profile plot.

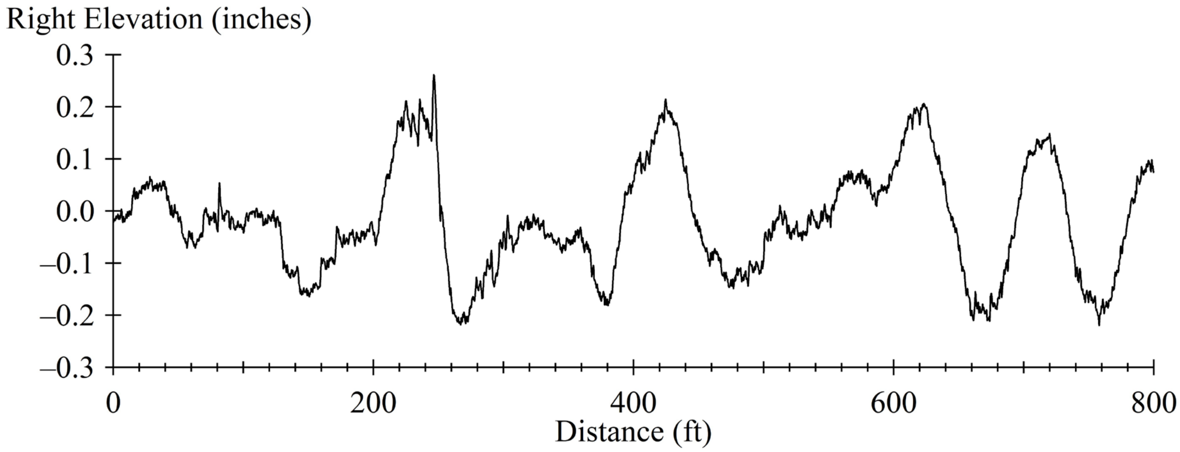


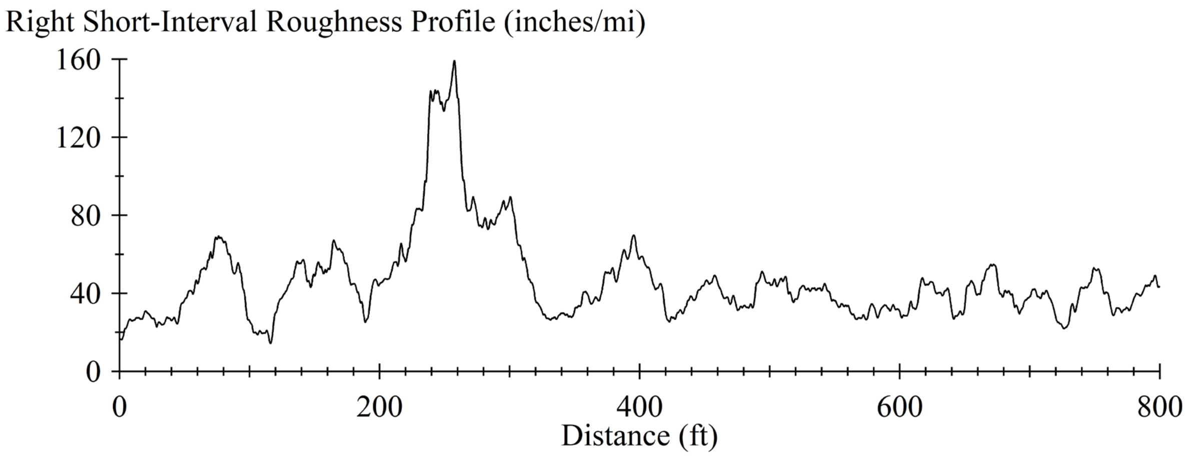
Note that the roughness profile covers an area shorter than the input profile. In this case, the input profile ran from 0 to 528 ft. Since the averaging used a 25-ft base length, no values appear in the roughness profile for the first 12.5 ft and the last 12.5 ft of the measured segment.

A short-interval roughness profile helps pinpoint localized roughness.

The previous example used a base length much shorter than a typical pavement segment, which helped identify the location and severity of a rough feature of the profile. Although no universal definition exists for a short-interval roughness profile, American Association of State Highway and Transportation Officials (AASHTO) R 54-14 defines localized roughness as "any 25-ft segment of roadway that contributes disproportionately to the overall roughness index value."

The following example demonstrates the use of a short-interval roughness profile for construction quality control. The first plot shows 800 ft of profile measured on a newly paved asphalt surface. The contractor who placed the surface observed that the paver temporarily came to a stop in the location that corresponds to 247 ft in the profile. The roughness profile shows that most of the wheel track is very smooth within this pavement segment, but localized roughness appears in the position of the paver stop.





A short-interval roughness profile is a tool for quality control and quality assurance.

A contractor may use a roughness profile for two reasons. First, some quality assurance specifications for new pavement surfaces use a short-interval roughness profile to identify localized roughness. Values in the short-interval roughness profile above a particular threshold may trigger a negative pay adjustment or a requirement for corrective action. Quality assurance specifications that include localized roughness provisions guard against localized surface defects and roughness that degrade user satisfaction on an otherwise smooth roadway.

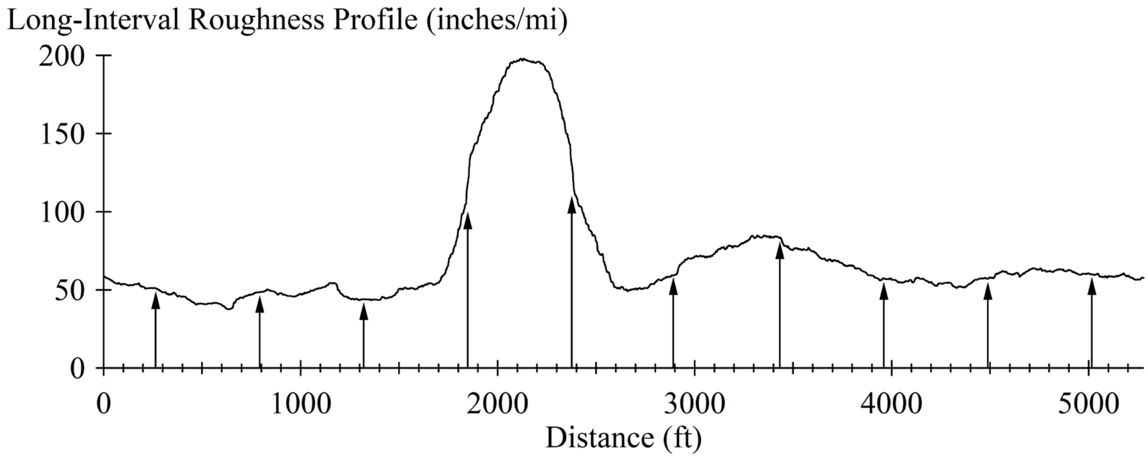
Second, the short-interval roughness profile helps identify features that contribute the most to the overall roughness. If the overall IRI of a road segment is higher than expected, the short-interval roughness profile provides a way to determine the location and severity of the surface defects that contribute to roughness the most. With either motivation for using a short-interval roughness profile, inspection of the elevation profile at the location of localized roughness helps diagnose their cause. Be advised that a peak in the short-interval roughness profile often appears a short distance downstream of the feature that causes it.

A short-interval roughness profile is a tool for analyzing in-service pavements.

For in-service pavements, a short-interval roughness profile helps identify pavement surface features that contribute heavily to the overall roughness of pavement sections, as well as roughness at built-in features, such as utility covers, drainage inlets, etc. In some cases, sources of localized roughness appear less than 25 ft apart, and a shorter base length may be needed to pinpoint them.

A long-interval roughness profile helps inspect long road segments.

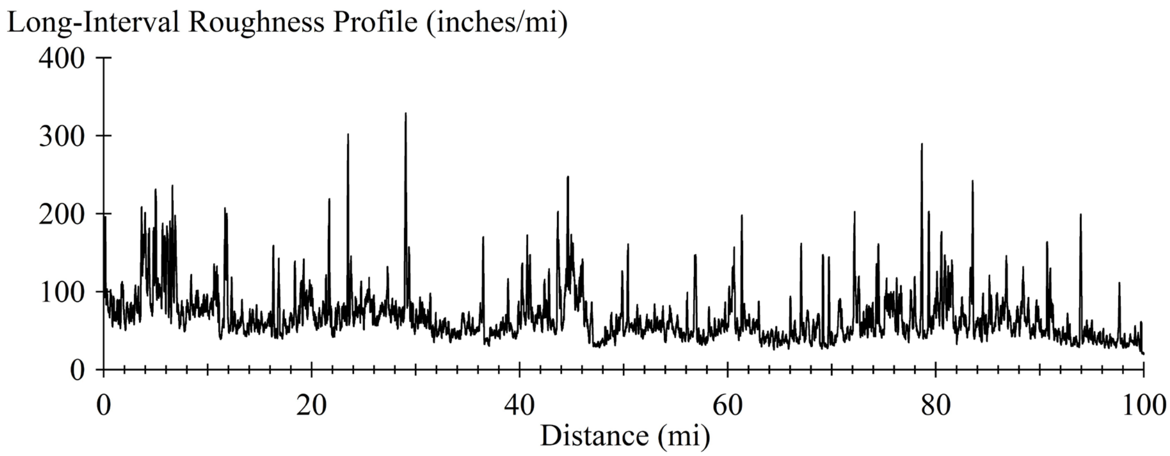
A long-interval roughness profile is calculated the same way as a short-interval roughness profile. The only difference is the use of a longer base length. The following figure shows a long-interval roughness profile produced using a base length of 528 ft. This long-interval roughness profile used the same mile-long profile as input that produced the IRI values for 0.1-mile-long pavement sections in the discussion of segment length.



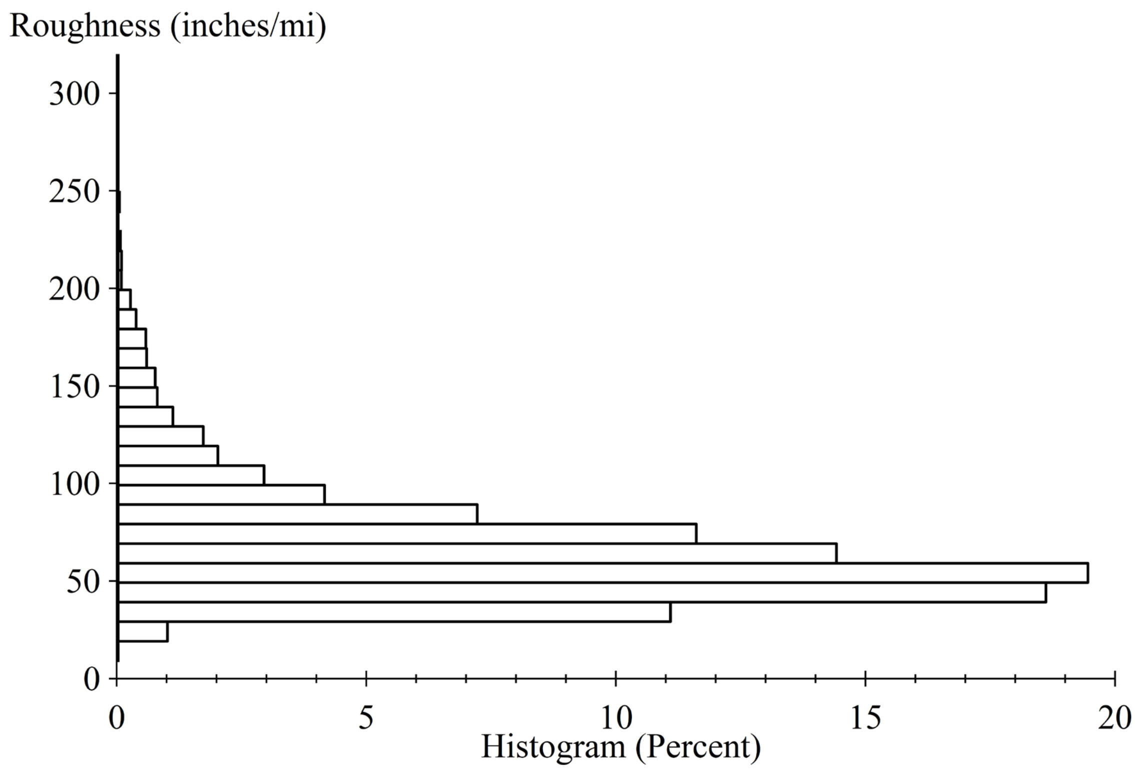
Each point in the long-interval roughness profile represents the average IRI for a segment of the profile starting at a location 264 ft prior to that point and ending 264 ft past that point. The long-interval roughness profile provides the IRI of every possible 528-ft-long road segment. The arrows in the figure correspond to the IRI values from the discrete tabulation in the discussion of segment length. The arrows illustrate the potential lost information when reporting roughness at discrete intervals, because the roughest 528-ft-long segment is not among the discrete roughness values.

A histogram helps summarize the information from a roughness profile.

The next figure shows a long-interval roughness profile for the 100-mile-long profile used in the discussion of length. The long-interval roughness profile used a base length of 528 ft. The profile was measured on an interstate highway. As shown, a large proportion of this profile has an acceptable level of roughness, with several locations that stand out as rough. Most of the isolated rough areas appear at bridges.



A histogram of values from a roughness profile provides a snapshot of the roughness status for long road segments. The next figure shows the roughness histogram from the long-interval roughness profile over the 100 miles of interstate. For this example, 11.92 percent falls within the FHWA "acceptable" category, and 2.01 percent falls within the FHWA "poor" category. The prior example, which used discrete segments that were 0.1-mile-long, approximated the roughness distribution derived using the roughness profile and histogram discussed here.



For construction quality control and assurance, using a long-interval roughness profile and its histogram provides a way to examine the overall status of a newly placed pavement surface. Using the short-interval roughness profile and its histogram provides a way to identify localized roughness and track their overall range.

What Are Errors?

