Enhancement of Network Level Macrotexture Measurement Practices through Deterioration Modeling and Comparison of Measurement Devices for Integration into Pavement Management Systems

Kyle F. Maeger

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Master of Science
In
Civil Engineering

Gerardo W. Flintsch, Chair
Kevin P. Heaslip
Samer W. Katicha

November 12, 2018
Blacksburg, VA

Keywords: Pavement, macrotexture, surface properties

Enhancement of Network Level Macrotexture Measurement Practices through Deterioration Modeling and Comparison of Measurement Devices for Integration into Pavement Management Systems

Kyle F. Maeger

ABSTRACT

This research sought to integrate measurement and prediction of surface macrotexture for use in pavement management systems. This was achieved through two experiments, the first modeled the behavior of a binder-rich stone matrix asphalt when subjected to traffic loading using a heavy vehicle simulator to report the effect on pavement macrotexture. The second experiment compared high-speed macrotexture measurement devices on a variety of surfaces and under various operating conditions. The change in macrotexture due to traffic loading showed that as the cumulative load increased, the macrotexture decreased due to bleeding on the pavement's surface. A regression model determined that, on average, the macrotexture's root mean square (RMS) decreased 0.14 mm per million equivalent single axle load applied. A comparison of RMS and mean profile depth (MPD) outputs indicated that RMS was more sensitive to changes in macrotexture due to bleeding. In comparing devices, pairwise device agreement was evaluated using a Limits of Agreement. The results demonstrate good repeatability for each of the devices tested. The agreement analysis showed that not all high-speed devices can be used interchangeably for all pavement surfaces. Data acquisition speed was found to be a factor in macrotexture parameter calculation for two of the devices. The effect of speed was found to be worse on randomly textured surfaces than on transversely textured surfaces.

Enhancement of Network Level Macrotexture Measurement Practices through Deterioration Modeling and Comparison of Measurement Devices for Integration into Pavement Management Systems

Kyle F. Maeger

GENERAL AUDIENCE ABSTRACT

This thesis sought to integrate the collection and prediction of a pavement surface property known as macrotexture for use in the management of pavement networks. This was achieved through two experiments, the first of which modeled the behavior of asphalt concrete with a higher than typical asphalt content when subjected to simulated traffic to determine the effect on pavement macrotexture. The second experiment compared five high-speed macrotexture measurement devices on a variety of pavement surface types and under various operating conditions. The change in macrotexture due to traffic loading showed that as the cumulative traffic increased, the macrotexture decreased due to the asphalt coming out on the surface, referred to as bleeding. For the comparison of measurement devices data were processed using current industry standards. The results demonstrate good repeatability for each of the devices tested. The analysis also showed that not all high-speed devices can be used interchangeably for all pavement surface types. Vehicle speed was found to be a factor for two of the devices. The effect of speed was found to vary by surface type. Finally, vehicle acceleration was shown to influence the parameters produced by one of the devices, demonstrating that care should be taken to gather high-quality datasets for the critical pavement characteristic of macrotexture.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to the faculty of Virginia Tech, in particular my advisory committee, Dr. Gerardo Flintsch, Dr. Kevin Heaslip and Dr. Samer Katicha who have provided guidance along the way to completion of this thesis. I would also like to thank the leaders and Civil Engineers in the United States Air Force that have enabled me to be at Virginia Tech including fellow contributor to this research Vinnie Bongioanni. Finally, I would like to thank my wife Heather and my kids Clark and Marilou for their love and support of my education and career goals.

ATTRIBUTION

Authorship for the journal manuscripts presented in chapters 3 and 4 was shared with Vinnie Bongioanni. The delineation of contributions is presented in Appendix W.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

The pavement management process requires database inputs of both pavement condition measurement and pavement condition prediction. Of the various condition parameters developed, surface macrotexture is one that is not widely collected or utilized in pavement management systems (PMS) in the US.

The macrotexture of a road surface is an important parameter to understand how vehicle tires interact with the pavement. In particular, the macrotexture influences key driver safety issues including the friction of both wet and dry pavements which in turn can affect a driver's ability to stop efficiently, as well as the effect on splash and spray which can limit driver visibility (*1; 2*). Other effects of pavement texture include noise, both inside and outside the vehicle, and rolling resistance (*1*).

Because pavement texture influences these road characteristics, it is of great value to pavement engineers and managers to be able to collect and utilize macrotexture data at the network level. While there are standards within the US and abroad that govern the calculation of various texture parameters, there is little guidance for network-level macrotexture collection and processing. Furthermore, to fully utilize macrotexture data collected at the network level it should be integrated into a PMS. This requires an understanding of how macrotexture fits into current pavement deterioration models and the ability to predict its performance.

There is a wide array of devices and techniques available to measure macrotexture, including both dynamic and static types. Dynamic devices collect data in motion and vary in operation from manually pushed walking speed devices to vehicle mounted devices that can operate at highway speed. Static measurement methods require hand-placement and may be more time consuming and labor intensive to produce a smaller amount of data. Furthermore, the variety of devices and techniques may not yield the same results, which makes determination of the true value of macrotexture a challenge.

Previous research has been done on how macrotexture will change on pavement surfaces over time with traffic. However, the research is not comprehensive for all surface types and pavement distresses. This limits the ability of pavement deterioration models to incorporate macrotexture as an input parameter.

Problem Statement

There is no standard for network-level macrotexture data collection and application in the US. The variety of devices used to measure macrotexture can may yield different results, which makes comparing them difficult in the absence of a known true value of macrotexture. Furthermore, it is not clear how macrotexture will change with time and traffic. Modeling how macrotexture changes enables pavement managers to predict and plan for preserving and maintaining pavements with adequate macrotexture. The optimal integration of macrotexture into network level PMS requires that it be collected and processed to generate parameters that are accurate and repeatable.

Objective

The overall objective of the thesis is to enhance network level pavement macrotexture measurement practices for use in a PMS. This will be achieved through the following secondary objectives:

1) Contribute to the current body of knowledge concerning the effects of traffic on pavement macrotexture. This study seeks to model the behavior of binder-rich Stone Matrix Asphalt (SMA) surfaces, when subjected to traffic loading using a Heavy Vehicle Simulator (HVS) to isolate the effect of traffic from other environmental effects and report the effect on pavement macrotexture.
2) Evaluate the ability of various macrotexture measurement devices to measure typical pavement surfaces to determine if they can be used interchangeably within an acceptable range. As network-level measurements are envisioned to be made on a reoccurring basis, the ability of a device to produce consistent results on given surfaces (its repeatability) is also studied. Finally, since data collection does not occur in an ideal world, the effects of operational factors such as speed, acceleration, and deceleration on the measurements are also evaluated.

Significance

The significant contributions of this research are presented in two journal manuscripts (Chapters 3 and 4).

Chapter 3 presents the results of an experiment that observed the change in macrotexture of a binder rich stone matrix asphalt surface under simulated traffic loading. It is expected that pavement bleeding will decrease macrotexture as binder fills in voids between aggregates, however, the rate of change with traffic and the most appropriate parameter to observe this change is not known. Modeling this behavior enables prediction of when a pavement will reach a predetermined trigger value for maintenance or rehabilitation.

Chapter 4 presents the results of an experiment that evaluated macrotexture measurement devices on a variety of surfaces under various operating conditions. The reproducibility and repeatability of various macrotexture devices has not been widely assessed nor has the effect of various operational factors likely to occur in data collection at traffic speed. This knowledge is necessary for the development of protocols for network level macrotexture collection.

The knowledge garnered from this research of modeling the change in macrotexture and the evaluation of macrotexture measurement devices supports the primary pavement management inputs of condition prediction and condition measurement respectively. While the prediction of future condition and measurement of current condition have value to pavement managers separately, together they can be used in a PMS to optimize budgets and select best alternatives for project level decisions (*3*).

Thesis Organization

This thesis is presented in five chapters. The first chapter is an introduction to the research. Chapter 2 presents the necessary background for this thesis that was researched in the literature review process. Chapters 3 and 4 are adapted from journal manuscripts and represent the main body of research presented in this thesis. Finally, Chapter 5 presents a summary of the work, engineering and scientific significance, conclusions and recommendations for future research.

References:

[1] ISO. ISO 13473-2 Characterization of pavement texture by use of surface profiles. Vol. Part 2 Terminology and basic requirements related to pavement texture profile analysis.

[2] Hall, J. W., K. L. Smith, L. Titus-Glover, J. C. Wambold, T. J. Yager, and Z. Rado. Guide for Pavement Friction. 2009, p. 257p.

[3] Shahin, M. Y. *Pavement Management for Airports, Roads, and Parking Lots*. Kluwer Academic Publishers, 1994.

CHAPTER 2: LITERATURE REVIEW

To ensure this research is unique and builds upon the existing body of knowledge related to macrotexture, a thorough literature review was conducted. The topics reviewed include pavement management, texture wavelengths and their effects on vehicle-road interactions, the methodologies of collecting and calculating macrotexture, previous research on macrotexture changes due to traffic, accelerated pavement testing, previous macrotexture equipment comparisons, and a methodology for comparing measurement devices.

Pavement Management

Roadway pavements are complicated systems that must be properly maintained to provide adequate transportation for the public good. Construction and maintenance of roads represents a significant portion of public fund expenditures, mandating that pavement managers strive to provide acceptable levels of service while seeking the maximum possible benefit from investment. Figure 1 shows the general pavement management process. Fortunately, modern computing technology enables pavement managers to the handle the complexity of inputs to the decision-making process through software packages known as Pavement Management Systems (PMS). Once the pavement network is defined and a database is created, the two key elements that influence decisions are the measured condition of the pavement and prediction or modeling of the condition. These elements are reflected at the network level to develop project priorities, and at the project level to determine the best course of action. Periodically collecting pavement condition information at the network level ensures that the information in the database is accurate. It also important to have accurate models to predict the pavement condition over time. This enables the optimal timing of maintenance actions to maximize the benefit. A PMS can determine benefit by modeling the deterioration of a condition index to failure over time, estimating the improvement to that index by a maintenance action, and modeling the condition index deterioration after this maintenance action. The area under the curve created by the maintenance action is the benefit area, where the goal of a PMS is to maximize this area through the optimal timing of a maintenance action (*1*). This is illustrated for two consecutive maintenance actions in Figure 2.

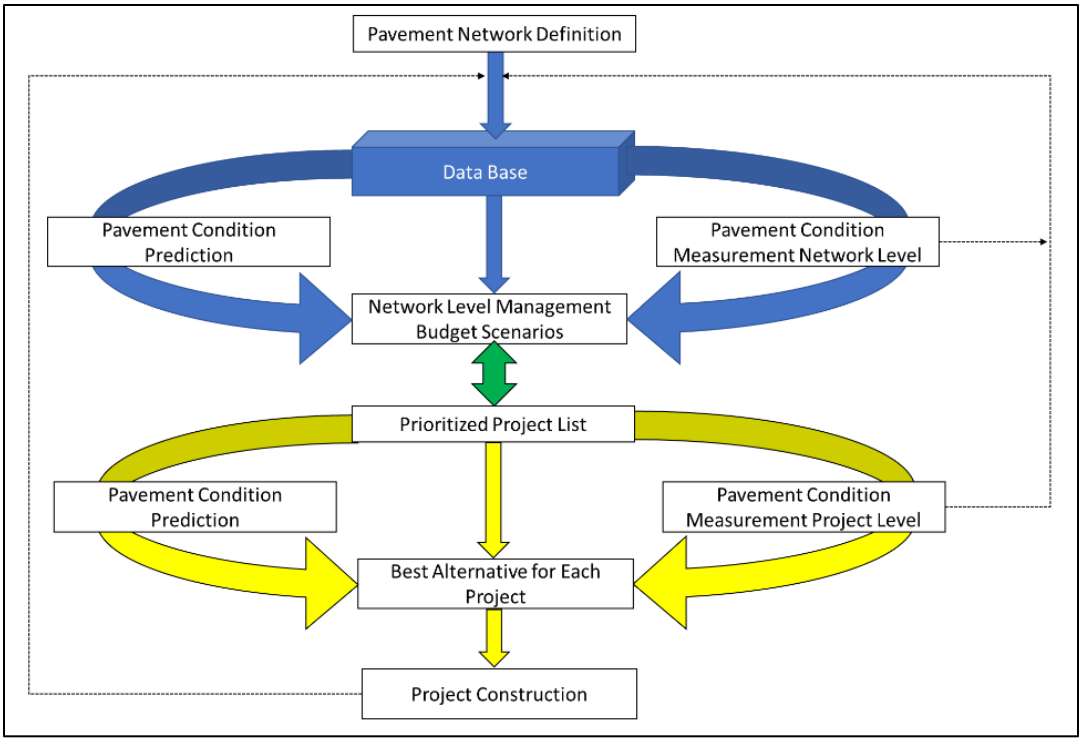


Figure 1 - Pavement Management Process (adapted from Shahin (*2*))

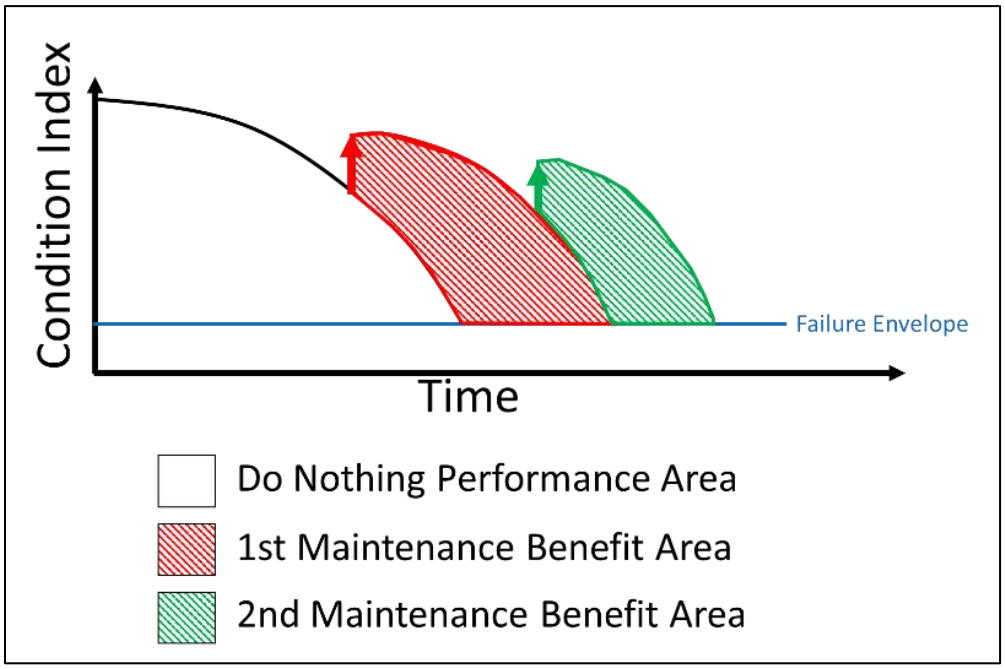


Figure 2 - Benefit of Maintenance (adapted from Peshkin et al. (*1*))

There are a variety of pavement distresses that agencies collected for use in pavement management. NCHRP conducted a survey of the distresses that agencies collect for pavement management. The survey showed that the primary distresses collected across agencies include: rutting, transverse cracking, fatigue cracking, longitudinal cracking, block cracking, raveling, faulting, spalling, bleeding, edge cracking, punch-outs, shattered slab, durability cracking, pumping, and other unspecified distresses with rutting, fatigue cracking, transverse cracking, longitudinal cracking and bleeding as the most commonly collected. In addition to these distresses agencies may also measure the roughness as an indicator of ride quality, conduct structural evaluations and collect surface friction data which can include macrotexture (*3*).

Figure 3 shows the percentage of agencies in the US and Canada that collect different types of condition data at the network and project level. At the network leve, surface distresses and smoothness are collected by most agencies, while structural capacity and surface properties are not as widely collected. Conversely, since project level data is typically more detailed and specific, surface friction properties (which may include macrotexture) and structural capacity are more prevalent at the project level among agencies (*3*).

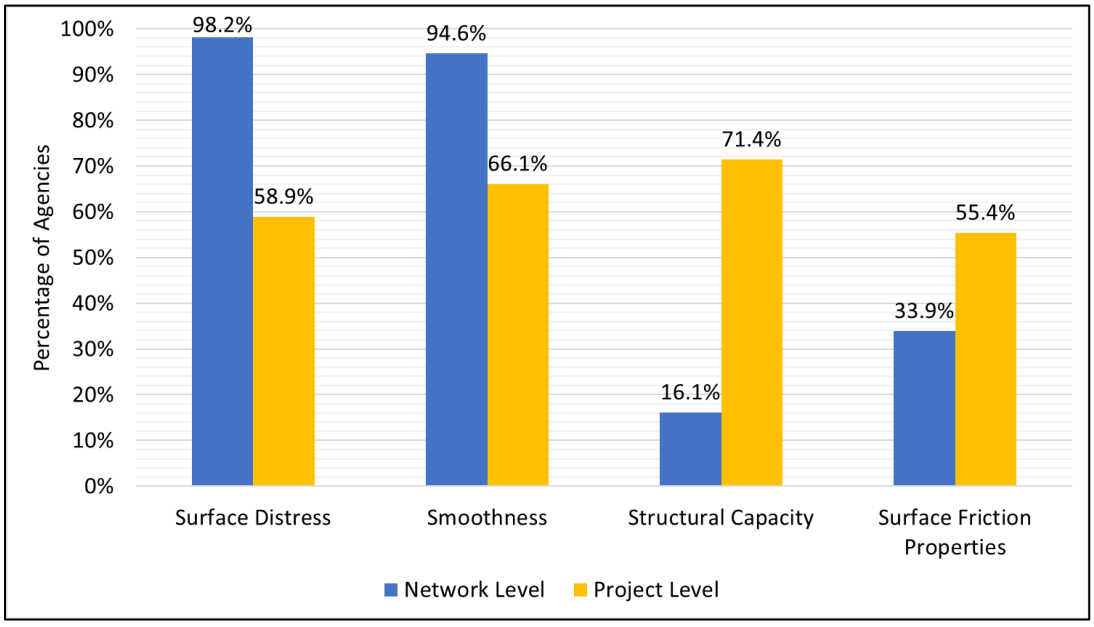


Figure 3 - Pavement Condition Collected by Agencies (Adapted from NCHRP (*3*))

While macrotexture is not universally collected and utilized, it has been implemented in some locations. A survey of state highway agencies was conducted by NCHRP to determine which states were collecting macrotexture at the network and project levels, the results of which are shown in Figure 4 and Figure 5 (*4*). Since this survey was conducted the Virginia Department of Transportation began collecting network level macrotexture through a contract with Virginia Tech.

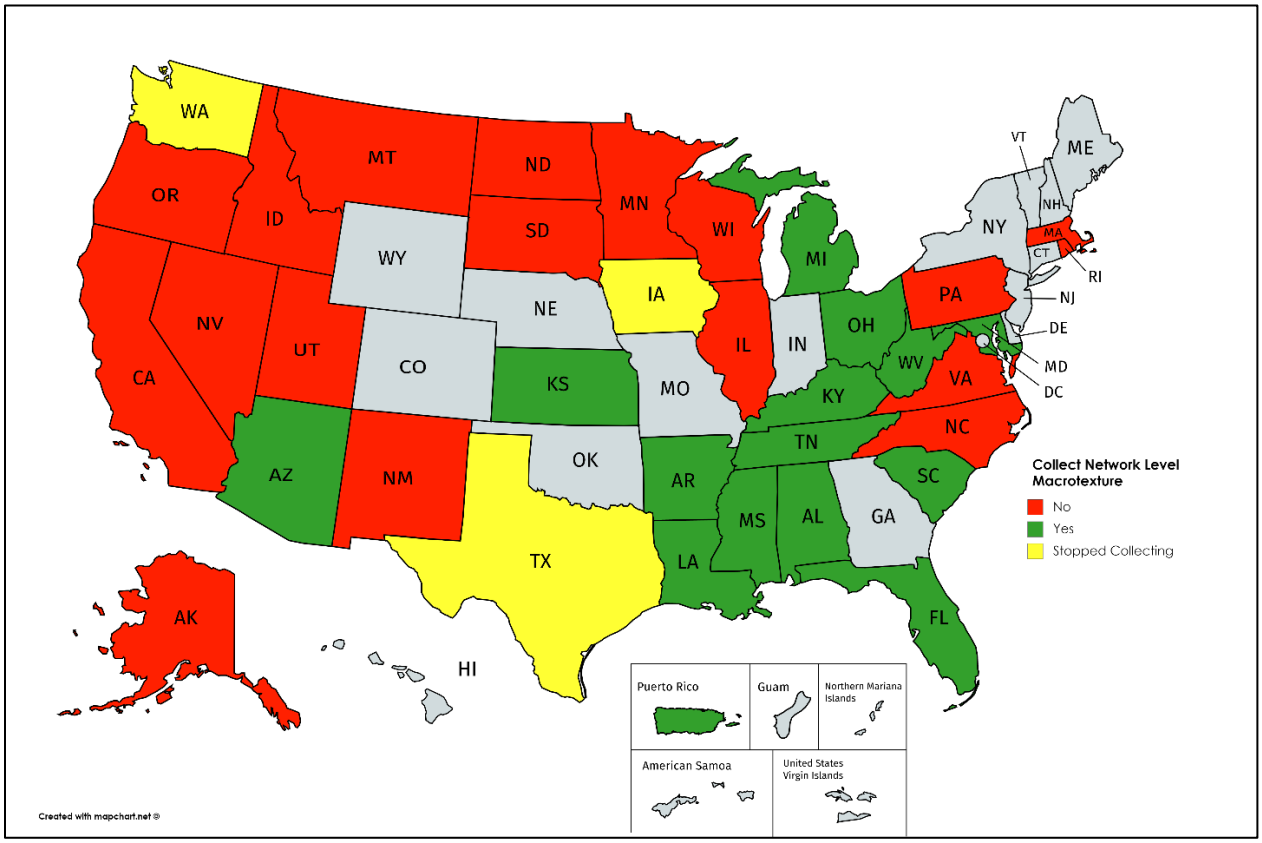


Figure 4 - States Collecting Network Level Macrotexture (Adapted from NCHRP (*4*))

For implementing macrotexture into PMS, macrotexture can be added to new or existing data collection contracts. The United Kingdom has implemented these type of contracts to gather macrotexture for both highways and local networks (*5*). As an example of project level application of macrotexture by agencies, the New Zealand Transport Agency uses macrotexture as a performance specification for chip seals and seal coats, requiring an average texture depth of greater than 0.9 mm (*6*).

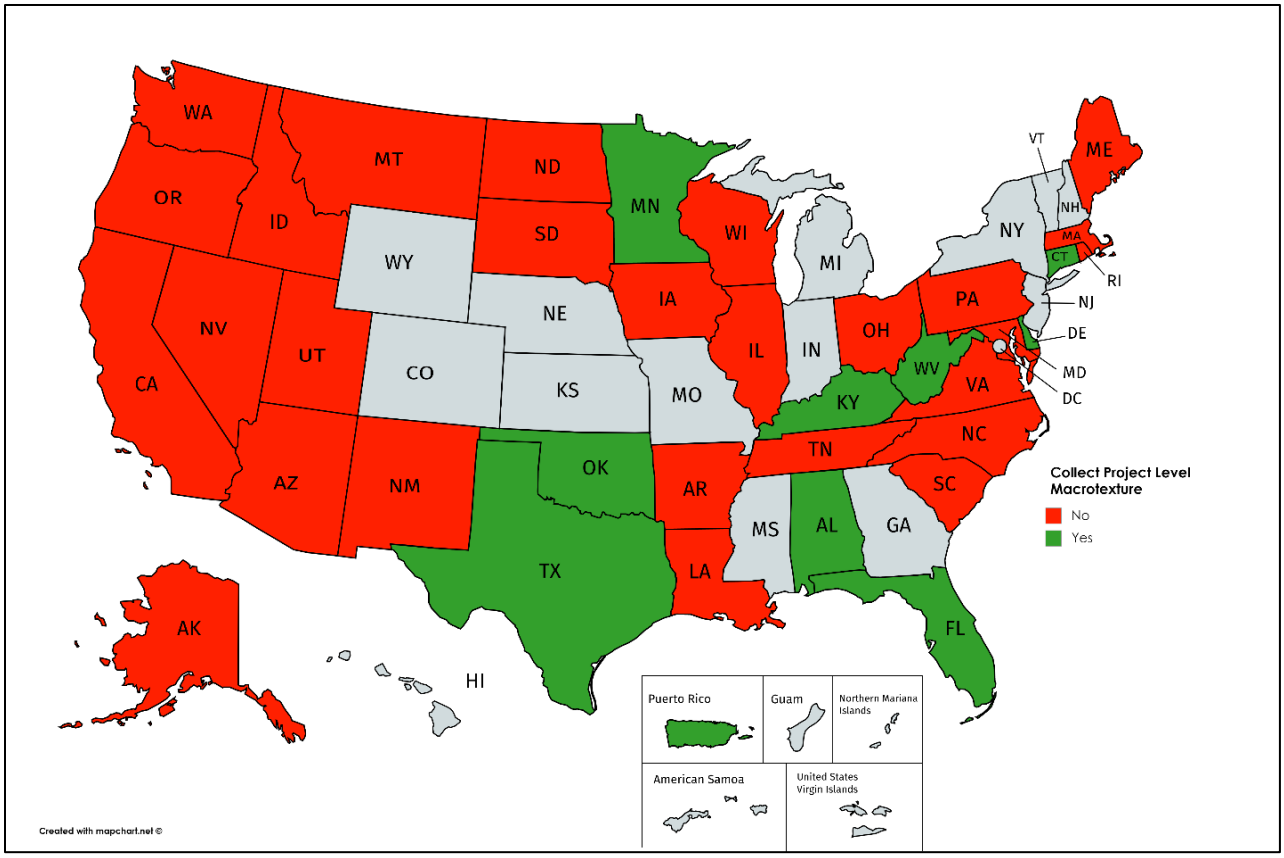


Figure 5 - States Collecting Project Level Macrotexture (Adapted from NCHRP (*4*))

Texture Wavelengths

In non-technical terms, pavement macrotexture is the texture created between aggregates or grooves that can be felt on the surface. The World Road Association defines macrotexture as "surface irregularities of a road pavement with horizontal dimensions ranging between 0.5 mm and 50 mm and vertical dimensions between 0.2 and 10 mm" (*7*). Alternatively, the International Organization for Standardization defines macrotexture as "deviation of a pavement surface from a true planar surface with the characteristic dimensions along the surface of 0.5 mm to 50 mm, corresponding to texture wavelengths with one-third-octave bands including the range 0.63 mm to 50mm of center wavelengths" (*8*).