Up to this point, we have covered what you can do with profile measurements that are valid. The assumption is that statistics such as IRI, RN, PSDs, and roughness profiles calculated from actual profile measurements are the same as would be obtained by analyzing the true profile. How do you know if you can trust your profile measurements? Verification testing, sensor calibration, system checks, and validation ensure the quality of measured profiles. The rest of this book addresses errors and variations in profile measurements.

In an ideal situation, repeated passes by a profiler down the same imaginary line on a road would produce the same profile and profile-based statistics each time. Further, the results would match those obtained with other profilers. In practice, the agreement is not perfect.

Accuracy is a lack of error.

Optimistic engineers and equipment developers speak of the accuracy of their systems. However, accuracy is defined by error. The smaller the error, the better the accuracy.

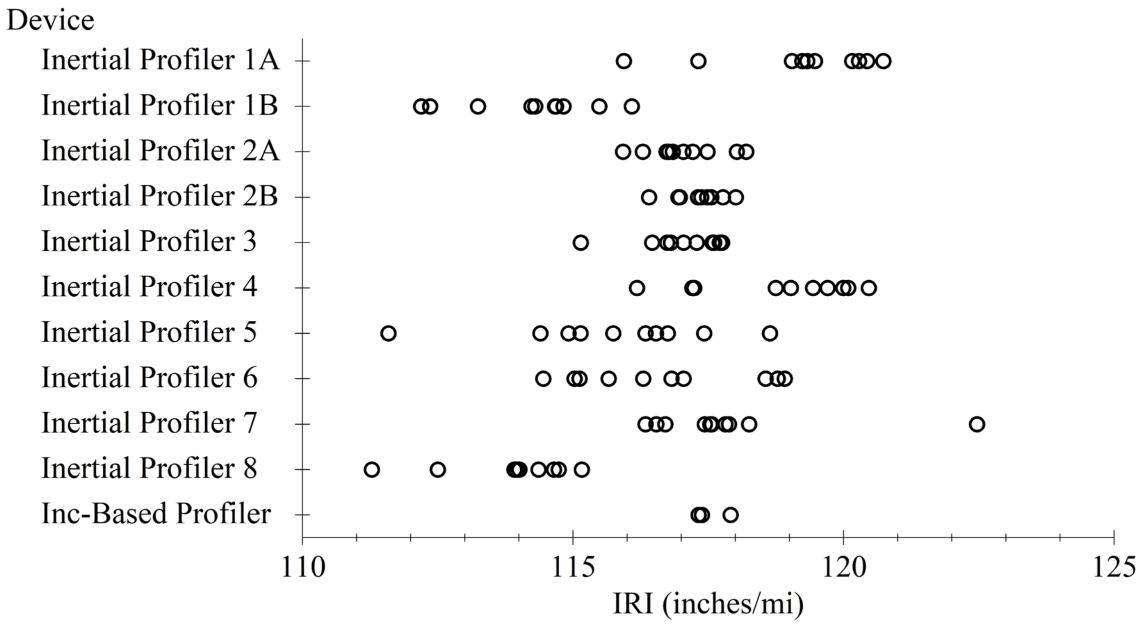
A lack of systematic error in a profile measurement means it is not biased relative to the result that would be calculated from the true profile. A lack of random error in a profile measurement means it produces the same result consistently in multiple passes, without excessive variation. Performing a sufficient number of passes and averaging the results can overcome random errors in profile-based roughness index values. Performing multiple passes cannot overcome bias.

Repeatability is the ability to obtain consistent measurements with the same device at (nearly) the same time.

Suppose we make repeated measurements with a profiler along the same line, over and over, multiple times in succession. If we were to measure the variations in the profile or profile-based statistics of each, the results would characterize the repeatability.

It is common to characterize the variation in a summary index such as IRI by taking the standard deviation. The units of the standard deviation match those of the index. For example, if we process the profiles to obtain IRI with units of inches/mi, the lack of repeatability will be in terms of inches/mi. To scale the variation as a percentage, divide the standard deviation by the mean value and multiply by 100.

The figure below shows the IRI values produced in ten passes by each of ten inertial profilers on an asphalt test section with smooth texture. The figure also shows the IRI values produced by three passes of an inclinometer-based profiler.



Repeatability requirements depend on the application.

No profiler is perfectly repeatable, and roughness index values from some measurements are higher than the expected average for a given line on the road, while others are lower. For quality assurance on newly placed road surfaces, where pay adjustments based on roughness may be imposed, only a small level of variation may be tolerated. A very repeatable profiler may be needed, or multiple repeated measurements may be required to overcome run-to-run variations. For surveys of the roughness of in-service roads, a higher level of variation is tolerable, because it may average out over many road segments. (Bias, in contrast, will not average out.) In any application, an understanding of the expected repeatability is important.

The number of repeated measurements that should be taken to characterize roughness depends on the repeatability of the instruments and the testing methodology. If an instrument is very repeatable, a single measurement might be sufficient. For a less repeatable instrument, enough measurements should be taken to obtain a mean value of the roughness index with some confidence.

Reproducibility is the ability to obtain consistent measurements with different devices at (nearly) the same time.

Suppose two profilers make repeated measurements along the same line and calculate a roughness index, such as the IRI, from the profile measured in each pass. The difference in the average index value obtained by each device characterizes the reproducibility. In the previous figure, inertial profiler 1A and 1B are different profilers of the same design from the same manufacturer. Comparing their output is an example of reproducibility among copies of that design. The same is true of inertial profilers 2A and 2B, which are also two profilers of similar design from the same manufacturer.

Portability is the ability to obtain consistent measurements with devices with unique designs.

All the inertial profilers in the figure above use the same measurement principle. Comparing their output is an example of reproducibility among inertial profilers. Comparing their output is also an example of portability between inertial profilers of unique designs. Design variations between inertial profilers include differences in sensor types (e.g., height sensor or accelerometer types), host vehicle, mounting location and hardware, and processing techniques (e.g., filter type and settings).

The previous figure includes IRI computed from three repeated profile measurements by an inclinometer-based profiler. Comparison of output from the inertial profilers to the output of the inclinometer-based profiler is an example of portability between the two measurement concepts. The issue of portability of profile measurements and measurement of roughness index values is likely to arise more often as novel methods of roughness measurement or estimation emerge.

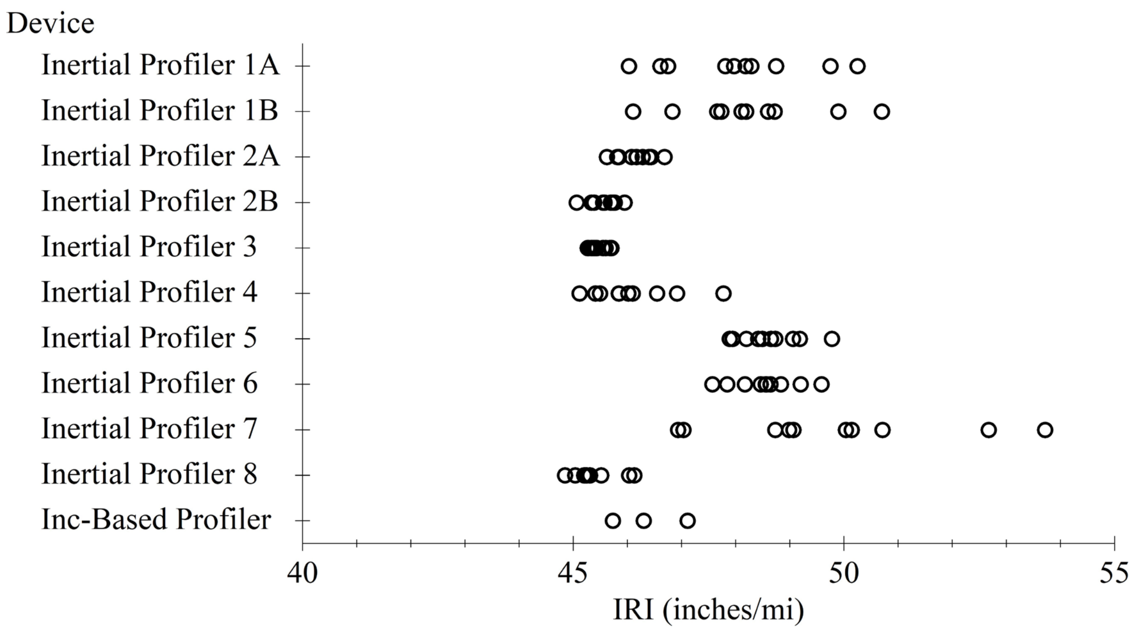
The standard for portability is the true profile. Recall that a profiler is considered valid if values of a given roughness index are neither higher nor lower, on average, than values of the roughness index obtained from true profile. Because any valid profiler is linked to the true profile, its measurements are portable by definition.

Agencies often designate an inclinometer-based profiler as a reference device and verify the output of inertial profilers based on their ability to reproduce its measurements. This is a legitimate practice, so long as the inclinometer-based device is valid. That is, the inclinometer-based device must capture important aspects of the true profile, such as the content that affects the IRI, with sufficient accuracy. When that is the case, the ability of the inertial profilers to reproduce the designated reference measurements is a surrogate for characterizing their accuracy.

Surface properties affect the repeatability, reproducibility, and portability of profilers.

To characterize the repeatability of an instrument or the reproducibility and portability between instruments, run repeat tests for different road conditions of interest. In many cases, different roughness levels and different surface texture types affect the output of profilers differently.

The next figure shows a set of measurements on a smooth asphalt test section with smooth texture. The repeatability of each profiler, the reproducibility between inertial profilers, and portability between the inclinometer-based profiler and the inertial profilers is different than it was on the rougher section.



Pavement surface properties affect the quality of profile measurements in several ways. For example, very smooth pavements may pose a challenge to profilers if their sensors lack the resolution to capture small variations with sufficient precision, or if variations in the measured signals are not large enough relative to sources of sensor error, such as noise or drift. Very rough pavement may cause motion of the host vehicle large enough to exceed the valid range of the sensors. Profiles of rougher pavement typically exhibit a higher level of lateral variation than smoother pavement, which increases variation in profile measurements when the same line is not followed precisely in each pass. Coarse surface texture poses a particular problem for profilers that do not sense the road surface over a sufficiently wide area.

Repeatability and accuracy can be calculated for profile.

If one is using profile plots to diagnose sources of pavement roughness, then the profile needs to be repeatable and accurate, not just a profile-based roughness index. One method, described in AASHTO R 56-14, produces a rating of repeatability and accuracy of the profile and is described in one of the following sections.

What Is Verification Testing?

Verification testing is used to confirm that a piece of measurement equipment is operating properly. Typically, verification testing will neither determine if the equipment was properly calibrated nor validate its design. Calibration and validation tests are more difficult to conduct and interpret and are normally done less frequently.

Before placing a new profiler into service, verify that its output seems reasonable.

A profiler provides large amounts of information that are difficult to obtain any other way. However, roughness indices such as IRI and RN relate well to the public's perception of ride quality. As a member of the public, you can evaluate whether roughness index values measured by a profiler make sense. Tests of this sort are sometimes called reality checks or sanity checks.

These simple tests cannot determine the accuracy of the system, but they can verify that it is producing reasonable output.

For example, nearly all profilers include software to compute IRI. IRI values for highways typically range from 30 inches/mi for a very smooth new pavement, to 90 inches/mi to 120 inches/mi for average sections, to 190 inches/mi and higher for sections that should be considered for repair. Highways rarely get to roughness levels higher than 220 inches/mi. In 2018, the FHWA Highway Performance Monitoring System database included IRI values for 0.1-mile-long segments that covered about 232,200 miles of roadway. More than 77 percent of the distance had IRI values less than 120 inches/mi, and over 92.5 percent had IRI values below 190 inches/mi.

In the field, use an output that you understand through experience.

If you are using the profiler to measure IRI values, verify its operation in the field by inspecting IRI values. If you are using it to view profiles, verify its operation by inspecting plots of profiles.

Determine the repeatability.

No profiling device is perfect. Errors exist. If you profile the same imaginary line on the road several times, you will not get exactly the same result each time. However, you should get almost the same result with repeated measurements.

The repeatability of roughness index values usually improves for longer profiles. If you are processing the profiles to get IRI, you should be able to repeat the measured values within 2 percent for profiles that are one mile in length. For shorter profiles, larger variations occur. Learn the variability of your profiler during repeat measurements and watch for a change over time.

If the output is not reasonable, contact the manufacturer.

If you obtain questionable outputs from a profiler, read the profiler manual and check that everything is connected properly. Perform all recommended field calibrations and system checks. If you continue to suspect the profiler is not collecting accurate measurements, there is little you can do to compensate for it. Something in the system is not working correctly and must be fixed.

Periodically verify that the measurements from the profiler remain reasonable.

A profiler has many components that can fail and place the accuracy of the measurements in jeopardy. Like any complex measurement system, your profiler should be tested periodically to ensure it is working as intended. You should establish periodic “sanity checks” to verify that the profiler is working properly.

A good practice is to establish control sections that can be used identify emerging sources of measurement error. Profile a set of pavement sections with stable roughness on a regular basis and compare the new profiles and roughness index values with those obtained in the past. Try to include a pavement section that is smooth and another section that is rough.

The readings from control sections are not used to adjust future outputs. The instrument is either a valid profiler, or it is not. If it is not, get it fixed by addressing the source of error or stop using it.

What Is Calibration?

Much of engineering measurement involves conversion between different physical variables (voltage, inches, etc.) and conversion between different units applied to the same type of measure (inches, meters, etc.)

Calibration is a process of correcting the scale of a measuring device.

Sensor calibration involves determining the relationship between the outputs of an instrument to known inputs. To check the calibration of a laser height sensor, set the distance between the laser and its target to a known value and read the output of the sensor. Perform the same test for several distances over the useful range of the sensor and compare the outputs to the inputs. If the outputs do not agree sufficiently well with the inputs, the sensor is not properly calibrated.