Texture is also measured at other wavelengths, all of which are important to pavement tire interactions. Wavelengths smaller than macrotexture are called microtexture (less than 0.5mm), and wavelengths larger than macrotexture are called megatexure (50mm to .5 m). The largest wavelength category is referred to as roughness or unevenness (.5m to 50m) and is

commonly related to ride quality of a roadway. Hall et al. (*9*) describe microtexture as texture provided by the surface of the aggregates, and macrotexture as provided by the shape, size and gradation of aggregates in asphalt pavements or manufactured grooving in concrete pavements. They further state that microtexture provides friction through adhesion and macrotexture through hysteresis, with the added effect that macrotexture provides paths for water to drain from the surface reducing water film in wet conditions allowing the adhesion from microtexture. This leads to the observation that at high vehicle speeds macrotexture is the primary contributing factor to friction, and at low vehicle speeds the adhesion from microtexture controls friction (*9*). Figure 6 illustrates the effect of texture wavelengths of various vehicle-road interactions.

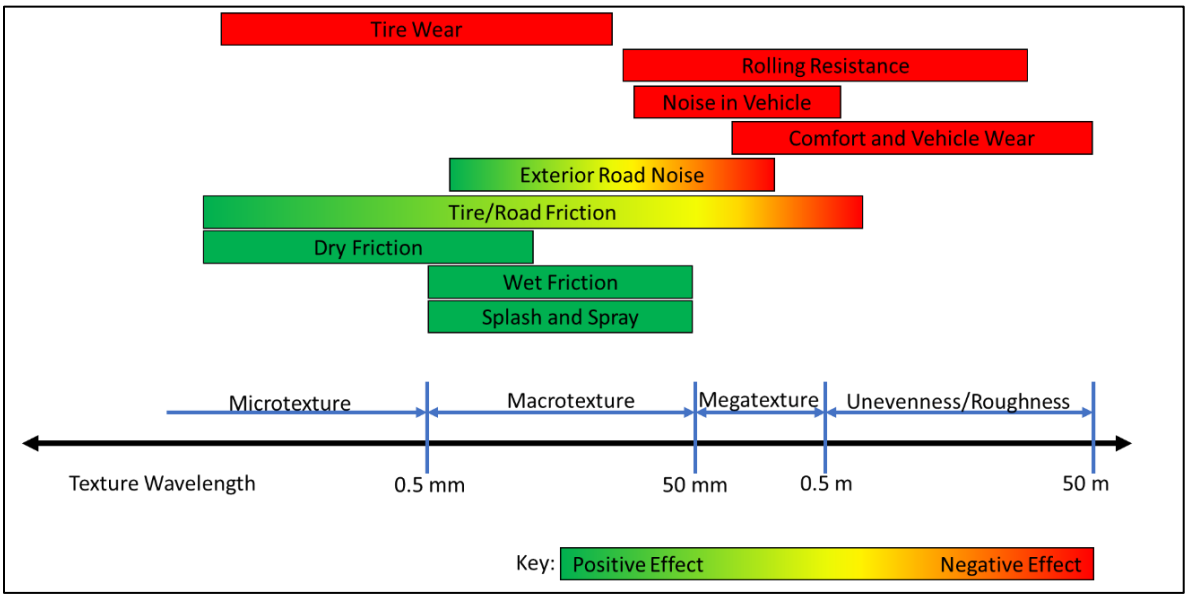


Figure 6 - Texture Effects (Adapted from PIARC and ISO 13473-2 (*8*))

Measurement Methods

The International Organization for Standardization (ISO) outlines the basic methods of macrotexture measurement, these methods are summarized following. Macrotexture can be measured in two ways, through volumetric methods or through laser-based systems. The volumetric methods are conducted by using a known volume of material and spreading it evenly over an area that can be measured. The volume of material is divided by the area yielding the average depth of the texture, called the Mean Texture Depth (MTD). Various materials have

been used for this method including grease, putty, sand, and small glass beads. The use of sand is commonly referred to as the sand patch test. Laser based systems can generate profile data in 2-D or 3-D, which is then processed mathematically to produce macrotexture parameters (*8*).

Macrotexture Parameters

Two of the most common macrotexture parameters that are derived from profiles are Mean Profile Depth (MPD) and Root Mean Square (RMS). These are the primary measures used and evaluated in this thesis.

MPD is the average of mean segment depths for all segments within a profile, where mean segment depth is defined as the distance from the average level to the average level of two highest peaks, one found in each half of a 100 mm segment as illustrated in Figure 7. Because MPD is based on peaks, it is sensitive to outliers and spikes within collected data. This is addressed through various filtering and outlier removal techniques. Some macrotexture measurement devices or accompanying software have automated data processing routines that perform the data filtering and provide MPD as an output (*10-14*).

ASTM E1845, the standard for calculation of MPD covers how data should be processed to account for outliers and spikes in three main steps:

1) The processing includes removal of outliers, mostly caused by dropouts from devices. These outliers (defined as values higher or lower than the range of the surrounding profile) are removed, and the value is replaced with an interpolation of the values preceding and following the outlier;
2) the profiles are filtered to reduce noise;
3) the slope is suppressed by subtracting out a linear regression of the profile for the segments (*15*).

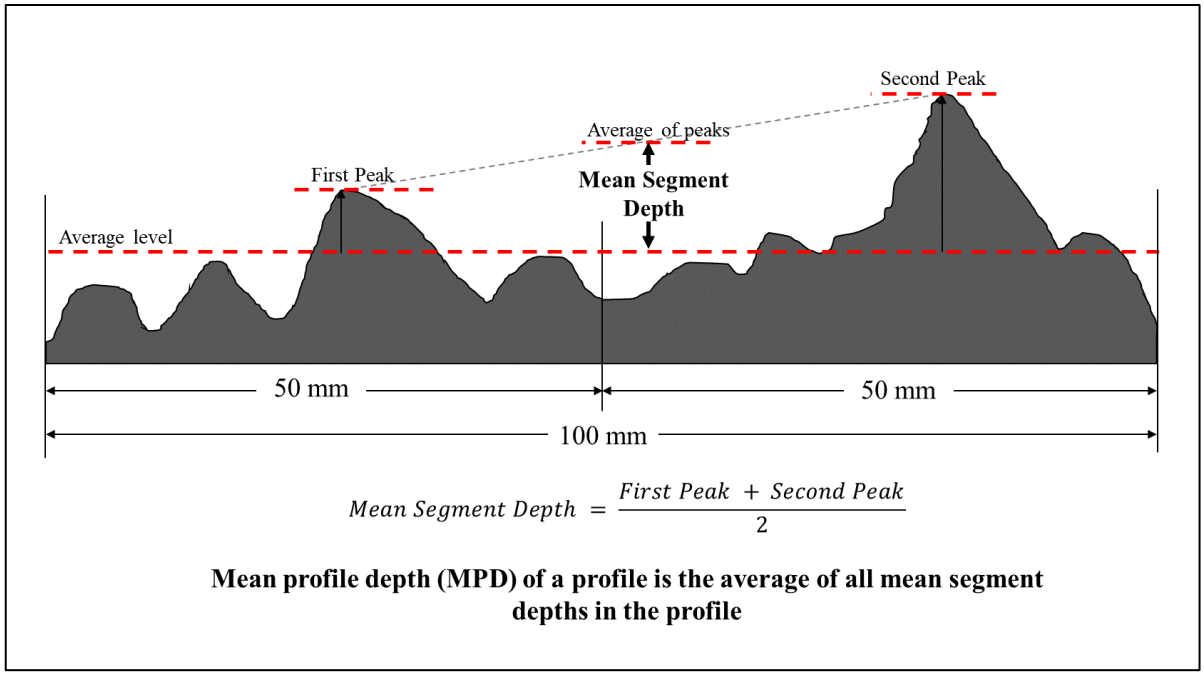


Figure 7 - Mean Profile Depth Calculation (Adapted from ASTM E1845 2015 (*15*))

Root Mean Square in simple terms is the standard deviation of the profile data for a segment and is in theory more robust against outliers and spikes than MPD.

$$RMS = \sqrt{\frac{1}{l}\int_0^l Z^2(x)dx} \quad (1)$$

where,

RMS = root mean square of ordinate values *Z(x)* with evaluation length of *l*

Outlier removal methods

As mentioned in the ASTM guidance for MPD calculation, outlier removal is an important part of macrotexture calculation. Goubert and Bergiers developed a method for outlier removal which has since been adopted by ISO (*4*). Their method looks at two consecutive profile amplitudes defined as Z_i and Z_{i-1}. A spike is defined as any value of the absolute difference of these amplitudes that is greater than a constant factor α, multiplied by the step size (space between device readings) Δx. The value of constant α is prescribed to be between 2 and 10, with careful consideration that if α is too low it could overly smooth the profile and if it is too high it may fail to appropriately identify spikes (*16*).

$$|z_i - z_{i-1}| \geq \alpha \cdot \Delta x \quad (2)$$

Katicha et al. (*17*) proposed an alternative method that identifies values that are outliers of the distribution of the measurements. This is done in two steps; the first of which is to fit a distribution to the data set using Generalized Gaussian Distributions (GGD). The probability density function of distribution is:

$$p(x) = \frac{\beta}{2\alpha\Gamma\left(1/\beta\right)} \exp\left(-\left(\frac{|x-\mu|}{\alpha}\right)^{\beta}\right) \quad (3)$$

where,

β = shape parameter,

α = a scale parameter and

μ is the location parameter

β is determined by fitting the distribution to 90th to 97th percentiles of the data. The second step of this procedure uses the False Discovery Rate (FDR) method developed by Benjamini and Hochberg (*18*) to "control the proportion of false discoveries." This is done by calculating the p-vale of all measurements, *n*, ordering them from least to greatest, selecting the value that controls the FDR (such as 0.01, 0.05, or 0.1), and setting *k* as the maximum value of *i* in the following, so that any measurement with p-value less than or equal to p_k is identified as a spike (*17*).

$$p_i \leq \frac{\mathrm{i}}{n} q \quad (4)$$

Change in Texture with Traffic

Macrotexture of a pavement changes over time from its initial constructed condition. Various modes of deterioration from the environment and traffic loading will affect the macrotexture. However, little research has been done to isolate the effect of traffic loading on macrotexture particularly for binder rich overlays.