For sensors with a linear calibration, two coefficients may be adjusted: the offset and the gain. A proper value of offset ensures that the sensor outputs are not shifted relative to the actual physical values. A proper value of gain ensures that the sensor captures the magnitude of changes in the physical values correctly.

Calibration of some profiler sensors is only done in the laboratory.

A typical inertial profiler includes an accelerometer, a non-contacting height sensor, a longitudinal distance sensor, a computer, and assorted electronics to power the sensors and connect them to the computer. Each sensor is independently calibrated. If any of the components do not work properly or the sensors are out of calibration, then the system cannot provide valid profiles.

Depending on the system's design, the operator may not calibrate some of the individual components. Special equipment is usually needed. In most systems of similar complexity, the sensors are calibrated at the factory and remain in calibration throughout their life. For example, many inertial profilers provide a way to check the calibration of height sensors using a block test, but no way to adjust the calibration coefficients. If the height sensor fails the calibration check, re-calibration in a laboratory may be needed.

The manufacturer may require the operator to perform periodic calibrations of some parts of the system. For example, the calibration of the distance measuring instrumentation may be performed in the field without advanced equipment. However, if the manufacturer does not provide instructions for calibrating parts of the system, the operator is not meant to make adjustments.

You cannot calibrate a profiler by measuring roughness.

Given that you calibrate a sensor by giving it a known input and reading the output, you might suppose that a profiler should be calibrated by measuring a profile with a known roughness level (e.g., IRI). This is wrong.

The exact conditions that contribute to the roughness index of a true profile are usually unique for that profile. The contributions of various features to the roughness index value are different for another true profile even if it has the same overall roughness level.

When you check the value of a roughness index computed from your profiler by comparing it to a reference, this is a verification test. If the agreement is not satisfactory, then the profiler is not valid for that condition. The next step is to diagnose and address the source of the error. Using the difference between measured roughness and the reference value to establish an adjustment factor is not valid.

We do not calibrate equations or computers.

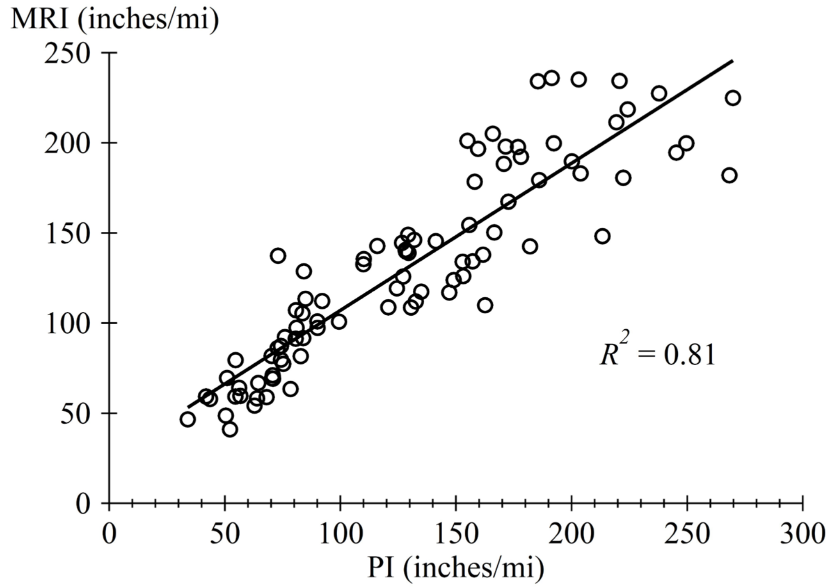
Remember, half of the profile measurement process involves interpretation of the profile data. The analyses applied to the profiles are fixed in the computer software. The equations are either programmed correctly the first time, or they are not. The analysis part of the process is not something that can change with time and use, and it is not a part of calibration.

What Is Correlation?

Correlation is a mutual relation or connection between different variables. Statistically, it is the degree of association between two data sets. Many statistics books further define correlation as the degree of linear association between two data sets.

Correlation analysis describes how much of the variation in one quantity is related to variation in another quantity.

Consider data comparing MRI with the PI used to calculate RN on several test sections. Each data set represents roughness index values calculated from the same profiles. Values of one index (MRI) are plotted on the ordinate (Y) axis against values of PI, plotted on the abscissa (X) axis. PI is expressed in inches/mi in the example to keep the plots on compatible scales. The figure shows that, in general, increases in values of MRI are linked with increasing values of PI. However, the relationship is not perfect. Both MRI and PI are based on the outputs of mathematical transforms applied to measured profiles, and each index responds differently to sinusoids according to their wavelengths; i.e., each index is sensitive to separate set of features within the profile. Each index may produce a different rank order of roughness for a group of pavement sections.



Regression creates a model used to predict one quantity from another.

Regression analysis fits a function to predict Y as a function of X. The constants in the function are calculated in the analysis to minimize the squared differences between the actual Y values and Y values predicted by the fitted function. In many applications the function is a linear equation with an offset and gain. For example, linear regression on the data shown in the figure above produces the following fitted line, which is shown in the figure:

$$f(PI) = 25.49 + 0.8165 \bullet PI = E(MRI)$$

The gain (i.e., slope) and offset (i.e., Y intercept) values minimize the sum of the squared differences between the values of *MRI* predicted by $f(PI)$ and the actual values. Since the prediction is not perfect, the output of the function is considered the expected value of *MRI*, or *E*(*MRI*), for a given value of *PI*.

Correlation analysis quantifies how much of the variation in one quantity is accounted for by another.

A typical measure of the quality of the linear fit between quantities is called the "R squared" value, or R^2. The R^2 value is the square of correlation coefficient, which has a specific mathematical definition. (See your favorite statistics textbook.) In the context of the example above:

$$R^2 = \frac{\text{variance of } f(\text{X})}{\text{variance of X}}$$

The R^2 value is normalized to stay in the range between 0 and 1. If the fitted equation predicts Y perfectly from X, then all of the data points in the figure would lie on the regression line, and R^2 would equal 1. If R^2 is zero, it means there is no relationship between X and Y and that the assumed form of the regression equation cannot use the values of X to improve the estimate of Y. In that case, the best estimate of Y is its mean value, regardless of the value of X.

The standard error summarizes the level of scatter around the fitted function.

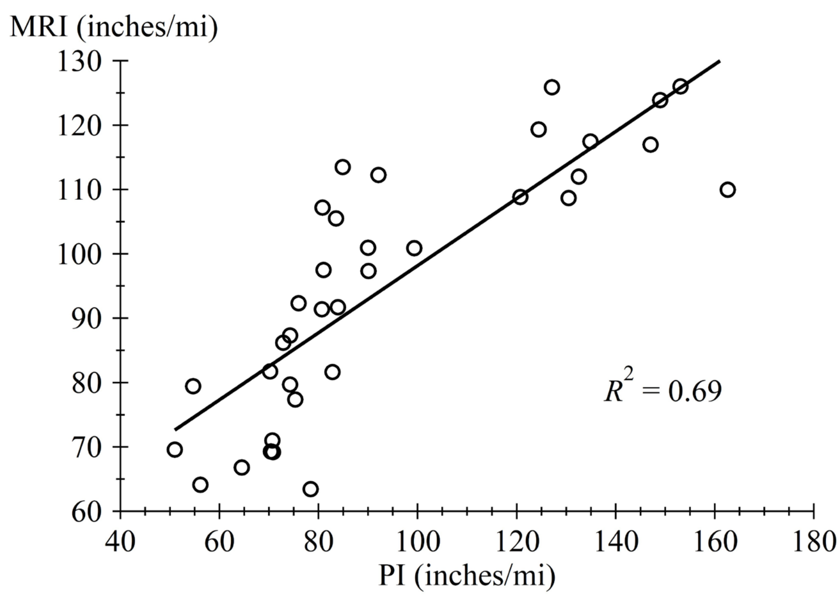
For each sample, the residual is the actual Y value minus the predicted Y value. In the example of linear regression above, the residual is the distance of each point on the plot above the best-fit line in the units of the vertical axis. Standard error of estimate is based on the sum of the squared residuals over the number of observations (*n*):

$$\text{Standard Error} = \sqrt{\frac{\sum\left(Y - f(X)\right)^2}{(n-2)}}$$

The use of *n–2* in the denominator has an explanation that only an expert in statistics would find remotely satisfying or intuitive. In any case, the standard error provides a very useful estimate of the expected error in each prediction of Y in the units of Y. Often, this is at least as important as R^2 for engineering purposes. The standard error for prediction of MRI by PI in the example above is 23.56 inches/mi.

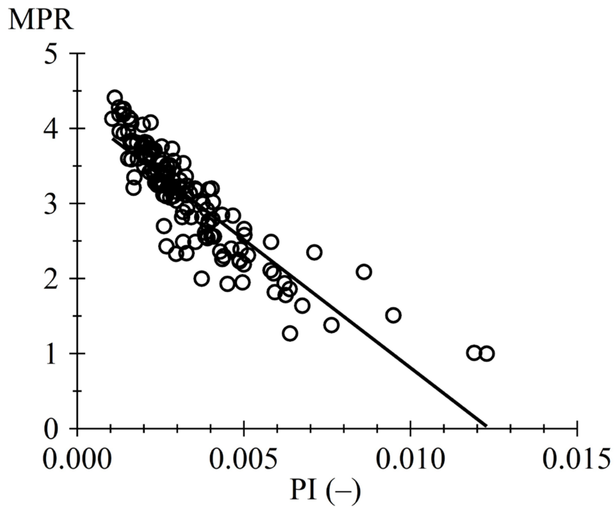
Correlation coefficient depends on the range of data.

The plot in the next figure shows the result of a linear fit of MRI to PI using a subset of the data from the previous example, which is confined to a third of the original range of MRI values. The level of scatter for this subset is about the same as it is for the full set. For example, the standard error for a linear fit to the subset is 23.57 inches/mi. However, using a limited range reduced the R^2 to 0.69, and changed the slope of the best-fit line to 0.5211.

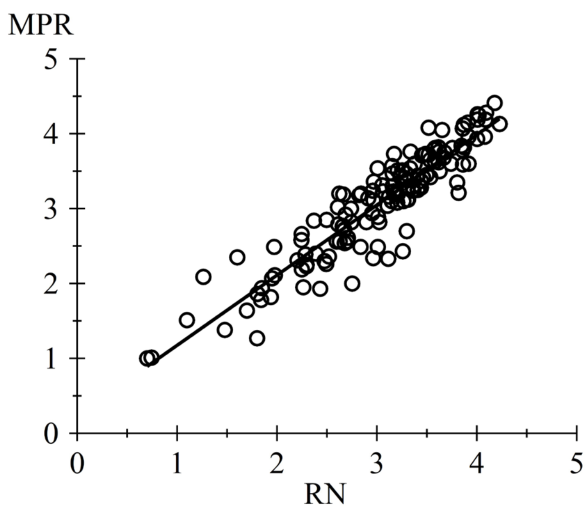


Correlation coefficient depends on the form of the model.

The following plots involve the use of correlation to predict MPR from a profile index. The plots show linear correlation of MPR to index values derived from measured profiles on 139 test sections used in Ohio in the development of RN. The first plot shows a linear fit of MPR to the PI used to calculate RN. These two quantities have a clear correlation. In this case, they are negatively correlated, because MPR shows a decreasing trend as PI increases. However, linear correlation is not particularly useful. For these data, the R^2 value is 0.77. For many applications in civil engineering design, an R^2 value in this vicinity is considered acceptable. However, two difficulties exist with the use of linear correlation for this dataset. First, the PI values are not distributed evenly. Many of the data points are clustered toward one side of the range, which reduces the legitimate implications that can be made from the best-fit line. Second, the residuals have a systematic bias relative to the best-fit line in various regions of the plot. That is, the data are not scattered symmetrically around the line. For example, a considerable proportion of the data points appear above the best fit line for low values of PI but appear below the best-fit line in the middle range.

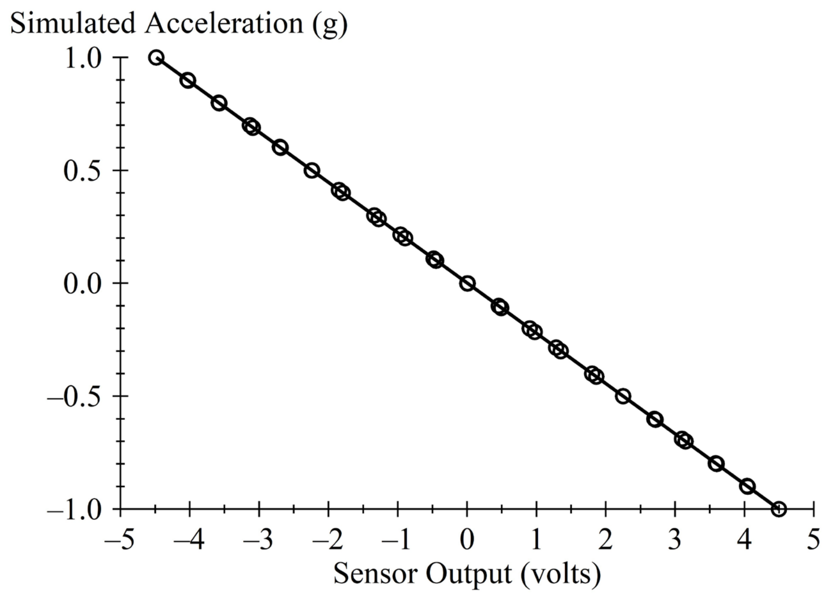


The next plot shows a linear fit of MPR to RN. RN is calculated using a non-linear transformation of PI to limit it to a 0-to-5 scale and cast it onto a scale that directly predicts MPR. The transformation to RN improved the R^2 to 0.85. The standard error of estimate for linear correlation improved from 0.348 for PI to 0.292 for RN. More importantly, the uniformity of the samples over the range and the symmetry of the scatter around the best-fit line improve the legitimacy of the correlation.



Correlation is often part of the sensor calibration process.

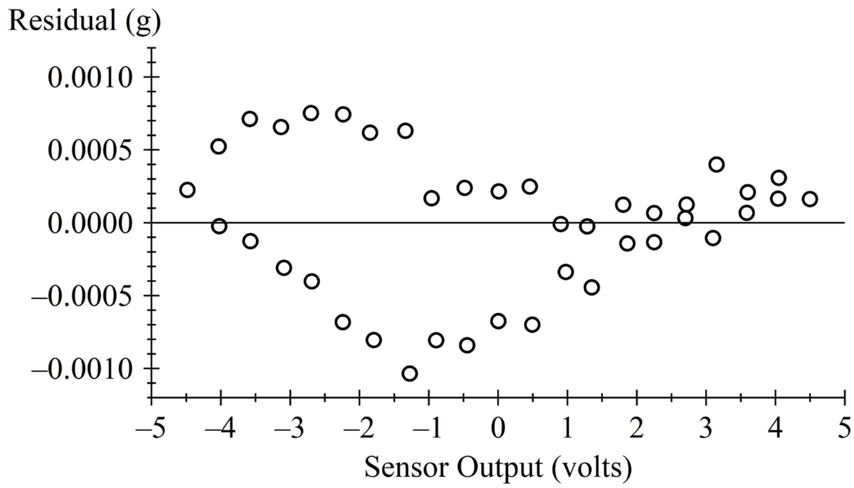
The figure below shows data collected during a laboratory calibration of a servo-type accelerometer. Servo-type accelerometers use a measurement of force to infer acceleration. As such, they detect the influence of gravity, even when they are still. If you align the sensitive axis of a properly calibrated servo-type accelerometer vertically, it will report 1 g. If you then turn it upside down, it will report –1 g. The calibration used tilt of the accelerometer at various angles around a full 360-degree rotation to simulate actual acceleration. The calibration was performed with 5 volts supplied to the sensor. At each of the tilt angles, the output of the sensor in volts was recorded.



Estimating simulated acceleration from its statistical relationship to sensor output is an example of calibration by correlation. In this example, the relationship between the output (voltage) and the input (simulated acceleration) is linear, and the R^2 value is virtually 1. That means a linear equation, with a slope and an intercept, can be used to convert voltage from the sensor to acceleration in a measurement device, such as a profiler:

$$A_{meas} = 0.0011423 - 0.22274 \bullet V_{sensor}$$

An important practice in sensor calibration is inspection of the residuals. The plot below shows the residual for each point, which is the difference between each value of simulated acceleration and the value predicted by the calibration equation. The plot helps characterize the precision and bias of the calibration in a way this is not visible in the calibration plot. If the residuals are too high for the intended application or have a systematic relationship to the input, the use of a given sensor or the calibration process needs to be re-examined. In this case, the RMS residual is 0.00047 g. The residual is systematic and is different in the range where the accelerometer was moved from upward to downward than in the range from downward to upward. If the associated bias is too high, the source of error needs to be identified. (In this example, the systematic residual is attributed to imprecision in the device used to slowly change the tilt of the accelerometer.)



This example only included simulated acceleration over a range from –1 g to 1 g. It is important to make sure the calibration of sensors includes the entire range of physical inputs for which valid readings are needed.

Do not use calibration by correlation for profilers.

Users of response-type systems calibrated them to a reference that was reproducible and stable with time. The reference value was a statistic defined for true profile. Usually, values of IRI calculated from measured profile were used. Calibration by correlation was used for response-type systems because their output could not be cast onto a comparable scale any other way. Response-type systems depended on the dynamic response properties of the vehicle in which they were installed.

Calibration by correlation is not needed for inertial profilers because the dynamics of the host vehicle are not a factor when they are functioning as designed. Vehicle motions are eliminated from the profile produced by a valid, working system. The instrumentation and electronics are calibrated separately, using accurate laboratory rigs and special test equipment at the factory.

If the outputs of your instrument do not agree with a profile-based reference, it's time to contact the manufacturer.

What Are System Checks?

A wide majority of road profile measurements are conducted on in-service roadways or within active construction sites. Road profile measurement requires the reliable operation of delicate sensors and electronics in harsh environments. System checks help detect and address issues with various sensors in profilers before time and effort is wasted gathering invalid or unusable measurements.

Consult the profiler user's manual and apply the recommended system checks regularly. Each system is different, and each manufacturer may recommend a different set of system checks or a different procedure for each type of system check. This section describes some highly recommended system checks for inertial profilers.

Maintain a log of all system checks.

Keeping a log of all regular system checks and on-site calibration activities helps other operators and data analysts determine the status of the profiler. A log also helps protect the credibility of past measurements whenever a new sensor or system failure is discovered and provides useful information for diagnosis. Some of the system checks described below involve collection of profile data under special conditions. Save the data and keep the raw files for those runs.

Establish a standard operating procedure that includes the details of when and how to perform calibrations, verification testing, and system checks. Many of the system checks described here should be performed as a standard part of a daily measurement routine.

Keep height sensor optics clean.

Inertial profilers typically use non-contacting height sensors that project light onto the pavement surface and detect the projected light with a specialized camera. It is important to keep the light source and the cover over the detector free of dust, mud, marking chalk, slurry, and other contaminants. Clean the sensors using the recommended solvents and wiping cloth, since scratches on the lenses or the windows that shield optical components may reduce the quality of the readings. Be careful to observe the recommended procedures for avoiding eye damage when working near lasers and clean the lenses with the laser power off.

Check the stand-off height and lateral sensor spacing.

Typical height sensors used in inertial profilers have a finite vertical range. It is best to mount the profiler hardware so that the road surface is near the middle of the height sensor range when the host vehicle is at rest. Checking the stand-off height is particularly important for inertial profilers that stow during transport, or inertial profilers that are not installed on a dedicated host vehicle. These are often called "portable profilers." Most portable profilers have provisions for adjusting the height as a part of the installation procedure.

Some procedures for inertial profile measurement specify a lateral spacing between the left and right wheel track. If the profiler hardware allows for adjustment of the lateral spacing, make the adjustment accordingly. Consistent lateral spacing between the measured wheel tracks (and consistent lateral tracking) is important when comparing measurements from different profilers.

Use gage blocks to test the height sensor calibration.

Any system, such as an inertial profiler, that uses non-contacting ranging sensors most likely obtains calibration values that were established by the sensor manufacturer or at a laboratory. For example, it is typically not necessary to calibrate height sensors while they are mounted to the profiler. However, it is very useful to check the calibration by verifying that the height sensors can accurately read a known change in height. This procedure is called a block test.

A block test is performed while the profiler is parked. Typically, the profiler's operational software includes a dedicated mode for conducting a block test. In most cases, the operator places a flat plate under the height sensor on a flat, level, and stable area, which is free of debris. A reading is collected on the base plate and used as a zero reference. Readings are then collected to observe the difference in height with gage blocks of various thicknesses placed on the base plate. The blocks are machined to very precise thickness. A typical set of gage blocks includes thickness values of 1/4 inch, 1/2 inch, 1 inch, and 2 inches. Note that height sensors have a limited operating range. If the top surface of the gage block is too close to the sensor housing, the reading will not be valid.

If the height sensor does not measure the thickness of the blocks to within the recommended tolerance, you may need to repeat the block test with more careful placement of the flat plate and gage blocks. You may also need to inspect the underside of the sensors for debris or scratches on optical components (with the power to the sensor off). If careful application of the block test procedure with clean sensors does not produce satisfactory results, repairs may be needed.

Perform the recommended accelerometer checks.

Some inertial profilers provide a method for field verification of the accelerometer calibration. The procedure typically involves inspection of accelerometer readings when the sensor is upright and again when the accelerometer is upside down. In that case, the sensor should read 1 g and –1 g, respectively. This may seem imprecise. However, for accelerometers that detect the influence of gravity, such as servo-type accelerometers, this is a robust way to make sure the sensor is well calibrated. The reason is that the error in the upright reference and upside-down reference is proportional to 1 minus the cosine of the misalignment error between the sensitive axis of the accelerometer and the direction of gravity. For example, 2 degrees of misalignment from vertical corresponds to an error of only 0.0006 g.

Calibrate your longitudinal distance measurement instrument (DMI).

Inertial profilers, and many other profilers, estimate longitudinal distance using encoders that record the rotation of one or more of the device's supporting wheels. Typical encoders read thousands of pulses per revolution, and changes in the pulse count must be converted to travel distance using a calibration factor.

The operational software of most profilers provides a module for calibrating the DMI in the field. To do this, lay out a section of known length and carefully mark the starting and ending point. A length of 1000 ft is typical. Measure the distance between the starting and ending points using the available mechanism for automated triggering. Typically, the operational software in the profiler will use the distance measurement and the known length to adjust the DMI calibration.

For profilers mounted on host vehicles with pneumatic tires, inflate the tires to a standard pressure. Warm up the tires by driving the host vehicle at the expected operating speed for the recommended length of time. (Tires warm up very slowly, and any warm-up that takes less than 20 minutes is most likely insufficient.) Keep in mind that the DMI calibration obtained on one type of pavement texture may not be perfectly accurate on another. Depending on the internal construction of the tire mounted to the wheel with the DMI, the calibration may also be sensitive to speed.

Calibrate the DMI as directed by the device manufacturer, and any time a change is made that affects the rolling radius of the tire(s) used to monitor longitudinal distance (such as installing new tires or moving the profiler to a different host vehicle).

Static collection of road profiles is an important way to test the health of a profiler in the field.

The operational software of many profiler designs provides a mode for conducting static testing. The operational software of every inertial profiler should include this feature. Static testing provides a way to identify a substantial proportion of system health issues and sources of measurement error before effort is wasted collecting invalid profiles.

During static data collection, the profiler remains still. The profiler simulates forward motion using an artificial signal in place of the DMI readings. Data are collected from the remaining sensors and are used to compute a "static profile" to help identify potential errors. Two types of static tests are common: quiescent and bounce.

Static testing requires a choice of simulated speed and distance.

Creating the artificial DMI signal usually requires the user to select a simulated forward speed for the device. Select a speed that corresponds to the expected operating speed or the middle of the expected speed range for production runs. The user must also select the duration for the run unless the profiler software enforces a standard choice. Some standards require data collection until at least 528 ft of simulated distance is accumulated. A simulated distance of 528 ft corresponds to 9 seconds of run time for an expected operating speed of 40 mi/hr. For a walking-speed profiler, several minutes may be needed to accumulate 528 ft of simulated distance.

A quiescent test measures noise in the system.

In a quiescent test, profile data are collected while the device is motionless. The quiescent test provides a way to assess the system noise (from power, signal conditioning hardware, sensors, etc., and their interaction) on the measured profile. Typically, the user assesses the noise level by computing the IRI of the quiescent profile. For inertial profilers, a typical expectation is an IRI value of 3 inches/mi or less, but values below 1 inch/mi are common. Comparing the results of many quiescent tests for changes over time is important. Any substantial change in the results should prompt investigation.

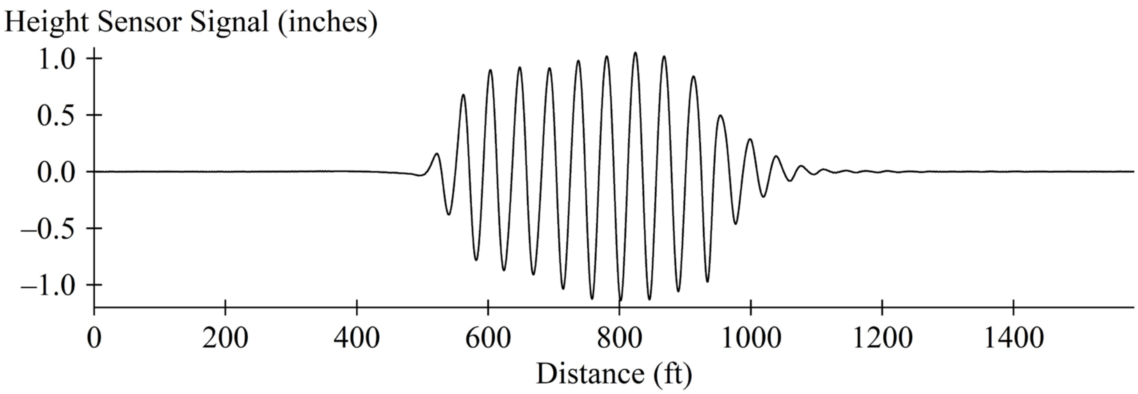
For inertial profilers, it is common to perform the quiescent test with the engine running and while the profiler instrumentation is powered by an on-board source. When unsatisfactory results are obtained, additional tests with the engine off may provide diagnostic information.

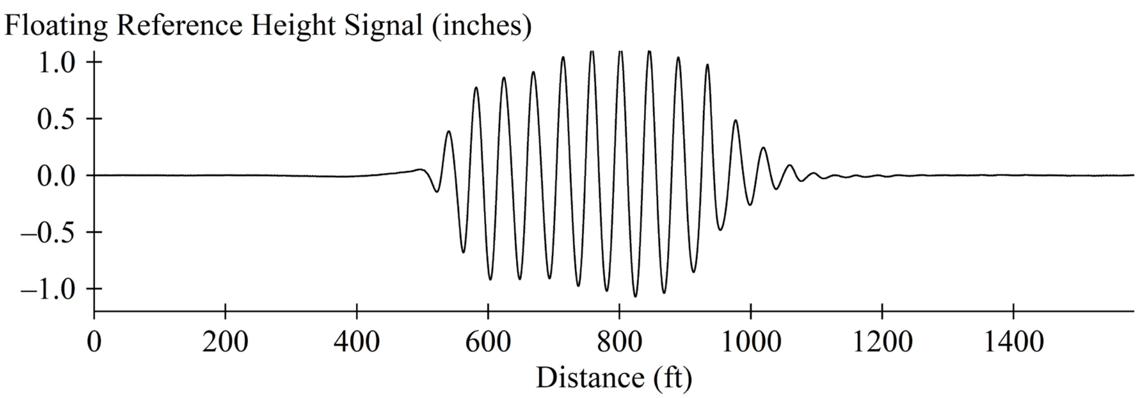
For a rolling inclinometer-based device, a quiescent test may characterize the contribution to profile measurement errors of inclinometer noise or drift. However, it omits the influence of out-of-roundness of the supporting wheels or the potential for contamination of the inclinometer readings during operation at inconsistent speed or with tilting.