Previous research has shown that the change in macrotexture of a pavement with traffic loading can vary depending on the surface type, as well as the distresses observed. For example, macrotexture has been used to measure chip seal performance, and models developed have shown that there is a decrease in chip seal macrotexture with traffic and it is best fit with

logarithmic functions (*19*). It has also been demonstrated that the macrotexture of seal coats as measured in Mean Texture Depth (MTD) using the sandpatch method decreased with time and traffic, with this decrease attributed to wear of aggregates and embedment (*20*). Riemer and Pittenger conducted an experiment to test the deterioration of macrotexture for various pavement preservation treatments. Their study also showed a logarithmic relationship between the macrotexture and the independent variable of time to model deterioration for use in pavement management systems (*21*). Using a Model Mobile Load Simulator (MMLS33), Lin and Tongjing showed that macrotexture decreased with simulated traffic load initially but reached a stable level at around 30,000 loading cycles (*22*). Alternatively, Powell and Buchanan observed the change in macrotexture for thin asphalt overlays and found that traffic loading can lead to an increase in macrotexture due to raveling of aggregate (*23*). For stone mastic asphalt, a study showed a logarithmic relationship between macrotexture decrease and the number of gyrations in a gyratory compactor used to compact test samples, however when simulated traffic was applied using a road test machine (RTM), the samples with lower initial macrotexture showed little change as a result of traffic while samples with "greatest values of initial macrotexture showed an increase in texture in the later stages of simulated trafficking," which was attributed to raveling (*24*). The Highway Development and Management software system (HDM-4) uses a complex logarithmic equation to model decrease in macrotexture with traffic. It enables calibration of the model by observing the incremental macrotexture change during an analysis year (*25*).

The research indicates that the effect of traffic on macrotexture depends on the surface type as well as the distresses induced. Furthermore, the effect of traffic loading can either decrease or increase macrotexture over time, and that the behavior may be best modelled linearly or logarithmically.

Accelerated Pavement Testing

This thesis relies on accelerated pavement testing (APT) to simulate traffic loading onto the pavement using a Dynatest Mk VI Heavy Vehicle Simulator (HVS) (Figure 8). The HVS environmental chamber was also used to control temperature and protect the pavement from weather effects. The HVS is capable of testing sections 40 ft long and up to 6 ft wide and can

produce a pavement loading between 7000 and 22,500 pounds at 10,000 passes unidirectionally or 20,000 passes bidirectionally over 24 hours (*26*). Figure 9 shows the inside of the chamber.

Figure 8 - Heavy Vehicle Simulator at Virginia Tech Transportation Institute

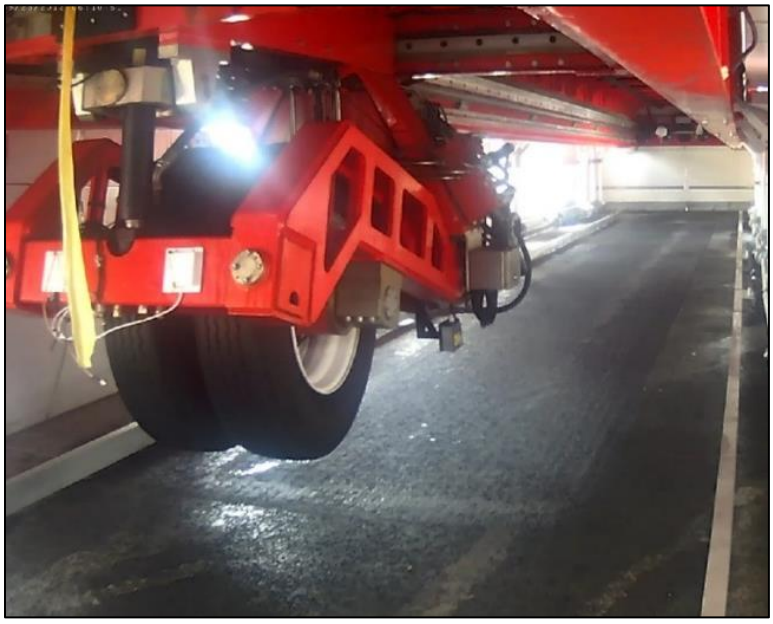


Figure 9 - HVS (Inside)

The traffic load applied to the pavement can be analyzed in a variety of ways. This research analyzed the effect of cumulative Equivalent Single Axel Loads (ESALs). ESALs is a preferred parameter to measure traffic as it is the primary measure of traffic loading used in

mechanistic-empirical pavement design methods. Theoretically, the cumulative ESALs is the summation of the Equivalent Axel Load Factor (EALF) for axle load groups multiplied by the number of passes of that axle load group for all axle load groups applied to pavement. The EALF is defined as the ratio of pavement damage induced from a pass of the axle load group to the damage from a standard 18-kip single axle.

$$EALF = \frac{Damage\ from\ pass\ of\ axle\ in\ load\ group}{damage\ from\ a\ 18{-}kip\ single\ axle} \quad (5)$$

$$ESAL = \sum F_i N_i \quad (6)$$

Where F_i *is* the EALF and N_i is the number of passes for the ith axle load group (*27*)

The ESALs applied by each pass of the HVS dual wheel are estimated by (*28*):

$$ESAL = \left(\frac{HVS\ Load\ Applied}{9\ Kips}\right)^{4.2} \quad (7)$$

Where a pass is simply a single run of the loaded HVS dual wheel from one end of the test strip to the other.

Equipment to Measure Macrotexture

Macrotexture measurement equipment is available in a variety of types with different advantages and operation. In general, there are three main categories: static, walking-speed and high-speed, with certain models of each type offering opportunity for 2-D or 3-D profile data collection. Following is a sample selection of currently available devices presenting a 2-D and 3-D device of each type.

Static Devices

Static devices require manual placement and operation while stationary over the surface being measured. This can lead to a more labor-intensive effort with a smaller sample of data for network level collection.

One of the most common static devices and the one utilized in this research for measurements within the HVS is the Nippo Sangyo Co. Circular Track Meter (CT Meter). The CT Meter is a static device that uses a single laser on a circular track of radius of 142mm to

collect a 2-dimensional profile. It collects 1,024 data points in 8 arc segments (111.5 mm) with 128 points each (*11*). The macrotexture is reported in MPD and RMS.

Figure 10 shows a photograph of the CT Meter, Figure 11 a schematic of the measurement footprint of the device, and Figure 12 sample profile output from the device.

Figure 10 - CT Meter

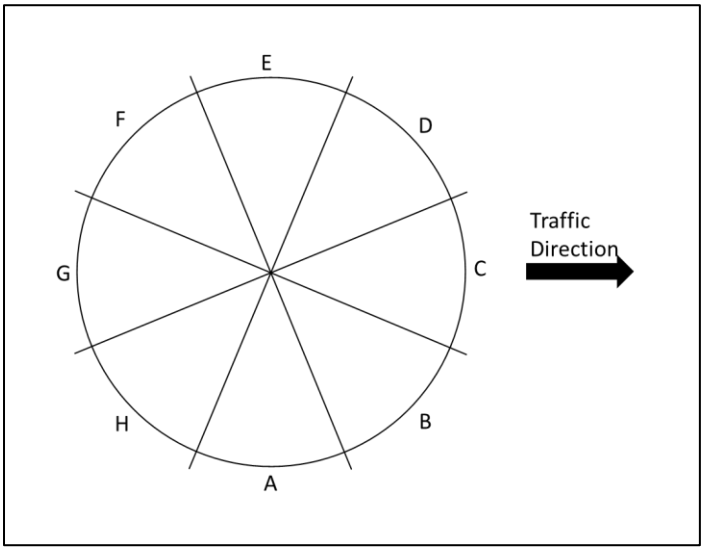


Figure 11 - CT Meter Measurement Footprint (Adapted from Nippo Sangyo Co. 2018 (*11*))

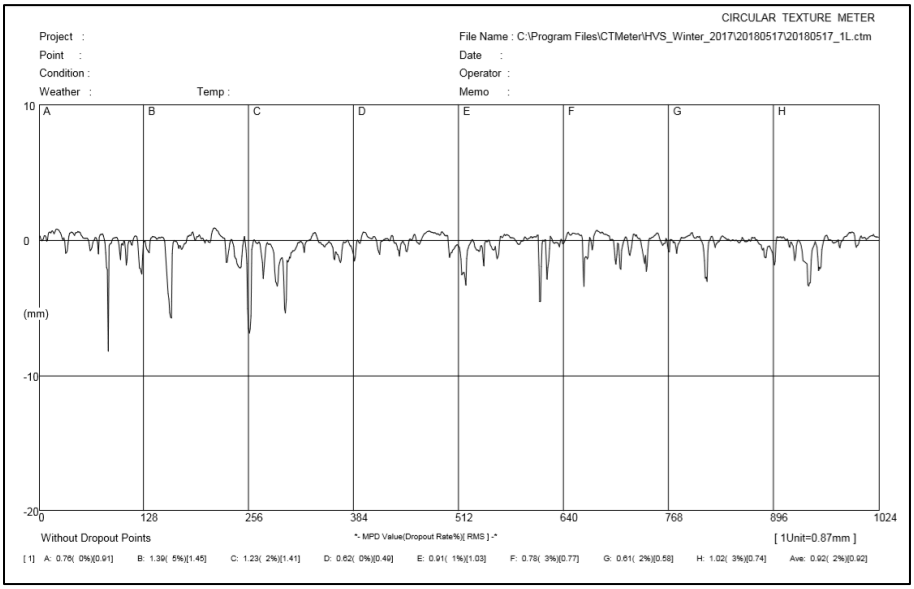


Figure 12 - Sample CT Meter Output

Another example of a static device is the AMES Laser Texture Scanner. Two versions of this device are shown measuring a manufactured surface in Figure 13. While the CT meter collects a 2 dimensional profile, this device collects data in 3 dimensions with a length resolution of .0496 mm, width resolution of .0415 mm and vertical resolution of .01 mm (*13*).

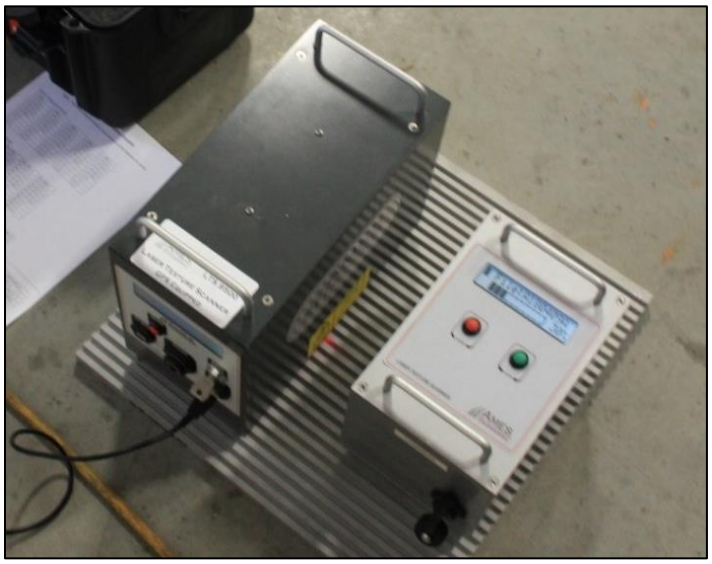


Figure 13 - AMES Laser Texture Scanner

Walking Speed Devices

Similar to static devices, walking speed devices are manually operated, however, they collect profile data as they are pushed over the surface being measured allowing for a larger area of coverage to collect data from. A commonly used walking device is the TM2 from WDM shown in Figure 14. This device uses a line laser projected transverse to the direction of travel and measures the profile at a spatial interval of 2 mm when operated at an average walking speed of 3.5 kilometers per hour (*14*).

Figure 14 - WDM TM2

Another option for walking speed devices is the ARRB G3 Walking Profiler shown in Figure 15. This device is designed to produce profile outputs of roughness (given in International Roughness Index) as well as macrotexture reported in MPD (*29*).

Figure 15 - ARRB G3 Walking Profiler

High Speed Devices

High speed devices are designed to operate mounted on a motor vehicle and moving at traffic speed. This is advantageous for network level collection because data can be collected over a much larger area than with static and walking speed devices.

The AMES AccuTexture 100 is an example of a high speed device that collects 2-D profiles using a single spot laser with a sample rate of 100 KHz shown mounted on the front of a vehicle in Figure 16 (*30*).

Figure 16 - AMES AccuTexture Device

An example of a high speed device that collects 3-D data is the Surface Systems & Instruments, Inc. (SSI) CS9350 Partial Lane Survey-Profiler (pictured mounted in Figure 17) which is capable of collecting at 25 longitudinal spacing with a line laser that collects at 0.3 mm transversely (*31*).

Figure 17 - SSI CS9350 Partial Lane Survey-Profiler

Equipment Comparisons

Several macrotexture equipment comparisons along with attempts to harmonize with friction measurements have been previously performed (*4*).

PIARC

The World Road Association conducted the International Experiment to Compare and Harmonize Texture and Skid Resistance Measurement in the fall of 1992. The goal of this study was to compare various macrotexture and friction measurement devices for producing an international standard to harmonize them, presented as the International Friction Index (*32*).

NASA

Over the course of several years beginning in 1994 NASA held the Annual NASA Tire/Runway Friction Workshop at the Wallops Flight Facility in Virginia. Over the years various friction, texture and profiler devices have been tested under various conditions (*33*).

United Kingdom (UK)

The UK has been measuring friction since the early 1930's and introduced texture into their data collection in the 1970's using the sand patch test. Two annual contracts, TRACS (Traffic Speed Condition Surveys) and SCANNER (Surface Condition Assessment of the National Network of Roads) are implemented to collect surface texture among other measurements for the highways and local roads respectively in the UK. They have also studied the relationships among speed, texture and friction (*5*).

ROSANNE

ROSANNE (Rolling Resistance, Skid Resistance and Noise Emission Measurement Standards for Road Surfaces) is a project funded by the European Union, and as the name implies is focused on harmonizing skid resistance, noise emission, rolling resistance as well as studying texture and its influence on them (*34; 35*).

TYROSAFE

TYROSAFE (Tyre and Road Surface Optimisation for Skid Resistance And Further Effects) was another European Union project to harmonize and "optimize the assessment and management of essential tire/road interaction parameters" (*36*). This project was focused on the interaction between tires and the road surface (*37*).

Pavement Surface Properties Consortium

The Pavement Surface Properties Consortium has conducted annual equipment comparisons since 2007 in order to evaluate methods and technologies (*38*). In 2008 a project was conducted on the Virginia smart road to compare various friction devices including a locked-wheel skid trailer, a grip tester and a dynamic friction tester over a variety of pavement surface types. The project evaluated the application of IFI from the PIARC project to the devices tested (*39*).

Limits of Agreement

One of the challenges of conducting comparisons of device measurements is the lack of a known true value being measured, often referred to as the "reference" measurement. To account for this, a methodology that was pioneered by Bland and Altman for use in the biopharmaceutical fields called Limits of Agreement (LOA) can be employed(*40-44*). This

method can assess the agreement or closeness between readings of the devices evaluated as well as the repeatability of each individual device (*43*). In the field of pavements properties Katicha et al. (*45)* used this method to compare Falling Weight Deflectometer (FWD) and Traffic Speed Deflectometer (TSD) measurements. It was also used by de Leon, Flintsch and McGhee (*46*) to compare friction measuring devices.

This type of analysis enables the determination of both the repeatability and reproducibility of devices. Repeatability is defined as "the ability of the equipment/technology/raters to produce the same values on repeated measurements" and reproducibility as "a measure of how well two different devices/methods/raters are able to measure the same pavement condition value on the same road segment" (*3*).
The procedures for LOA are as follows:

The LOA method used assumes that the true value measured does not change and that measurements are made in quick succession to minimize effects caused by changes in experimental conditions. To ensure that the true value of the pavement macrotexture does not change, the standard deviation of data should be constant and not related to the magnitude of the measurements taken. This is checked by plotting the standard deviation against the mean for each device (*42*). Then the 95% confidence LOA are calculated as prescribed by Bland and Altman (*42*):

$$LOA = 1.96 * S_c \quad (8)$$

where:

S_c = corrected standard deviation of differences = $\sqrt{S_D{}^2 + f_1 \cdot S_1{}^2 + f_2 \cdot S_2{}^2}$ (9)

S_D = standard deviation of the difference between the mean of the repeated observations for each subject by two devices compared;

Multiplication factors f_1 and f_2 depend on the number of observations (or repeat runs) m_i on a particular subject *i* and the number of subjects n, such that:

$$f = \left(1 - \frac{1}{n}\Sigma\frac{1}{m_i}\right) \quad (10)$$

if the number if observations is the same for all subjects the equations is simplified to

$$f = f_1 = f_2 = \left(1 - \frac{1}{m}\right) \quad (11)$$

and S_1 and S_2 are the variances of the devices which is the Mean Square Error (MSE) determined from an ANOVA of the device with the subjects as the model input (*46*). These variances are also used to determine the repeatability coefficient, where the square root of the MSE is used as the within-device standard deviation and as proposed by Bartlett and Frost can be interpreted to mean that two measurements made on a subject by a device should differ by no more than the repeatability coefficient, c_r, 95% of the time assuming a normal distribution of differences between measurements (*47*):

$$c_r = 1.96 * \sqrt{2} * SD \quad (12)$$

The boundaries for LOA are the mean of the differences for the devices for all the subjects ± the LOA. The LOA are plotted with the mean of the two devices' measurements on the *x*-axis and the difference between the devices on the *y*-axis for each subject.

Summary of Background Research

Pavement management requires inputs of both current condition and prediction of future condition. The various wavelengths of pavement texture including microtexture, macrotexture, megatexture and roughness or unevenness play a key role in the interactions of tires on the roadway. Macrotexture is an important parameter for safety issues related to friction and splash and spray. While previous research has been conducted to model macrotexture changes with traffic, there is room for further research to focus on different surface types and distresses. Furthermore, while comparisons of macrotexture equipment have been previously conducted, no standard for network level macrotexture collection has been implemented in the United States.

References:

[1] Peshkin, D. G. H., T.E.; Zimmerman, K.A. NCHRP Report 523, Optimal Timing of Pavement Preventive Maintenance Treatment Applications. *Transportation Research Board*, 2004.

[2] Shahin, M. Y. *Pavement Management for Airports, Roads, and Parking Lots*. Kluwer Academic Publishers, 1994.

[3] Flintsch, G., and K. McGhee. NCHRP Synthesis 401 Quality Management of Pavement Condition Data Collection.In, 2009. p. 143.

[4] NCHRP. NCHRP Project 10-98 Protocols for Network-Level Macrotexture Measurement Interim Report. 2017, p. 108.

[5] Ferne, B. UK Experiences on Pavement Texture Measruement and Interpretation. *Transportation Research Circular,* Vol. E-C216: International Experience and Perspective of Pavement Texture Measurements and Evaluation, 2015, pp. 25-41.

[6] NZTA. Performance Based Specifications for Reseals.In, New Zealand Transport Agency P17, 2012.

[7] PIARC. *Road Dictionary*. http://www.piarc.org/en/Terminology-Dictionaries-Road-Transport-Roads/. Accessed 16 December 2016.

[8] ISO. ISO 13473-2 Characterization of pavement texture by use of surface profiles. Vol. Part 2 Terminology and basic requirements related to pavement texture profile analysis.

[9] Hall, J. W., K. L. Smith, L. Titus-Glover, J. C. Wambold, T. J. Yager, and Z. Rado. Guide for Pavement Friction. 2009, p. 257p.

[10] ISO 13473-1. Characterization of pavement texture by use of surface profiles - Part 1: Determination of Mean Profile Depth.In *Part 1: Determination of Mean Profile Depth*, International Organization for Standardization, 1997.

[11] Nippo Sangyo Co. *Circular Track Meter Product Guide*. http://www.nippou.com/en/products/ct.html. Accessed 1 June 2018, 2018.

[12] AMES. *Laser Texture Scanner Model 9400/9400HD.* AMES Engineering. https://amesengineering.com/products/laser-texture-scanner-model-9400/. Accessed 4 Oct 2018.

[13] AMES. *Laser Texture Scanner Model 9500.* AMES Engineering. https://amesengineering.com/products/laser-texture-scanner-model-9500/. Accessed 4 October 2018.

[14] WDM. *TM2*. WDM. https://www.wdm.co.uk/equipment/equipment-tm2/. Accessed 4 October 2018.

[15] ASTM. E1845 Standard Practie for Calculating Pavement Macrotexture Mean Profile Depth. *ASTM International*, 2015.

[16] Bergiers, L. G. a. A. About the Reproducibility of Texture and the Problem of Spikes. *7th Symposium on Pavement Surface Characteristics: SURF 2012*, 2012.

[17] Katicha, S. W., D. E. Mogrovejo, G. W. Flintsch, and E. D. d. L. Izeppi. Adaptive Spike Removal Method for High-Speed Pavement Macrotexture Measurements by Controlling the False Discovery Rate. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2525, 2015, pp. pp 100–110.

[18] Benjamini, Y., and Y. Hochberg. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society,* Vol. Series B (Methodological), Vol. 57, No. 1, pp. 289-300.

[19] Aktaş, B., M. Karaşahin, and M. Tiğdemir. Developing a macrotexture prediction model for chip seals. *Construction and Building Materials,* Vol. 41, 2013, pp. 784-789.

[20] Roque, R., D. Anderson, and M. Thompson. Effect of Material, Design, and Construction Variables on Seal-Coat Performance. *Transportation Research Record No. 1300, Asphalt Pavement and Surface Treatments: Construction and Performance 1991*, 1991.

[21] Riemer, C., and D. Pittenger. Modeling Pavemnt Texture Deterioration as a Pavement Preservation Management System Tool. *Transportation Research Record: Journal of the Transportation Research Board,* Vol. 2306, 2012.

[22] Lin, C., and W. Tongjing. Effect of fine aggregate angularity on skid-resistance of asphalt pavement using accelerated pavement testing. *Construction and Building Materials,* Vol. 168, 2018, pp. 41-46.

[23] Powell, R. B., and S. Buchanan. Long Term Performance of a Thin Asphalt Overlay on the NCAT Pavement Test Track. *Transportation Research Record: Journal of the Transportation Research Board,* Vol. 2306, 2012.

[24] Woodward, D., P. Millar, C. Lantieri, C. Sangiorgi, and V. Vignali. The wear of Stone Mastic Asphalt due to slow speed high stress simulated laboratory trafficking. *Construction and Building Materials,* Vol. 110, 2016, pp. 270-277.

[25] Rainsford, S., and C. Parkman. Predicting Texture Deficiency in Pavement Management. *MWH NZ Ltd, Transit New Zealand*.

[26] Dynatest. *Heavy Vehicle Simulator (HVS) Brochure*. https://dynatest.egnyte.com/dl/Qdm823XPyJ. Accessed 1 June 2018.

[27] Huang, Y. H. *Pavement Analysis and Design, 2nd Ed.* Pearson Prentice Hall, 2004.

[28] D. Jones, R. J. H. Reflective Cracking Study: First-LEvel Report on HVS Testing on Section 586RF -- 45 mm MB15-G Overlay.In, University of California Pavement Research Center UC Davis, UC Berkeley, 2006.

[29] AARB. *Walking Profiler G3.* ARRB Sytems. https://arrbsystems.com/walking-profiler-g3/. Accessed 4 October 2018.

[30] AMES. *Ames AccuTexture 100.* AMES Engineering. https://amesengineering.com/products/ames-accutexture-100/. Accessed 4 October 2018.

[31] SSI. *CS9350 Partial Lane Survey-Profiler.* Surface Systems & Instruments, Inc. https://www.smoothroad.com/cs9350-high-speed-portable-geoprofiler. Accessed 4 October 2018.

[32] Henry, J. J. Overview of the Internatinoal PIARC Experiment to Compare and Harmonize Texture and Skid Resistnace Measurements: The International Friction Index.In, The Pennsylvania State University, University Park, PA.

[33] Yager, T. J. An Overview of the Annual NASA Tire/Runway Friction Workshop and Lessons Learned.In, NASA Langley Research Center, Hampton, Virginia, 2005.

[34] *ROSANNE*. http://rosanne-project.eu/. Accessed 8 October 2018.

[35] Haider, M., M. Conter, R. Wehr, U. Sandberg, and F. Anfosso. Project ROSANNE: Rolling resistance, Skid resistance, and Noise Emission measurement standards for road surfaces. *Inter-noise*, 2014.

[36] Haider, M., and M. Conter. TYROSAFE Final Summary Report.In, Brussels, Belgium, 2010.

[37] Scharnigg, K., G. Schwalbe, and M. Haider. TYROSAFE (Tyre and Road Surface Optimisation for Skid Resistance and Further Effects. *3rd International Surface Friction Conference, Safer Road Surfaces*, 2011.

[38] NCHRP. *Transportation Pooled Fund Program, Study Detail View, Pavement Surface Properties Consortium: A Research Program*. https://www.pooledfund.org/details/study/371. Accessed 19 Nov 2018.

[39] Flintsch, G. W., E. D. L. Izeppi, K. K. McGhee, and J. A. Roa. Evaluation of International Friction Index Coefficients for Various Devices. *Transportation Research Record: Jornal of the Transportation Research Board (2094)*, 2009, pp. 136-143.

[40] Martin Bland, J., and D. Altman. Statistical Methods for Assessing Agreement Between Two Methods of Clinical Measurement. *The Lancet*, Vol. 327, No. 8476, 1986, pp. 307-310.