A bounce test measures the consistency between the height sensor and accelerometer in an inertial profiler.

A bounce test is designed to simultaneously verify the quality of signals from the height sensor and accelerometer in an inertial profiler that is mounted on a host vehicle with a suspended chassis such as a high-speed profiler or a lightweight profiler. In a bounce test, profile data are collected while the host vehicle is parked. As described above, forward motion is simulated with an artificial DMI signal. Data from the other sensors are collected while the profile is exercised upward and downward by inducing a bouncing motion in the host vehicle.

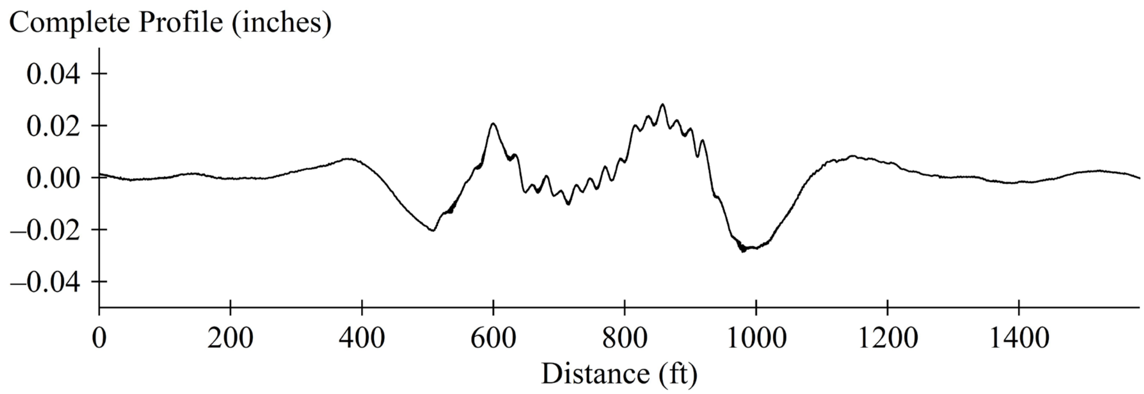
The following figures show the output from the height sensor and accelerometer of a profiler throughout a bounce test. The accelerometer signal has been integrated twice to show the floating reference height signal, which is the vertical motion of the profiler. Both signals have been high-pass filtered to remove drift.



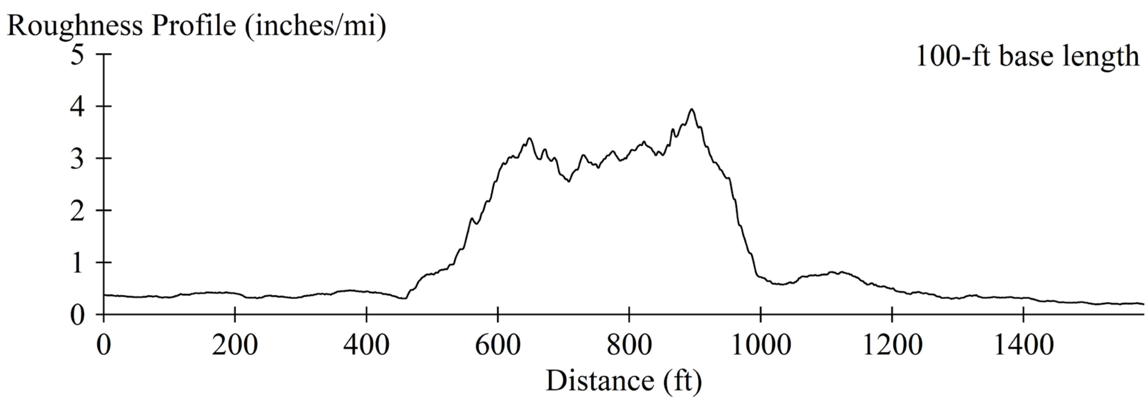


During the bounce test, the accelerometer detects the cycles of downward and upward vertical motion of the profiler. Simultaneously, the height sensors detect cycles in which the road surface becomes nearer and farther away from the profiler. Since the sensors remained above the same point on the road surface, the height sensor and accelerometer signals should combine to show no change in the road surface elevation. That is, the measured profile should be flat.

The next figure shows the profile, which was constructed by combining the height sensor signal with the floating reference signal. It is not perfectly flat. Some roughness is present, because of system noise, accelerometer drift, slight imprecision in the sensor gains, and errors caused during the bounce motion because the vehicle body did not remain perfectly level. However, the roughness is low, and the vertical scale of the plot is small relative to the amount of bounce motion.



Typically, the user assesses the quality of body motion cancelation in a bounce test by computing the IRI. The next figure shows a roughness profile for the bounce test computed using the 100-ft base length. Over the interval that included the bouncing motion, the IRI value was about 3.5 inches/mi. For inertial profilers, a typical expectation is an IRI value of 8 inches/mi or less. Monitoring changes in the IRI measured during a bounce test over time is important because a change may be a sign that something about the profiler has changed that needs to be addressed.



The example discussed here used a common procedure for conducting the bounce test. The simulated travel speed of 40 mi/hr was selected, because it represented the expected operational speed for the runs to follow. The test was conducted with a quiescent interval, followed by an interval with bounce motion, followed by another quiescent interval. The time of each interval was set to correspond to 528 ft of simulated motion. At 40 mi/hr, 528 ft corresponds to 9 seconds.

Other recommendations for conducting a bounce test include:

- Attempt to induce motion of the host vehicle body of about 1 inch in each direction.
- Attempt to induce true vertical motion of the profiler to the extent possible. (Minimize lateral or longitudinal motion and minimize roll.)
- Induce motion without putting force on hardware that is not designed to sustain the loading.
- Conduct the bounce test on a flat, level area of the pavement.
- Conduct the bounce test with a smooth, flat, non-reflective surface under the profiler sensors. Use of a clipboard is recommended.

An important function of the bounce test is to identify emerging problems with system health. These include damage to the sensors, unclean height sensor optics, degradation of electrical connections, electronic component failure, selection of incorrect filtering parameters, loose mounts, and mounting the height sensor too close to the limits of its measuring range. In addition, any design flaw that hinders the consistent measurement of vehicle motion by the height sensor and accelerometer causes roughness to appear in a profile measured during a bounce test. Examples include inconsistent signal timing, differences in the filtering applied to each sensor signal, poor mutual alignment of the sensors, inaccurate calibration of one of the sensors, and the use of a host vehicle with excessive dynamic motions in the range of frequencies that affect the IRI.

What Is Profiler Validation?

Independent validation of road profiler accuracy and repeatability is necessary to ensure data quality. An important feature of the process is examination of the profile, rather than simple comparison of IRI values or some other roughness index. Validation using profile is important for two reasons. First, it is possible to obtain the same roughness index value from two profilers through compensating error, even though the profile measurements themselves do not agree very well. In that situation, it is unlikely that roughness index values from two profilers will agree very well on a larger set of test sections. Second, valid measurement of the relevant details within the profile provides useful information about the pavement surface beyond the average roughness. That is, it supports the use of the methods in this book (and other publications) to diagnose sources of roughness, such as inspection of filtered profile plots, the use of short-interval roughness profiles for identifying localized roughness, and the use of PSD plots to identify repetitive roughness features.

A valid profiler is repeatable.

A valid profiler can obtain repeatable measurements of profiles and roughness index values. However, a profiler that produces repeatable profile and roughness measurements is not necessarily valid. The device could be reporting the same errors over and over. Even so, testing the repeatability of a profiler is an efficient way to identify many of the error sources known to affect accuracy. Repeatability testing is efficient, because it does not require other profilers to gather at the same site. Repeatability testing of inertial profilers should include multiple passes at each of two different speeds and using speeds as different as possible within the valid range for the device. This is because some error sources are very repeatable at one speed but affect the profiles progressively differently as the speed changes.

Valid profilers reproduce each other's outputs.

The need for obtaining reproducible profile and roughness measurements is an important motivation for validation testing. Reproducible roughness measurements allow us to compare results across different districts and agencies and over successive roughness surveys. Reproducible profile measurements help the community leverage detailed analysis methods for learning about their pavement surfaces.

Profilers are validated by comparison to the output of a reference device.

Validation testing typically requires a profiler to measure multiple test sections and reproduce measurements made by a designated reference profiler. The level of agreement to a reference profiler is often considered a measure of accuracy, even though no reference profiler is perfectly accurate.

It is important that the reference profiler captures the relevant information from the true profile with sufficient accuracy and repeatability. Repeatability can be established the same way as it is for any other profiling device. The issue of accuracy testing for reference profilers is an open research topic.

The purpose of the measurement may influence the device or method used to capture reference measurements. For example, if the purpose of the profiler under validation testing is to capture the IRI, the reference device must be valid in the wavelength range that affects the IRI. Many profiler validation experiments use rolling inclinometer-based profilers to measure reference profiles. When they are operated properly, inclinometer-based devices have the potential to measure the true profile in the waveband of interest for the IRI.

In many contexts, the IRI is used as a general measure of the effect of road roughness on vehicles, their cargo, and their occupants. In that case, use of a device that contacts the road surface with supporting wheels or feet may be a viable choice, so long as the supports envelop small asperities, such as texture, and bridge over narrow cracks or joints in a manner similar to a vehicle tire.

In other contexts, profile measurements are used to assess details about the pavement surface with very small dimensions. For example, a profiler used to measure joint faulting may also need to provide enough detail about the surface to identify the locations of the joints. For that purpose, a reference device that is insensitive to the depth of narrow dips by bridging over them might not be suitable.

Perform validation under the range of conditions that the profiler will encounter.

Different roughness levels and surface types challenge the performance of profilers differently. For example, a very smooth test section may challenge the resolution of the sensors on a profiler, because the signals will fluctuate over a small range. In contrast, a very rough test section may challenge the maximum range of the sensors.

Various texture types, cracks, or joints pose a challenge to some inertial profilers because of the way the height sensors interact with the road surface to determine its elevation. A typical example is the measurement of a pavement with longitudinal tines or grooves using a height sensor with a very narrow point of projected light. The profiler may demonstrate excellent agreement to an inclinometer-based reference device on a dense-graded asphalt section, but poor agreement on a longitudinally tined section.

Profiler certification is a formal version of profiler validation.

Profiler certification is a formal process used to deem a profiler suitable for a particular function, such as measurement of IRI for construction quality assurance. Typically, profiler certification expires on a regular schedule such as the end of each construction season. AASHTO R 56-14 and ASTM E950/E950M-22 address methods for profiler certification.

What Is Cross Correlation?

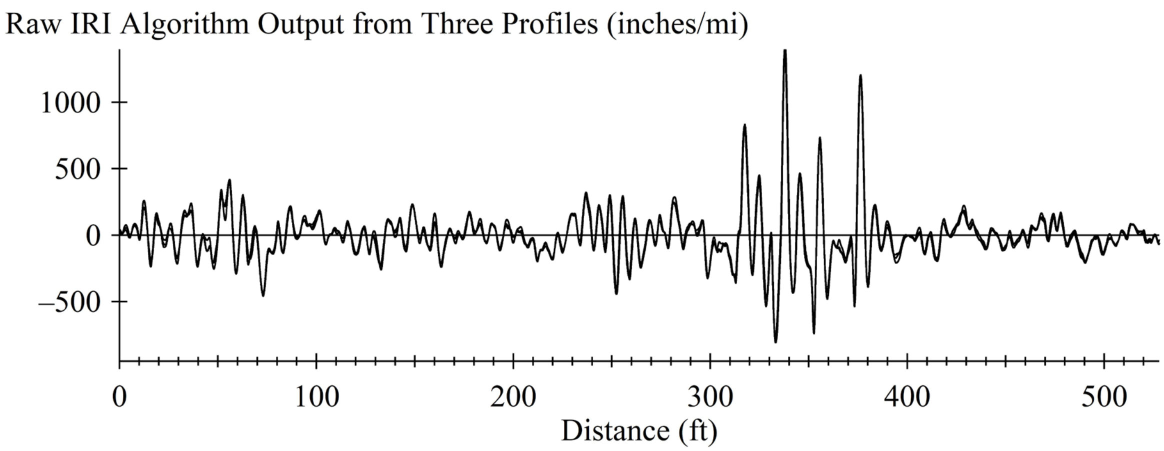
Cross correlation is a method used in the analysis of random data to examine the shared properties of two signals. Spectral analysis, cross correlation, and several other methods are all part of a standard set of techniques used in signal analysis that require substantial mathematical background. To examine that background, find a textbook on random data analysis, such as:

Bendat, J.S. and Piersol, A.G., *Random Data. Analysis and Measurement Procedures.* Second Edition. John Wiley and Sons, New York. (1986).

Cross correlation is one method of quantifying the level of agreement between two measured road profiles. This section describes a methodology for comparing measured profiles using cross correlation.

Compare the content from the true profile that is relevant for the application.

Recall the three raw profiles from the same test section shown early in this book. No two of those traces looked alike. (See "What Is a Profiler?") Once the profiles were filtered to remove fluctuations that took place over long distances, they looked exceptionally similar. (See "What Is Filtering?") A useful technique is to filter profiles to emphasize the waveband of interest before comparing them. If measurement of the IRI is of interest, transform the profiles using the IRI algorithm before comparing them. The next figure shows the same three profiles from the previous examples after application of the IRI filter.



In this case, the three traces are nearly identical after application of the IRI algorithm. That is due in part to the removal of long-wavelength content and the influence of differences in high-pass filtering procedures that affect the look of raw profile plots but do not affect the IRI. The plot above looks nothing like the raw profiles. In part, this is because the wavelength range is cropped to remove both long- and short-wavelength content that does not affect the IRI. In part, it is because the IRI algorithm converts the profile into units of slope, which is more closely tied to vehicle response.