[41] Bland, J. M., and D. G. Altman. Measuring agreement in method comparison studies. *Statistical methods in medical research*, Vol. 8, No. 2, 1999, pp. 135-160.

[42] Bland, J. M. A., Douglas G. Agreement between methods of measurement with multiple observations per individual. *Journal of biopharmaceutical statistics,* Vol. 17, No. 4, 2007, pp. 571-582.

[43] Barnhart, H. X., M. J. Haber, and L. I. Lin. An overview on assessing agreement with continuous measurements. *Journal of biopharmaceutical statistics*, Vol. 17, No. 4, 2007, pp. 529-569.

[44] Carstensen, B., J. Simpson, and L. C. Gurrin. Statistical models for assessing agreement in method comparison studies with replicate measurements. *The international journal of biostatistics,* Vol. 4, No. 1, 2008.

[45] Katicha, S. W., G. W. Flintsch, B. Ferne, and J. Bryce. Limits of agreement (LOA) method for comparing TSD and FWD measurements. *Journal of Pavement Engineering*, Vol. 15, 2014, p. 6.

[46] de Leon, E. D., G. Flintsch, and K. McGhee. Limits of Agreement Method for Comparison of Pavement Friction Measurement. *Transportation Research Record: Journal of the Transportation Research Board,* Vol. 2306, 2012, pp. 188-195.

[47] Bartlett, J., and C. Frost. Reliability, repeatability and reproducibility: analysis of measurement errors in continuous variables. *Ultrasound in Obstetrics and Gynecology: The Official Journal of the International Society of Ultrasound in Obstetrics and Gynecology,* Vol. 31, No. 4, 2008, pp. 466-475.

CHAPTER 3: CHANGE IN MACROTEXTURE DUE TO TRAFFIC AND BLEEDING IN AN ACCELERATED PAVEMENT TESTING MACHINE[1]

Abstract

This study sought to model the behavior of a binder-rich stone matrix asphalt overlay of jointed plain concrete pavement when subjected to traffic loading using a heavy vehicle simulator to report the effect on pavement macrotexture. The change in macrotexture due to traffic loading was measured periodically with a circular track meter. Overall, as the cumulative load increased, the macrotexture decreased due to bleeding on the pavement's surface. A regression model was produced for the center of the wheel path which determined that, on average, the macrotexture's root mean square (RMS) decreased 0.14 mm per million equivalent single axle load applied, with a coefficient of determination of $R^2 = 0.507$. The relatively low R^2 was a result of high variability within the collected data. A comparison of RMS and mean profile depth (MPD) outputs indicated that RMS was sensitive to macrotexture changes due to bleeding. No clear model was determined for measurements outside the center of the wheel path. In one case, the buildup of material caused measures of MPD to increase while RMS decreased when calculated from the same profiles. In another case, both RMS and MPD measurements showed macrotexture increasing due to buildup of material on the surface.

Keywords: Accelerated Pavement Testing, Macrotexture, Surface Properties

[1] This manuscript was coauthored with Vincent Bongioanni and Dr. Gerardo Flintsch and will be submitted to ASTM Journal of Testing and Evaluation.

Background

A roadway's texture is the primary influencing factor over several critical tire-pavement interactions. Chief among these are the important safety concerns of friction (*1*) and splash and spray (*2*). Several studies (*3; 4*) have shown a correlation between texture and vehicle crashes in wet weather. Secondary to safety concerns, in terms of pavement macrotexture, are user and environmental impacts such as rolling resistance (*5*) and road noise (*6*).

The World Road Association defines macrotexture as “surface irregularities of a road pavement with horizontal dimensions ranging between 0.5 mm and 50 mm and vertical dimensions between 0.2 and 10 mm” (*7*). Surface texture smaller than this is typically referred to as "microtexture," and larger texture as "megatexture" or, if larger still, as the "roughness" of a road. As water film thickness increases, the pavement’s macrotexture provides drainage paths for water beneath the tire to escape, reducing hydroplaning potential and allowing for greater tire-pavement adhesion (a function of the pavement’s microtexture). Macrotexture also provides friction through hysteresis (energy loss due to asymmetrical deformation of the tire). The hysteresis effect increases exponentially with increasing vehicle speed, accounting for 95% of available friction at speeds above 65 mph (*8*).

There are a variety of devices, both dynamic and static, that are used to measure macrotexture. Dynamic devices collect data while moving and vary in operation from manually pushed walking speed devices to vehicle-mounted highway speed devices. Static measurement methods require the data collection team to hand-place a device and manually perform the test on the surface of interest. Several data points are required on homogeneous pavement surfaces to find the surface’s representative characteristics. Operator judgment is needed to determine how many samples to collect, where to collect them, and what constitutes a “homogeneous pavement section.” These methods are generally regarded as “ground truth,” and can provide a high degree of measurement accuracy because they can record several samples at different orientations in one placement.

Two common measures of macrotexture are mean profile depth (MPD) and root mean square (RMS). MPD is the average of mean segment depths for all segments within a profile, where mean segment depth is the distance from the average level to the average level of the two

highest peaks in each half of a 100 mm segment, as illustrated below (Figure 18). Because MPD is peak-based, it is sensitive to outliers and spikes within the collected data. Various filtering and outlier removal techniques address these concerns. Many macrotexture measurement devices and software have automated data processing routines that perform the required data filtering.

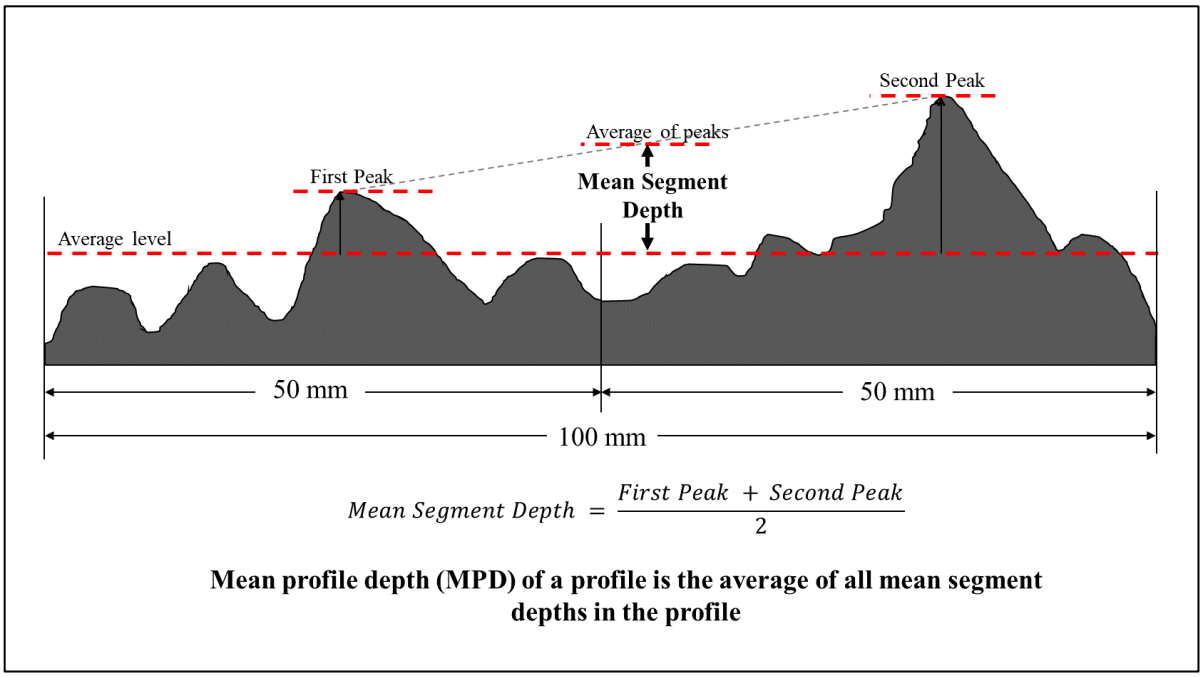


Figure 18 Mean profile depth – adapted from ASTM E1845 2015 (*9*).

RMS is akin to a standard deviation of the profile data for a segment and is in theory more robust against outliers than MPD because it is not based on single peaks within a given profile. It is calculated by the following equation (*10*)

$$RMS = \sqrt{\frac{1}{l}\int_0^l Z^2(x)dx} \quad (1)$$

Where:

Z(x) are ordinate values of texture height measurements

l is the evaluation length

Previous research on macrotexture has shown that the change in macrotexture of a pavement's surface can vary depending on the surface type, as well as the distresses observed. Macrotexture is a common measure of chip seal performance, and models for the decrease in chip seal macrotexture with traffic are best fit with logarithmic functions (*11*). It has been shown that the macrotexture of seal coats as measured in mean texture depth (MTD) using the sandpatch method decrease with time and traffic, with this decrease attributed to wear of aggregates and embedment (*12*). Another study also showed a logarithmic relationship between the macrotexture of various pavement preservation techniques and the independent variable of time to model deterioration for use in pavement management systems (*13*). Lin and Tongjing showed that macrotexture decreased with load initially but reached a stable level at around 30,000 loading cycles of the Model Mobile Load Simulator (*14*). Alternatively, it has been shown that traffic loading can lead to an increase in macrotexture due to raveling of aggregate (*15*). For stone mastic asphalt, a study showed a logarithmic relationship between macrotexture and the number of gyrations in a gyratory compactor used to compact test samples. However, when simulated traffic was applied using a road test machine, the samples with lower initial macrotexture showed little change as a result of traffic, while samples with "greatest values of initial macrotexture showed an increase in texture in the later stages of simulated trafficking," which was attributed to raveling (*16*).

Objective

This study sought modeled the behavior of binder-rich stone matrix asphalt (SMA) overlay of jointed plain concrete pavement (JPCP) when subjected to traffic loading using a heavy vehicle simulator to isolate the effect of traffic from other environmental effects and report the effect on pavement macrotexture. For this study, a decrease in pavement macrotexture was expected; however, as indicated, the distresses induced under loading can vary the effect on macrotexture. In binder-rich mixes, bleeding or flushing are commonly observed distresses. Therefore, the focus of this study was to model the effects of traffic-induced bleeding on macrotexture through analysis of variance (ANOVA) and linear regression.

Methodology

Pavement Structure

The structure of the pavement tested in this study is shown in Figure 19. The surface was comprised of two binder-rich (6.7%, by mass, binder content) lifts of SMA, each of identical mix design with a nominal maximum aggregate size (NMAS) of 12.5 mm and a total SMA thickness of 75mm. A JPCP underlaid the SMA founded upon an intermediate layer of dense-graded asphalt with a NMAS of 19 mm that ranged from 25 to 38 mm in thickness to create a level bedding course for the JPCP. The base course was a Virginia Department of Transportation aggregate blend "21-B," which was specified to be a non-plastic select material comprised of crushed materials with a minimum California Bearing Ratio (CBR) of 30 (*17*). The subbase is a select material with a minimum CBR of 7.5. The construction began with excavations to a depth 1.50 m across the entire test lane with two layers of geogrid material placed above the natural foundation and 150 mm of 21-B material placed above and below each geogrid.

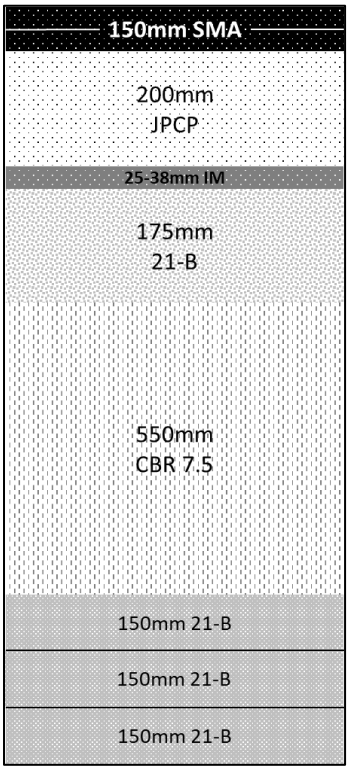


Figure 19 Pavement cross-section.

Equipment

This study utilized accelerated pavement testing to simulate traffic loading on the pavement using the Dynatest Mk VI Heavy Vehicle Simulator (HVS). Additionally, the HVS environmental chamber was used to keep the pavement surface temperature at approximately 50°F (10°C) and protect the pavement from weather effects.