Cross correlate the filtered traces to get an objective rating of the agreement between profiles.

The agreement score for any two filtered profile traces is a combination of a "shape" score and a "level" score. The shape score (ρ) is calculated as follows:

$$\rho = \frac{1}{\sigma_p \sigma_q} \sum_{i}^{N} p_i q_i$$

Where *p* and *q* are the two traces. The subscript "*i*" is the sample number within each profile. *N* is the number of points in the traces over the range where they are compared. The σ values are the standard deviation of the traces.

Dividing by the standard deviation of each trace limits the scale of the shape score to a value of 1 for perfect correlation and –1 if the traces are complete opposites. A shape score of 0 means the two traces show no correlation at all.

The point-by-point multiplication of readings increases the score where the two traces have the same sign and decreases the score where the two traces do not have the same sign. A more subtle property of the shape score calculation is that any inconsistency in the shape or placement of features between the two traces reduces the correlation level.

The level score (γ) guards against cases where the trace's shapes agree, but their magnitudes do not. That is, the level score penalizes differences in the overall roughness between the traces:

$$\gamma = \frac{\min(\sigma_p, \sigma_q)}{\max(\sigma_p, \sigma_q)}$$

If the standard deviation of the two traces is equal, the level score is 1. If not, the level score is the ratio of the smaller value of standard deviation to the larger value.

The agreement score is the product of the shape score and level score:

$$\text{Agreement Score} = \rho \bullet \gamma$$

A high agreement score means the measured roughness and the details that contribute to roughness agree.

The following table shows the mutual shape, level, and agreement scores for the three sample profiles and their IRI values.

Traces	Shape Score	Level Score	Agreement Score
Inclinometer-Based vs Inertial A	0.997	0.990	0.987
Inclinometer-Based vs Inertial B	0.990	0.982	0.972
Inertial A vs Inertial B	0.993	0.991	0.984

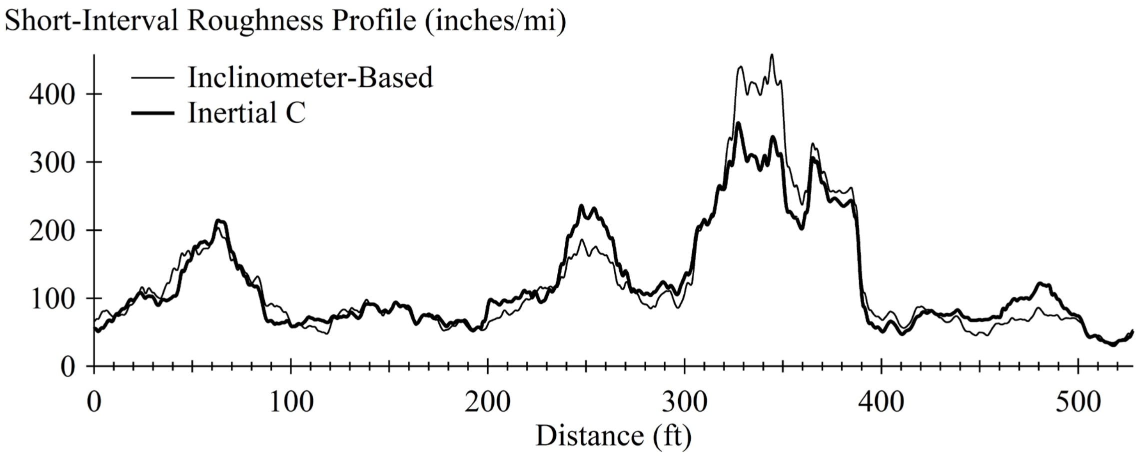
These scores are exceptionally high. Since the IRI-filtered profiles were used in this example, the implication of the high level scores is excellent agreement between the IRI values for each profile. The IRI of the profile from Inertial A was 123.5 inches/mi; the IRI of the profile from Inertial B was 125.7 inches/mi; and the IRI of the profile from the Inclinometer-Based profiler was 123.8 inches/mi.

The implications of the high shape score are consistency in the details of the profiles that affect the IRI and excellent agreement in the short-interval roughness profiles and identification of the severity localized roughness. The short-interval roughness profile for one of these profiles is shown in “What Is a Roughness Profile?” Like the plot above of raw IRI algorithm output, the three short-interval roughness profiles are nearly identical.

A low agreement score means the details that contribute to roughness do not agree.

Consider a fourth profile (Inertial C), which does not produce a high agreement score to the others. In a comparison to the inclinometer-based profile the level score is 0.957 and the shape score is 0.917, which is an agreement of 0.878. The overall IRI of the profile from Inertial C is 123.7 inches/mi, which agrees with the other profilers. However, the low agreement score, and the low shape score in particular, signifies that the distribution of roughness in the profile from Inertial C is not consistent with the others.

The figure below shows the short-interval roughness profile from the Inclinometer-Based profiler and Inertial C. Note the difference in the localized roughness severity near 330 ft.



Agreement scores can be used to quantify repeatability, reproducibility, and accuracy.

If a profiler makes two repeated passes over a test section, the agreement score between the two measured profiles is a measure of its repeatability on that section. Repeatability testing usually requires at least five passes over a test section. In that case, repeatability can be summarized as the average of the agreement scores from all ten of the possible comparisons.

The table above provides an example of using cross correlation to quantify reproducibility between three different profilers. However, if the measurement from the inclinometer-based profiler is designated as a reference profile, the agreement scores of each measurement from the inertial profilers are considered accuracy scores. Typically, accuracy scores are the average of the agreement scores from five or more repeat measurements by a profiler to the reference profile.

The algorithm used to apply cross correlation is complicated.

The example above used profiles that were already aligned longitudinally, had longitudinal distance measurement instruments that were in close agreement with each other, and were stored using the same recording interval. In practice, the algorithm for cross correlation must account for shortcomings in all three of these properties. For example, the typical algorithm for cross correlation resamples the profiles to a common interval before applying the summation for the shape score. The summation is usually applied multiple times with varying offset in one of the profiles to search for the alignment that produces the best agreement score. That is needed to remove small inconsistencies in the longitudinal offsets between profiles.

Details of the algorithm for producing agreement scores with cross correlation appear in:

Karamihas, S. M., "Development of Cross Correlation for Objective Comparison of Profiles." *International Journal of Vehicle Design*, Vol. 36, Nos. 2/3 (2004) pp. 173-193.

What Causes Profiling Error?

Longitudinal road profile measurement involves: (1) a user, (2) a profiler, and (3) a road.

Errors are caused by: (1) the user, (2) the profiler, and (3) the road.

We are profiling a different line each time.

A major source of variation in profile measurements is that the sensors follow a different line on the road in each pass, and the lines simply have different true profiles. For static profilers and profilers that operate at walking speed, the line is usually marked with chalk or paint, and this error is small. For inertial profilers, variations in choosing the line on the road are usually more significant than any other errors.

If we drive the profiler along a path parallel with the centerline of the road, there are two variables that locate the line being profiled:

1. the longitudinal starting position, and
2. the lateral position.

At 60 mi/hr, we cover almost 90 ft each second. The normal human reaction time corresponds to distances of 15 ft to 25 ft. In addition, when we trigger the profiler to start sampling, there may be a delay that depends on the design of the electronics and the computer software that controls the system during testing. The profile starts wherever the accelerometer and height sensor are located when recording of the sensor readings begins.

In modern profilers, the use of automatic triggering at landmarks and GPS position measurements improves the alignment of profile measurements to the desired road segment.

Lateral position and lateral sensor spacing are also important.

The path located directly under the height sensor defines the line covered by the profiler in each pass. In many profilers, the sensors are aligned laterally with the wheels of the host vehicle. In others, they are not. Differences in the spacing between the profiler sensors and the lateral position of the host vehicle within the lane are causes of both random error and bias in measured roughness.

Highly skilled drivers can maintain a lateral position within about 6 inches. However, variations of 12 inches or more are possible. Systems to assist the driver may improve accuracy and consistency in lateral tracking of profilers. However, they may distract the driver from the important job of operating safely on in-service roadways. At the time this book was written, the automotive industry was offering systems to assist with lane keeping. Some systems held the vehicle in a consistent lateral position within the lane, and others only intervened to keep the vehicle from crossing the lane boundaries.

The profiler has measurement error.

The valid measurement of profile requires that all parts of the system function correctly: the accelerometer, the height sensor, the speed sensor, the power supply, the electronic signal conditioning, the analog-to-digital converter, and the computer. Each part of the process involves some error. Ideally, the errors are small, but they do exist. In recent studies of profile measurement quality, the most common error sources were associated with the driver or operator. However, some unexplained errors remain, even in controlled experiments.

The sample interval may be too large for the analysis.

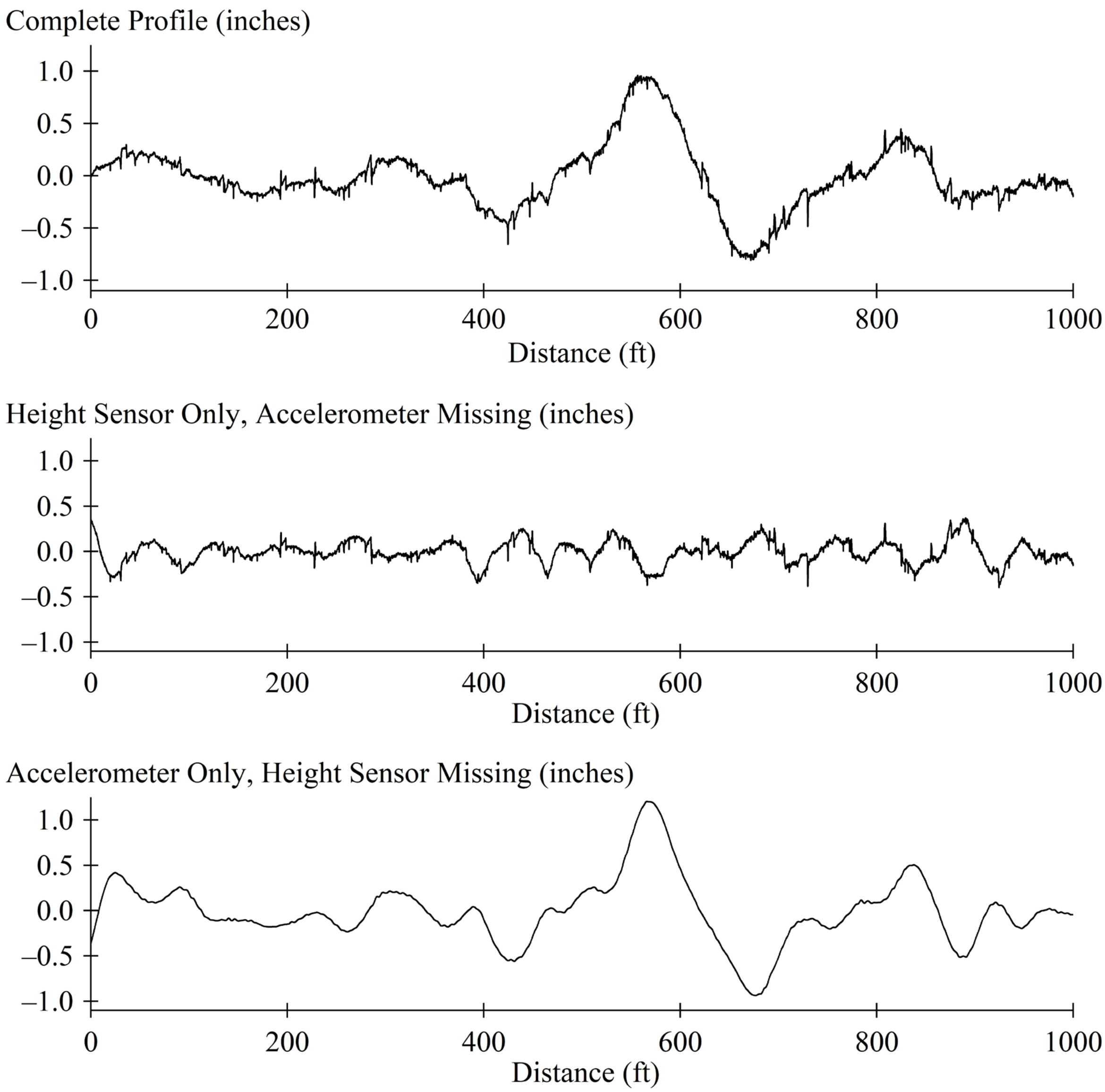
Most of the mathematical analyses applied to profile measurements are nearly exact for very small sample intervals but become approximate as the sample interval increases. In the past, users traded off the reduced equipment cost and data storage cost of using a longer recording interval with potentially larger errors in the roughness index values and other statistics calculated from the relatively coarsely sampled profiles.

Using modern sensors, computers, and data network infrastructure has reduced the pressure to limit the sampling rate and recording interval of profilers. However, the accuracy, repeatability, reproducibility, and relevance of a profile measurement depends on a complex combination of the data acquisition and filtering practices, as well as the specific way the height sensor determines the elevation of the road surface. A discussion of the effect of texture and cracks on profile measurement in a later section describes some of those issues.

A transducer may be disconnected or broken.

When a profiler malfunctions, it is often because one part has failed. One cause is a broken connection. Another is corruption of the power supply. In some cases, dust or mud obscure the optics used by a height sensor. If either the height sensor or accelerometer in an inertial profiler stops reporting valid readings, the profiler may still calculate and record a profile, but it will not be valid. Collecting a valid inertial profile requires valid signals from both sensors to cancel the influence of vertical motions of the host vehicle.

The plots below show a complete profile, the profile computed with the accelerometer signal missing, and the profile computed with the height sensor signal missing. The profile was measured at 66 mi/hr.



If one sensor's signal is absent from the profile calculation, the error in the profile may be difficult to detect for three reasons. First, using either signal alone will still result in a profile that fluctuates enough to potentially look like an accurate profile measurement in a plot. An inexperienced analyst may see either of the sample plots with a sensor missing and accept it as a valid profile. The height-sensor-only plot includes many of the most recognizable details, such as narrow dips at cracks and patches, which makes it difficult for an analyst to notice something is wrong.