This experiment expressed the traffic applied by the HVS in cumulative equivalent single axle load (ESAL). ESAL is a preferred parameter to measure traffic as it is the primary measure of traffic loading used in mechanistic-empirical pavement design methods. Theoretically, the cumulative ESALs is the summation of the equivalent axel load factor (EALF) for axle load groups multiplied by the number of passes of that axle load group for all axle load groups applied to pavement. The EALF is the ratio of pavement damage induced from a pass of the axle load group to the damage from a standard 18-kip single axle.

$$EALF = \frac{Damage\ from\ pass\ of\ axle\ in\ load\ group}{damage\ from\ a\ 18{-}kip\ single\ axle} \quad (5)$$

$$ESAL = \sum F_i N_i \quad (6)$$

Where:

F_i *is* the EALF

N_i is the number of passes for the i^{th} axle load group (*18*)

The ESALs applied by each pass of the HVS dual wheel are estimated by (*19*):

$$ESALs = Number\ of\ Passes \cdot \left(\frac{HVS\ Load\ Applied}{9\ Kips}\right)^{4.2} \quad (7)$$

The equipment used to measure the macrotexture was the Nippo Sangyo Co. Circular Track Meter (CT Meter), shown in Figure 20. The CT Meter is a static device that uses a single laser on a circular track with a radius of 142 mm to collect a two-dimensional profile. It collects 1,024 data points in eight arc segments (111.5 mm) with 128 points each (*20*). The macrotexture is reported in MPD and RMS.

Figure 20 - CT Meter

Experiment Setup

This experiment consisted of a single test lane trafficked using the HVS with the traffic loading applied at a normal distribution to simulate the effect of wander from the wheel path. The wander effect was applied by adjusting the transverse position of the HVS wheel apparatus after the prescribed number of passes. The distribution of passes in the HVS used the default wander from the Mechanistic-Empirical Pavement Design Guide and is depicted in Figure 21.

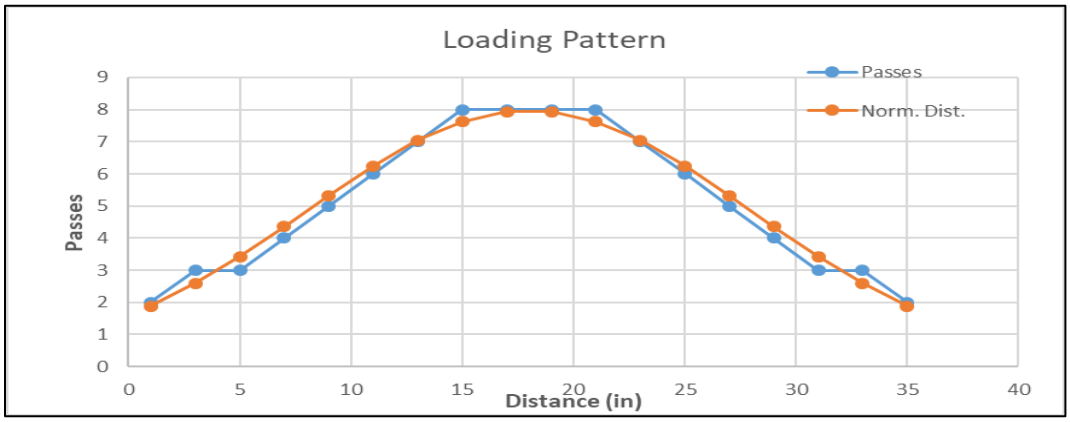


Figure 21 - HVS load distribution.

Initial macrotexture measurements were taken on the surface before traffic load was applied and were continued periodically, roughly once per week, as the cumulative traffic load increased on the pavement. Initially, daily measurements were taken; however, due to the low initial HVS load, there was little change over the initial 3 weeks of the experiment. These measurements have therefore been omitted from the analysis.

For each interval of loading at which macrotexture was measured a total of 20 measurements were taken on the traffic surface. The measurements were taken at five transverse positions along the width of the trafficked surface at each of four stations along the longitudinal length. One additional measurement was taken on a non-trafficked portion (Figure 22).

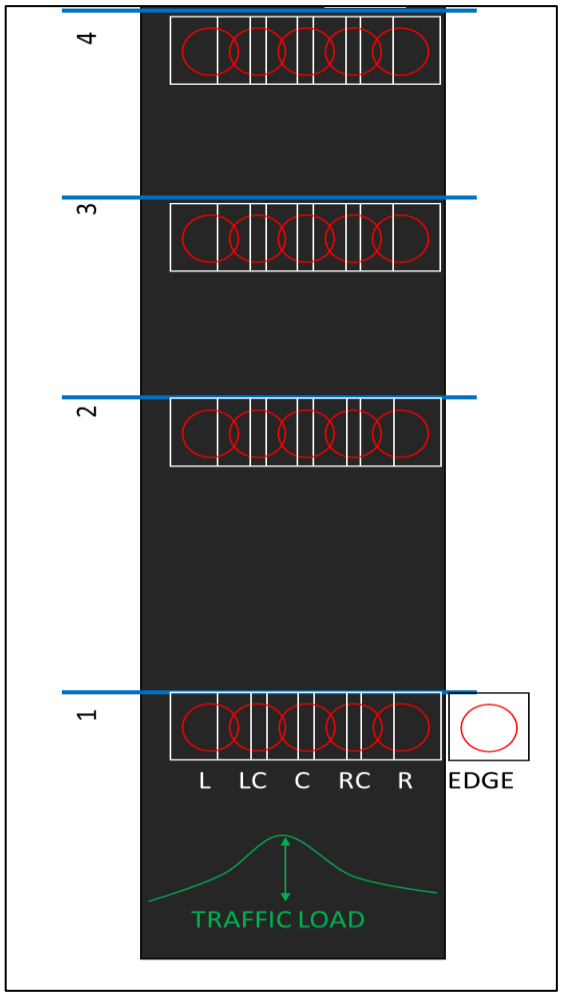


Figure 22 - Experiment measurement locations.

The red circles represent the path of the CT Meter laser. The centerline measurements identified by C in Figure 22 are of primary interest as they represent the center of the wheelpath, and therefore have the most impact to pavement-tire interactions. Additionally, network level macrotexture collection typically occurs in the wheelpath. The remainder of the measurements—L, LC, R, RC, which correspond to left, left-center, right and right-center as relative positions transversely with respect to the direction of travel—were used to determine if an overall pavement macrotexture effect occurred as a result of traffic along the entire width of wander around the wheel path. Additonally, the measurements outside of the wheelpath enabled analysis of distresses observed in those areas of the pavement. Herein, measurement locations are referred to by the station (1 through 4) and position (R, RC, C, LC, L, EDGE)—i.e., the center of station 4 would be "4C."

The methods of data analysis included statistical testing through ANOVA to determine if the application of traffic load effect on macrotexture was significant and the development of a regression model of observed behavior. JMP statistical software was used for all statistical analysis. Additionally, photographs taken at measurement locations over time as traffic load was applied were used to identify changes and distresses that developed in the pavement surface that indicated the causes of observed changes in macrotexture.

Results

Because the CT Meter produces both MPD and RMS macrotexture parameters, an initial analysis of the data to determine the most appropriate dependent variable response was conducted. The summary statistics of the data collected on the un-trafficked 1EDGE measurement show that MPD gave more consistent results with a smaller range and lower standard deviation (

Table 1). However, comparing regression models of macrotexture vs. ESALs, RMS yielded a better coefficient of determination, R^2 (Figure 23 and Figure 24).

Table 1: Comparison of MPD and RMS on Un-Trafficked Surface

Date	**Avg MPD**	**Avg RMS**
1/17/2018	0.81	0.96
1/24/2018	0.76	0.97
1/31/2018	0.79	0.96
2/8/2018	0.79	0.95
2/14/2018	0.87	1.38
2/21/2018	0.8	1.02
2/28/2018	0.8	1.02
3/7/2018	0.77	1.03
3/15/2018	0.78	1.01
3/21/2018	0.77	0.96
3/28/2018	0.82	1.15
4/3/2018	0.86	1.17
4/13/2018	0.85	0.97
4/18/2018	0.79	0.9
4/25/2018	0.8	1.03
5/10/2018	0.76	0.83
5/17/2018	0.75	0.89
Mean	**0.80**	**1.01**
Max	**0.87**	**1.38**
Min	**0.75**	**0.83**
Range	**0.12**	**0.55**
Std Dev	**0.04**	**0.13**

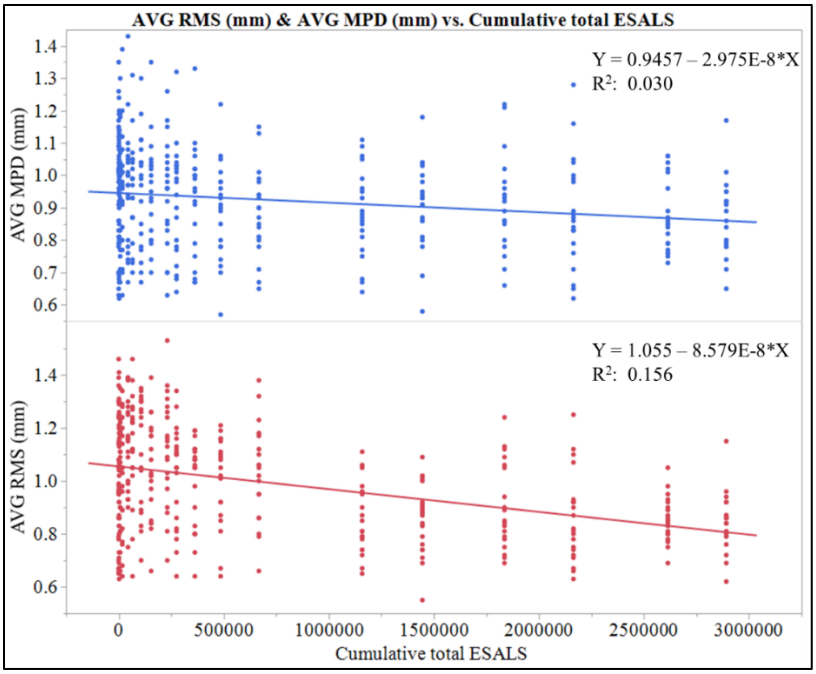


Figure 23 - Comparison of RMS and MPD for full data set.

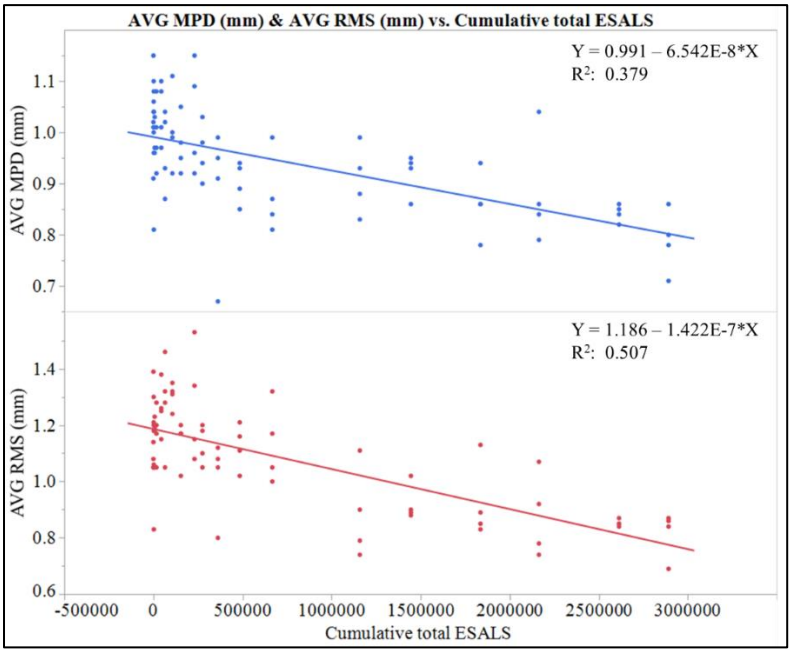


Figure 24 - Comparison of RMS and MPD for center of wheelpath only.

Because RMS yielded a better coefficient of determination for both the full data set and center of wheel path measurements, it was used as the primary macrotexture measurement for the remainder of the analysis.

Figure 23 also indicates a high degree of variability within the data collected, which was primarily due to the way in which the traffic load was applied. The measurement locations at the edges did not receive the same amount of simulated traffic as the center. It was therefore expected that the distresses would develop more quickly and have more severity closer to the center, and less so towards the edges. An analysis was conducted in which the fraction of the traffic load that was applied at the transverse positions was used as the independent variable instead of the total load; however, this analysis produced no significant improvement to the overall model and was therefore excluded from this study.