Second, the IRI values of the profiles calculated with a sensor missing may fall within a credible range. In the example, the IRI of the complete profile is 91.2 inches/mi. The IRI calculated from the profile using only the height sensor signal is 107.5 inches/mi, and the IRI calculated from the profile using only the accelerometer signal is 61.5 inches/mi. Detecting the absence of either signal may require knowledge of the expected IRI value, or inspection of PSD plots by an analyst who knows what to look for.

Third, it is possible to get a repeatable measurement of a profile with the accelerometer or height sensor disconnected, even though it will be invalid. The sample profiles created with a sensor missing include the influence of vertical vehicle motions in response to the roughness of the road. If repeated passes are collected at the same speed, the vertical vehicle motions will be very similar in each pass. Errors caused by a missing or malfunctioning sensor are easiest to detect using repeated passes at different speeds.

The height sensor might not work for other reasons.

Of all of parts that make up an inertial profiler, the height sensor is the most likely to fail in difficult conditions. Laser and optical sensors should be checked to ensure that the lenses are clear and not covered with dirt, rain, or mud that would prevent the continuous viewing of the image projected on the ground. Height sensors may also be restricted in the range of ambient temperature and moisture they can tolerate. Some height sensors depend on favorable ambient light conditions or produce invalid readings on highly reflective surfaces.

Road contaminants, debris, and moisture will affect the readings provided by non-contacting height sensors. To the extent possible, obtain profile measurements when the surface is clean. Avoid collecting profile measurements when the surface is wet, and mark any measurement collected when the texture is submerged in moisture as invalid.

Measurement speed, measurement timing, and the interaction of the road surface with height sensors all create confounding issues in profile measurement.

The following sections of this book address those issues.

What Is the Effect of Speed?

High-speed inertial profilers are built on ordinary highway vehicles, such as a passenger car, van, pickup, or light truck. Lightweight profilers use all-terrain vehicles or golf carts as host vehicles, which operate at a range of speeds lower than typical traffic speeds. When they are not measuring profile, profiler host vehicles can be driven at the same speeds that would be used if the vehicle did not have the instrumentation on board. But when measuring profile, what is the effect of speed?

Most profilers are valid for a range of speeds.

Most inertial profilers now in use produce profiles that are valid even if the speed changes modestly during the measurement. The range of speeds depends on the profiler design and the use to be made of the data. The manufacturer typically specifies a range of valid measurement speed for a particular profiler. It is best to verify that a profiler produces valid results for the range of speeds at which it is expected to operate.

True profile is static.

Remember that the true profile is a property of a line on the road. It has no associated speed. If your profiler is valid for a given purpose, then its speed during measurement is not a factor.

Under no conditions should you consider applying a “speed correction factor.” If different results are obtained from your profiler at different operating speeds, it is not valid at one (or both) of those speeds. Operate it within its valid speed range, disregard results collected outside of the valid speed range or make the necessary repairs or upgrades to ensure it is valid over the range of speeds at which it is operated.

One will finish sooner if one goes fast.

Drivers of inertial profilers tend to figure this out quickly. Safely driving the profiler host vehicle should always take the highest priority. Furthermore, sacrificing data quality to improve measurement speed may be expedient, but it is not efficient.

It is harder to track an exact line at high speed.

In applications where the intent is to capture a specific profile, defined by a specific (imaginary) line on the road, lower speeds make it easier for the driver to accurately follow the correct line.

A significant source of variability in repeated measurements is variation in the lateral position of the profiler between runs. In one run, the path followed by the profiler might pass over the worst part of a localized surface defect, increasing the roughness. In the next run, if the profiler tracks slightly differently, the defect may appear less severe. In a third run, the profiler may follow a line that misses the defect altogether. All three measured profiles might be correct in the sense that the profiler obtained a valid profile for the path it followed in each case.

Overall, variations like this often average out. Potholes and other causes of roughness are scattered about the road. Roughness features omitted in one profile are made up for with additional features that are not present in other profiles on the same road. Unless close agreement between multiple profile measurements is needed, some variation of the lateral position of the profiler is acceptable.

The profiler senses longer wavelengths at higher speeds.

Recall that the inertial profiler has an accelerometer to sense vertical movement of the vehicle and establish an inertial reference. Ideally, the output of the accelerometer would be valid no matter how little the vehicle moves. However, three sources of error in accelerometer readings may lead to errors in the profile, particularly if they are large enough compared to the legitimate part of the signal. The first source of error is electrical noise in the sensor and the instrumentation system. The second source of error is limitations in the resolution of the readings captured from the accelerometer. The third source of error, which is discussed below, is fluctuations in tilt of the host vehicle body from true vertical as it pitches and rolls in response to roughness of the road.

The accelerometer and its electronics in a profiler must be set to handle the largest vertical accelerations that are anticipated in normal use. Profiler designers specify the overall range based on the expected response of the host vehicle at the profiler mounting location during travel over a very rough road at high speed. For accelerometers mounted on the bumper of a commercial vehicle, the required range may be as large as 2.5 g in both directions, for a total of 5 g. Note that a 1-g vertical acceleration is the level where objects in the vehicle bounce in the air. Noise in accelerometers and limitations on their resolution often grow in proportion to their overall range. This poses a problem if the profiler must collect valid measurements on a rough road at high speed as well as on a smooth road at low speed.

Consider the three example sinusoids introduced earlier. The vertical acceleration that each sinusoid causes under the tire of a passing vehicle depends on speed. The table below shows the relationship between wavelength, amplitude, and vertical acceleration for three travel speeds. For all three sinusoids, the acceleration they cause diminished by a factor of 16 when the speed is reduced by a factor of 4. The 200-ft long wavelength generates 0.02 g vertical acceleration at 100 ft/sec (68.2 mi/hr), but only 0.0013 g at 25 ft/sec (17.0 mi/hr). Although the longer feature has the highest elevation amplitude, it has the lowest acceleration amplitude and is the most vulnerable to contamination by measurement errors at reduced speed.

Wavelength (ft)	**Amplitude (inches)**	**Speed (mi/hr)**	**Speed (ft/sec)**	**Frequency (Hz)**	**Acceleration Amplitude (g)**
200	0.8	68.2	100	0.5	0.02
50	0.2	68.2	100	2.0	0.08
10	0.04	68.2	100	10.0	0.41
200	0.8	34.1	50	0.25	0.005
50	0.2	34.1	50	1.0	0.02
10	0.04	34.1	50	5.0	0.10
200	0.8	17.0	25	0.125	0.0013
50	0.2	17.0	25	0.5	0.005
10	0.04	17.0	25	2.5	0.026

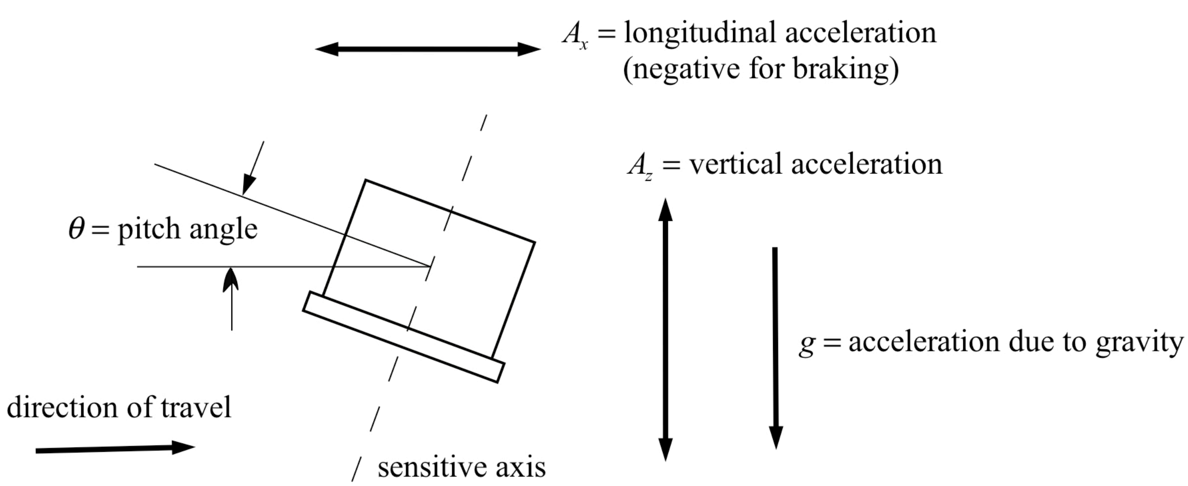
There is a low-speed limit for inertial profilers.

For very low speeds, the vertical acceleration diminishes to the point where errors submerge the legitimate part of the measurement over a substantial portion of the wavelength range of interest. A commonly specified low speed limit for high-speed inertial profilers used to measure the IRI is 20 mi/hr. The exact limit depends on several factors: (1) the intended use of the measured profiles, (2) the accelerometer specifications, (3) the quality of the instrumentation system, and (4) the dynamics of the host vehicle.

Using the brakes causes a measurement error.

Modern inertial profilers allow the speed to vary during the measurement process. The speed change is perfectly compensated if the accelerometer is oriented so (1) its axis is purely vertical, or (2) the longitudinal acceleration is zero. In fact, neither of these conditions is perfectly satisfied. For example, as the vehicle responds to roughness in the road, it pitches slightly forward and backward, changing the vertical axis of the accelerometer relative to pure vertical. The pitch angles are usually much less than one degree. However, during braking, the pitch can rise to one or two degrees. Except for braking or heavy use of the gas pedal, the longitudinal acceleration is also small.

A potential problem exists when the accelerometer is tilted while the host vehicle is accelerating longitudinally, as shown in the figure below. The figure shows tilting of the accelerometer due to tilting of the host vehicle body during braking.



The acceleration measured by the transducer (A_{meas}) is:

$$A_{meas} = (A_z + g)\cos(\theta) - g + A_x \sin(\theta)$$

Note that typical accelerometers used in profilers measure vertical acceleration with 1 g superimposed on it, because they are sensitive to the effect of gravity. The influence of gravity, and the tilt of the sensitive axis due to pitch, account for the first term to the right of the equal sign. The second term is present because 1 g is subtracted during processing to remove the effect of gravity before integrating the accelerometer signal. The third term is the erroneous sensing of longitudinal acceleration caused by tilt of the sensitive axis of the accelerometer.

Since the desired value of measured acceleration is equal to A_z, the error is:

$$A_{meas} - A_z = (A_z + g)\big(\cos(\theta) - 1\big) + A_x \sin(\theta)$$

For a 1-degree pitch angle, cos(θ) is 0.99985. For an actual vertical acceleration of 0.5 g, the error from the first term to the right of the equal sign is 0.00023 g. If the tilt was caused by 0.1 g of rearward acceleration during braking, the error from the second term is (–0.1 g)(0.0175), which is –0.00175 g.

Errors in profile caused by braking are proportional to the braking effort squared.

The severity of errors in profile caused by braking is roughly proportional to the braking effort squared. In the example above, braking for 1 second at 0.1 g causes a bias in elevation change of about 0.34 inches. Operating at 0.2 g deceleration with a pitch angle of 2 degrees increases the vertical acceleration error to –0.007 g, which causes errors four times as large.

The erroneous accelerometer readings caused by braking are temporary. Since the accelerometer signal is integrated twice, constant errors in the readings appear in the measured profile as errors in vertical curvature, leading to an error in the net elevation change during braking. The most concentrated errors occur where the brakes are released, because the artificial curvature in the profile changes over a short distance at that location. Often, the error in profile will register as localized roughness where the brakes are released. A good practice is to learn what level of braking causes enough error to justify marking the profile as invalid for each profiler.

Operating through a stop causes large measurement errors.

Profilers operating on urban and congested roadways inevitably need to stop during the measurement process. When a profiler comes to rest at a stop, the accelerometers are vulnerable to several small sources of error: misalignment of the sensitive axis with true vertical, minor errors in calibration, a difference between the actual force of gravity and the assumed value used to remove the offset, drift caused by temperature fluctuations, system noise, and limitations in the resolution of the readings. The overall bias in the accelerometer readings during the stop is mostly associated with the first four error sources listed above. However, they are difficult to predict or eliminate without using additional sensors.

When the profiler is moving at a typical traffic speed, slowly changing errors in accelerometer readings typically cause only small errors in profile. The slowly changing accelerometer errors translate to fluctuations in elevation that spread out over long distances and affect roughness indices like the IRI very little. High-pass filters often remove bias errors altogether. However, when the profiler is stationary, the effects of bias errors and slowly changing errors appear as concentrated changes in elevation at the location of the stop. For example, a constant accelerometer reading of 0.001 g at the stop that lasts 30 seconds translates to a change in elevation of about 174 inches.

The magnitude of the artificial elevation change in an inertial profiler measurement is difficult to predict. Typically, the result is localized roughness with a high magnitude at the location of the stop. Some profilers apply filters that spread out the influence of highly localized sensor errors over a larger area around the location where the error occurred. The concentrated elevation change at stops is a common example of this phenomenon. The contaminated range depends on the processing and filter algorithms used to calculate the profile. A good practice is to determine the range within the profile that should be marked as invalid around each stop for each device.

Augmented inertial profilers and other technology can help avoid errors caused by speed.

Some vehicle-mounted profilers have been designed and implemented that do not depend on an accelerometer-established inertial reference. Those designs are not vulnerable to the error sources described above for operation at low speed, during braking, and through stops. However, they may be vulnerable to error sources that do not affect inertial profilers.

At the time this book was published, several manufacturers of inertial profilers had developed methods to update the standard inertial profiler design to mitigate or eliminate measurement errors at low speed, during braking, and through stops. The updates included additional sensors to overcome accelerometer tilt and measure true vertical acceleration, or additional sensors to switch to a different measurement principle for operation at very low speeds and during stops.

What Are the Effects of Texture, Cracks, and Joints?

Textured surfaces may cause aliasing errors in profile measurements.

Coarse-textured surfaces such as chip-sealed or open-graded pavements pose difficulties for some profilers. Texture involves asperities that occur over short distances that correspond to wavelengths of several inches or less. These wavelengths are outside the range of interest for most profile analyses, but they affect the performance of some instruments. The specific nature of the error depends on the type of instrument and the type of texture.

Recall the discussion of aliasing in the "What Are Sine Waves?" section of this book. For profile shapes other than a sinusoid, the same effect exists. Very short wavelength variations in the true profile cause an aliased shape in the measured profile if the sample interval is not sufficiently small.

Aliasing can be reduced for static and rolling profilers.

Inclinometer-based systems that contact the road surface with wheels or supporting feet reduce some of the aliasing effect. Supports that cover an area larger than the distance between peaks and gaps in surface texture are less prone to aliasing errors. The actual reduction in aliasing depends on the "footprint" of contact with the ground. Averaging or smoothing out of the influence of texture by the length or width of the contact area helps avoid aliasing. For some applications, bridging over narrow gaps or dips within the contact area helps remove content that is less relevant to ride quality than other features. On the other hand, when a device makes unyielding contact and rests on the highest point under its footprint, it introduces another form of aliasing into the profile.

Aliasing can be reduced for systems with non-contacting sensors.

Height sensors on inertial profilers can operate at high sampling rates. For example, a common laser used as profiler height sensors in the past produced 16,000 readings per second. At highway speed, the readings were less than 0.1 inches apart.

Applying a low-pass (smoothing) filter removes the high-frequency content in a signal before the digitization process to protect against aliasing. Recall the discussion of Nyquist in the "What Are Sine Waves?" section of this book. To avoid aliasing and capture the content in a signal, all of the content at a frequency of half the sampling rate is required. At highway speed, a sampling rate of 1,000 Hz corresponds to about one sample per inch. At that sampling rate, a low-pass filter with a cut-off frequency of 500 Hz or lower is required. Since filters are not perfect, using a cut-off that is a quarter or a fifth of the sampling rate is a common recommendation.

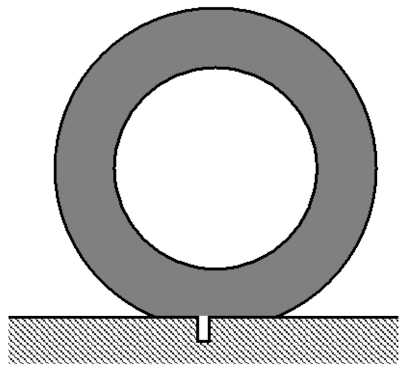
Directional textures cause problems for some height sensors.

Some laser height sensors project a very small dot onto the road surface. Each reading detects the height of a specific point on the road surface. If the profiler is tracking over a texture type with longitudinal channels, such as longitudinal tining, longitudinal grooving, diamond grinding, or a milled surface, the dot of projected light may wander into and out of the channels as the host vehicle passes over the pavement section. The true profile for the narrow, but wandering, slice of the road will include dips in the intervals where the height sensor detected the channels. This profile misrepresents the ride quality of the pavement because the tires of passing vehicles envelop the grooves.

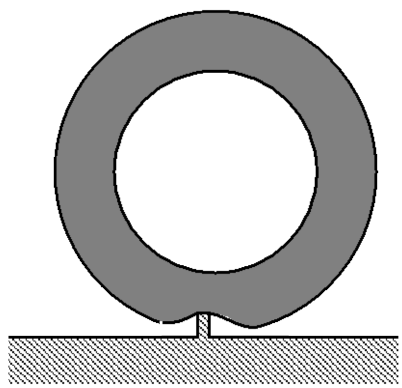
Height sensors with a wider footprint provide a more useful representation of the road surface height on pavements with longitudinal textures. Typical wide-footprint height sensors project (and detect) a line of light that is 4 inches or more wide on the road surface. Averaging the height within a wide footprint improves the relevance of the measured profile to vehicle response. On some textures, including longitudinal grooving and open-graded asphalt, the lowest portions under passing vehicle tires never contact the tread. In such cases, eliminating the influence of the lowest readings within a wide footprint is recommended.

Cracks and joints do not directly affect vehicle dynamics.

Although the presence of cracks in the road implies that the road is deteriorating, and may lead to increased roughness in the future, they are not necessarily "felt" by a vehicle. This is because they are usually very narrow compared to the length of a tire contact patch. The same is true of saw-cut joints on a concrete pavement.



Cracks do not directly affect tires and vehicles.



Bumps do.

So long as no change in elevation exists across a crack or joint, such as faulting, passing vehicles will bridge over them.

Cracks and joints do affect roughness indices.

Many height sensors project an image onto the ground that is small enough to go into a crack or joint. A short recording interval is needed to distinguish a crack or joint from a dip with a longer duration. Systems with a short sample interval are more likely to pick up a crack but provide enough information so that they can be identified as such. Often, however, this information is thrown away before the profile is recorded.

None of the profile analyses covered in this book include methods for handling cracks properly. They all treat a crack the same way as a bump that is as high as the crack is deep. This means that a road profile feature that is not relevant to the quality it is trying to predict can affect the index significantly.

Conventional filters do not completely remove cracks from the profile.

A low-pass filter that is applied to a profile that includes a large dip caused by a crack will reduce it but not remove it. A special "directional" algorithm, called a bridging filter, is needed that treats cracks differently than bumps. The algorithm will only be useful if the sample interval is close enough to identify a dip in the profile as a crack or joint.

Does the True Road Profile Change with Time?

Yes.

Changes in profile over time are a major motivation for collecting roughness measurements.

Pavement network engineers use periodic roughness measurements to identify sections of roadway experiencing degrading ride quality or where an increase in roughness may be a sign of structural degradation.

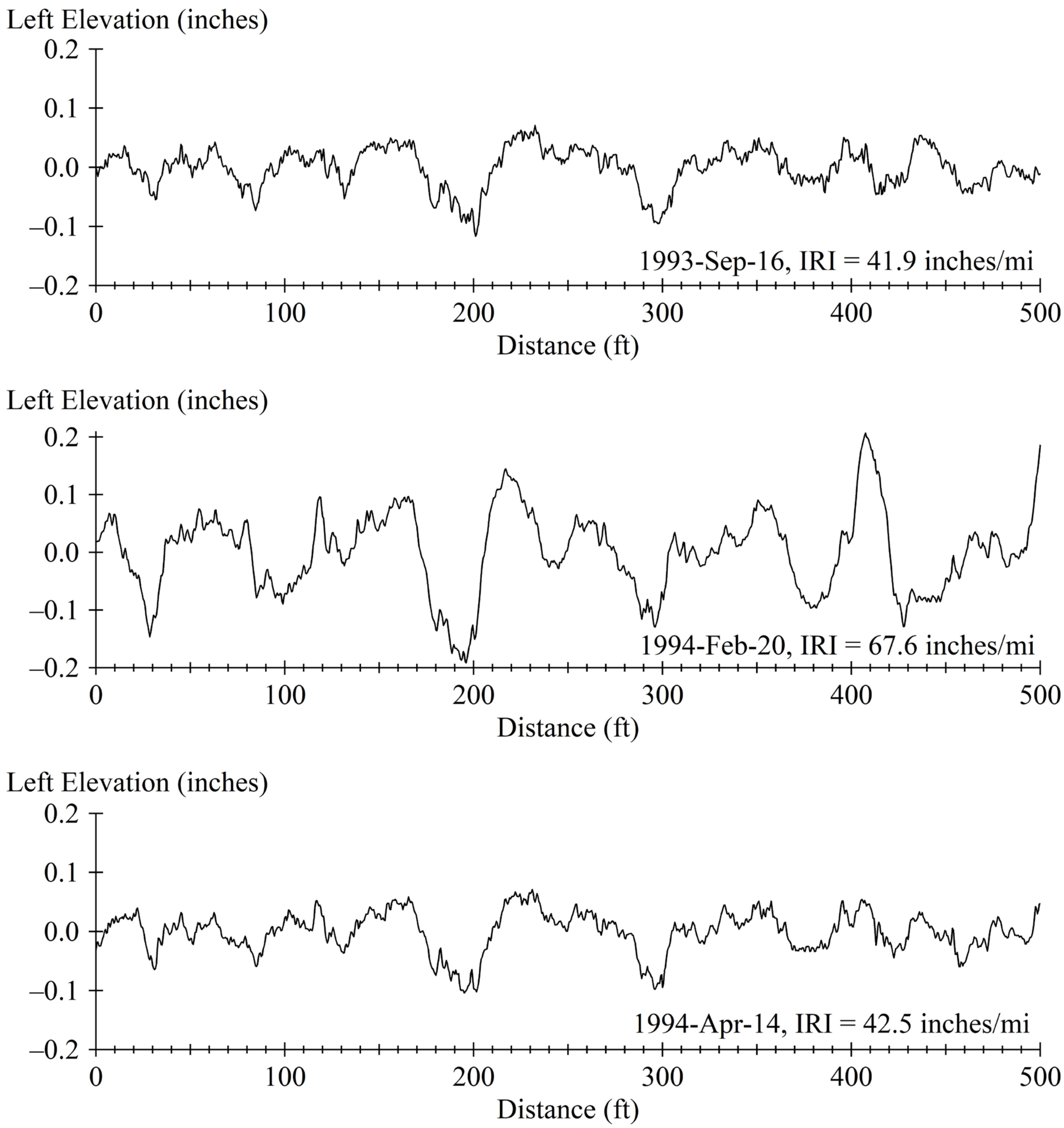
One motivation for using a profile-based roughness index for construction quality control is the opportunity to relate the roughness of a road segment at the time of construction to its progression over time. This provides a "cradle to grave" record of roughness on a consistent scale.

At the time this book was written, very little public data were available that examined the short-term changes over days or weeks in road profile and roughness in the earliest portion of a pavement's life. In asphalt pavements, compaction of the asphalt material by the traffic or settlement of the surface may cause changes to the profile. Changes in jointed concrete pavement due to early changes in slab strength or joint functionality are likely.

Typical pavement network surveys of road roughness take place annually or biennially. In some cases, roughness of pavement sections progresses so fast after the onset of structural failure that intervention is required before problems are detected in the regular network roughness survey. In other cases, emerging localized roughness appears in a network survey that requires attention but may go unnoticed until its contribution to segment-wide roughness is very large.

Seasonal changes in temperature and moisture affect the profiles of some pavements.

One example of seasonal changes to roughness is associated with changes in the volume of subsurface layers caused by very cold temperatures. The following figures show profiles of an asphalt pavement test section in New Hampshire from the LTPP experiment measured in three consecutive seasons from fall 1993 to spring 1994.

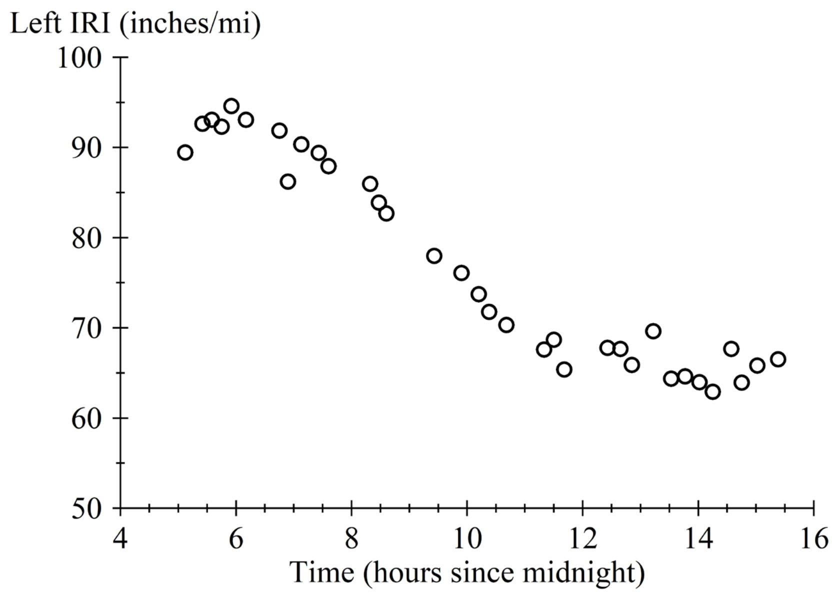


The section was a thin asphalt concrete pavement with a subgrade of poorly graded sand with silt. The profiles are similar in September 1993 and April 1994. The IRI values are 41.9 inches/mi and 42.5 inches/mi, respectively. In February 1994, the severity of several bumps and dips increases and other rough features appear that are not present in the other profiles. The IRI in February 1994 was 67.6 inches/mi. The likely cause of the temporary change in profile was frost heave.

Diurnal changes in temperature and moisture affect profiles of some pavements.

The profile of jointed concrete pavement slabs may exhibit upward curvature, in which the ends deform upward relative to the center, or downward curvature, in which the center deforms upward relative to the ends. The prevailing level of curvature depends on a complex combination of material properties, construction practice, and ambient conditions during placement and curing. Daily (and seasonal) changes in temperature and moisture cause changes in the temperature and moisture profile throughout the depth of the concrete pavement, which affects the magnitude and direction of slab curvature.

The figure below shows a set of IRI measurements collected over a 10-hour interval on a jointed concrete pavement in Michigan. The pavement section had doweled joints spaced 15 ft apart. The measurements were collected on a sunny spring day following a cool, clear night. The series of measurements occurred during a gradual change in air temperature from 38 degrees F to 74 degrees F. The profile of this pavement included net upward slab curvature in the early morning. Throughout the 10 hours that followed, heat from the sun progressively changed the temperature gradient so that the upper surface of the slabs accumulated more heat than the underside. The result was a reduction in upward curvature of the slabs throughout the section, and a commensurate decrease in roughness.



Variations in profiles due to changes in the true profile are, of course, not errors. The profiler is just capturing the current profile. Changes in profile with time confound the interpretation of roughness index values when the profile is not measured frequently enough. In addition, it is not always easy to determine if observed changes are in the true profile or in the measurement process.