Due to the high variability in the data and the difference in how the measurement positions received the traffic load, it is helpful to visualize the data for each measurement location separately (Figure 25). While generally all the measurement locations showed a decrease in macrotexture from traffic load, the slopes vary, and in one case (4R) show an increase in macrotexture, which is discussed further in the following section.

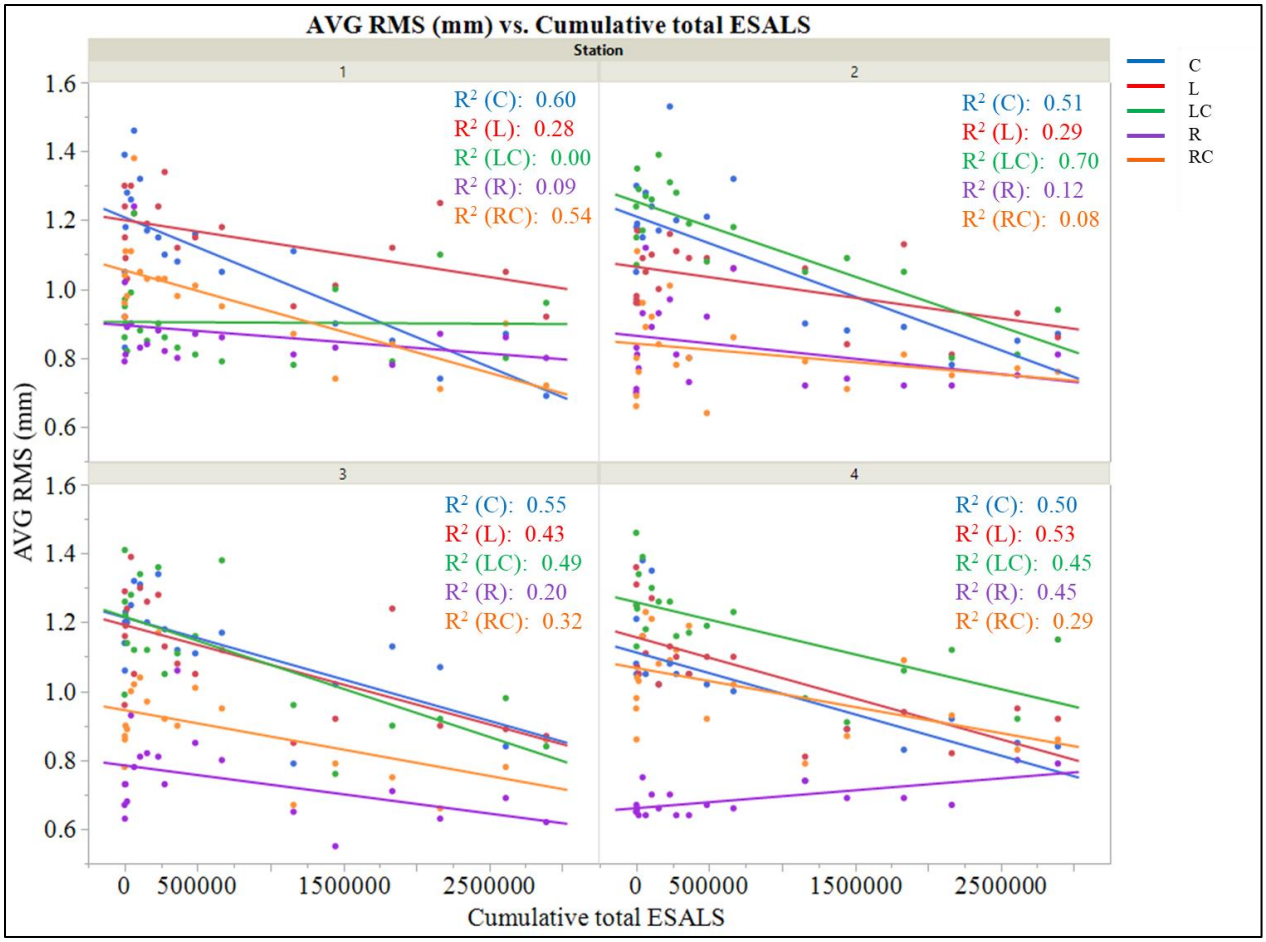


Figure 25 - Individual measurement location ESAL vs RMS.

Figure 26 shows the deterioration of the pavement over time in the center of the wheel path measurement locations. For each measurement location 1C through 4C, the picture on the left shows the pavement before traffic was applied and the picture on the right shows pavement after the final load of 2,891,714 ESALs. The middle picture shows pavement at an intermediate loading. The photographs demonstrate that as the cumulative traffic load increased, bleeding on the surface caused the gaps between the aggregates to be filled with binder, reducing the macrotexture.

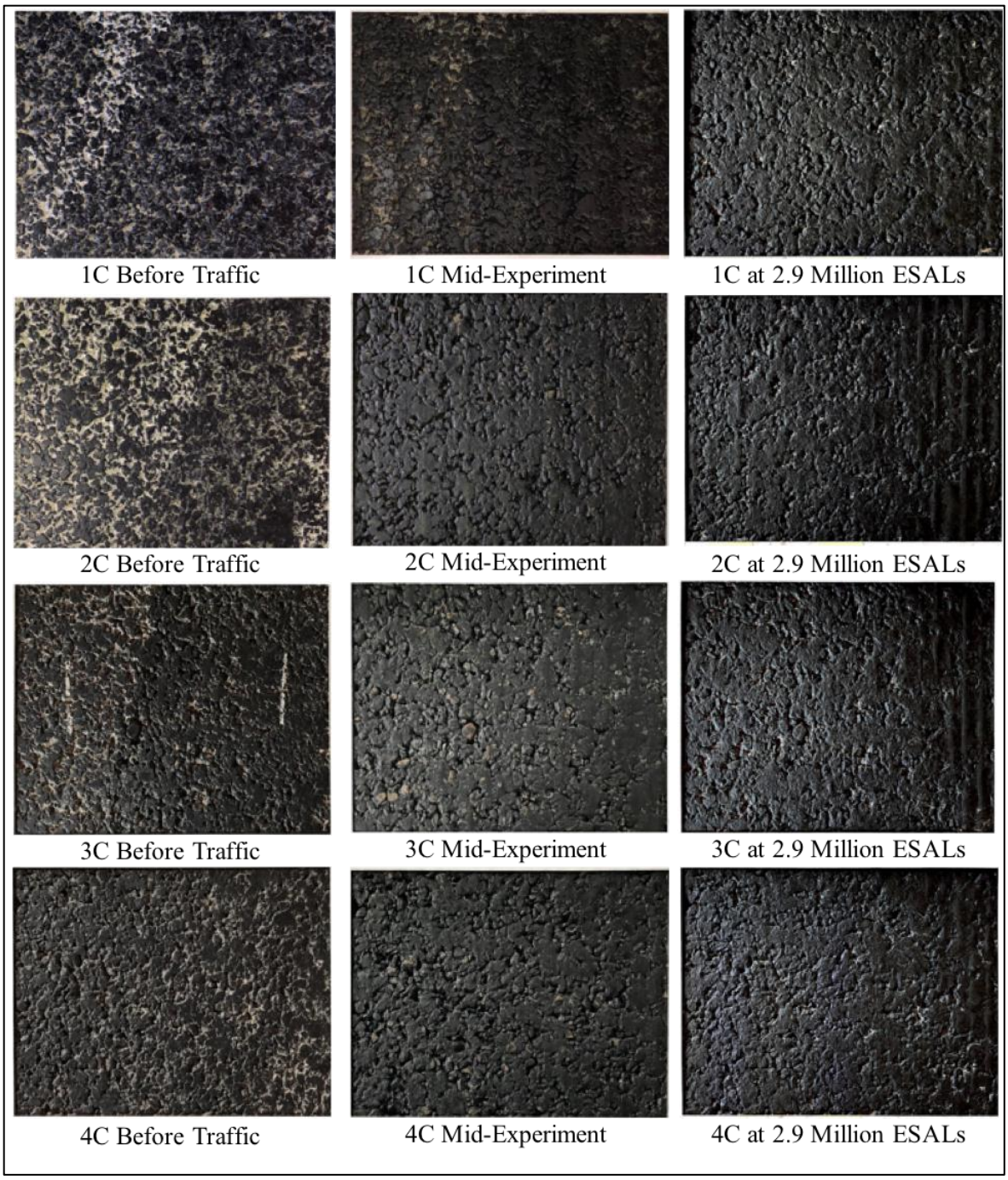


Figure 26 - Photographs of center measurements over time.

ANOVA Results

Table 2 shows the results of a 2nd degree factorial ANOVA of the full data set. The station and position of the measurements were used as blocking factors, the load in ESALs is the primary model effect, and RMS the response variable. The ANOVA showed a significant effect

of traffic load on RMS with a P-value less than 0.0001. However, the effect test of the blocking factors showed the transverse position to be significant as well with a P-value less than 0.0001, and as Figure 25 indicated, the different measurements are not completely co-directional. Additionally, the interaction terms that include the transverse position are significant, indicating that an overall effect is masked by either the variability, an unmeasured model effect, or both. For this reason, the full data set was not ideal for the development of a regression model.

Table 2: Full Data Set 2nd Degree Factorial ANOVA

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Model	27	11.1775	0.4140	30.7718	<.0001*
Error	371	4.9912	0.0134		
Cumulative Total	398	16.1687			
Effect Tests					
Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
Cumulative total ESALS	1	1	2.5326	188.2550	<.0001*
Station	3	3	0.0127	0.31570	0.814
Position	4	4	6.0141	111.7610	<.0001*
Cumulative total ESALS*Station	3	3	0.0336	0.8331	0.4763
Cumulative total ESALS*Position	4	4	0.4888	9.0834	<.0001*
Station*Position	12	12	2.1186	13.1236	<.0001*

Table 3 shows a one-way ANOVA and linear regression model of the load in ESALS and the effect on RMS with only the data collected from the center measurement locations. It shows that the effect of traffic is significant, with a P-value less than 0.0001. The regression model indicates that the macrotexture in RMS will decrease by -1.42E-07 mm per ESAL applied, or 0.142 mm per million ESALs. The coefficient of determination, $R^2 = 0.507$, indicates that a little more than half of the variability in the data is explained by the effect of traffic loading. The low R^2 is likely due to high variability in the data as well as some unique observed phenomena on the pavement surface during the study, which are discussed following.

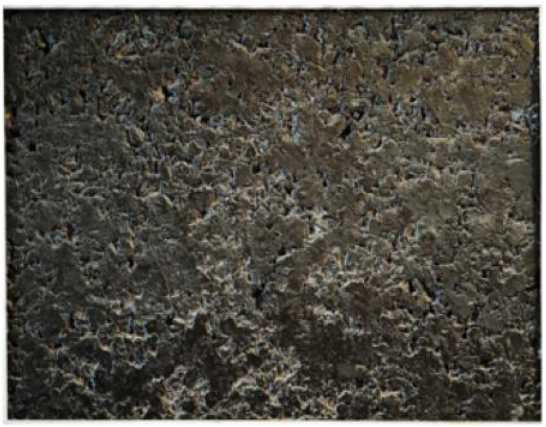
Table 3: Center of Wheelpath Data ANOVA

Summary of Fit				
RSquare	0.5072			
RSquare Adj	0.5010			
Root Mean Square Error	0.1313			
Mean of Response	1.083			
Observations	80			
Analysis of Variance				
Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	1.3838	1.3838	80.3040
Error	78	1.3441	0.0172	Prob > F
C. Total	79	2.7279		<.0001*
Parameter Estimates				
Term	Estimate	Std Error	t Ratio	Prob > |t|
Intercept	1.1864	0.0187	63.58	<.0001*
Cumulative total ESALS	-1.42E-07	1.59E-08	-8.96	<.0001*

Discussion of Results and Observed Phenomena

As previously indicated, change in macrotexture due to load can vary greatly based on the distresses that development in the pavement surface. This experiment showed that when bleeding is the primary distress that develops, the macrotexture will decrease with traffic. This is expected, as bleeding fills in the voids on the surface between aggregates. However, there were some anomalies in this study that did not follow that trend. One unique observation was that a buildup of material on the surface caused an increase in macrotexture at 4R.

Figure 27 shows photographs taken over the course of this study, where the top picture is before the application of traffic and the bottom represents the maximum load applied. These photographs clearly show the development of ridges on the pavement surface from buildup of the binder material bleeding up from the pavement matrix. Figure 28 indicates that the macrotexture as measured by RMS increased with the cumulative traffic.



4R Before Traffic

4R Mid-Experiment

4R at 2.9 Million ESALs

Figure 27 - Photographs from Position 4R.

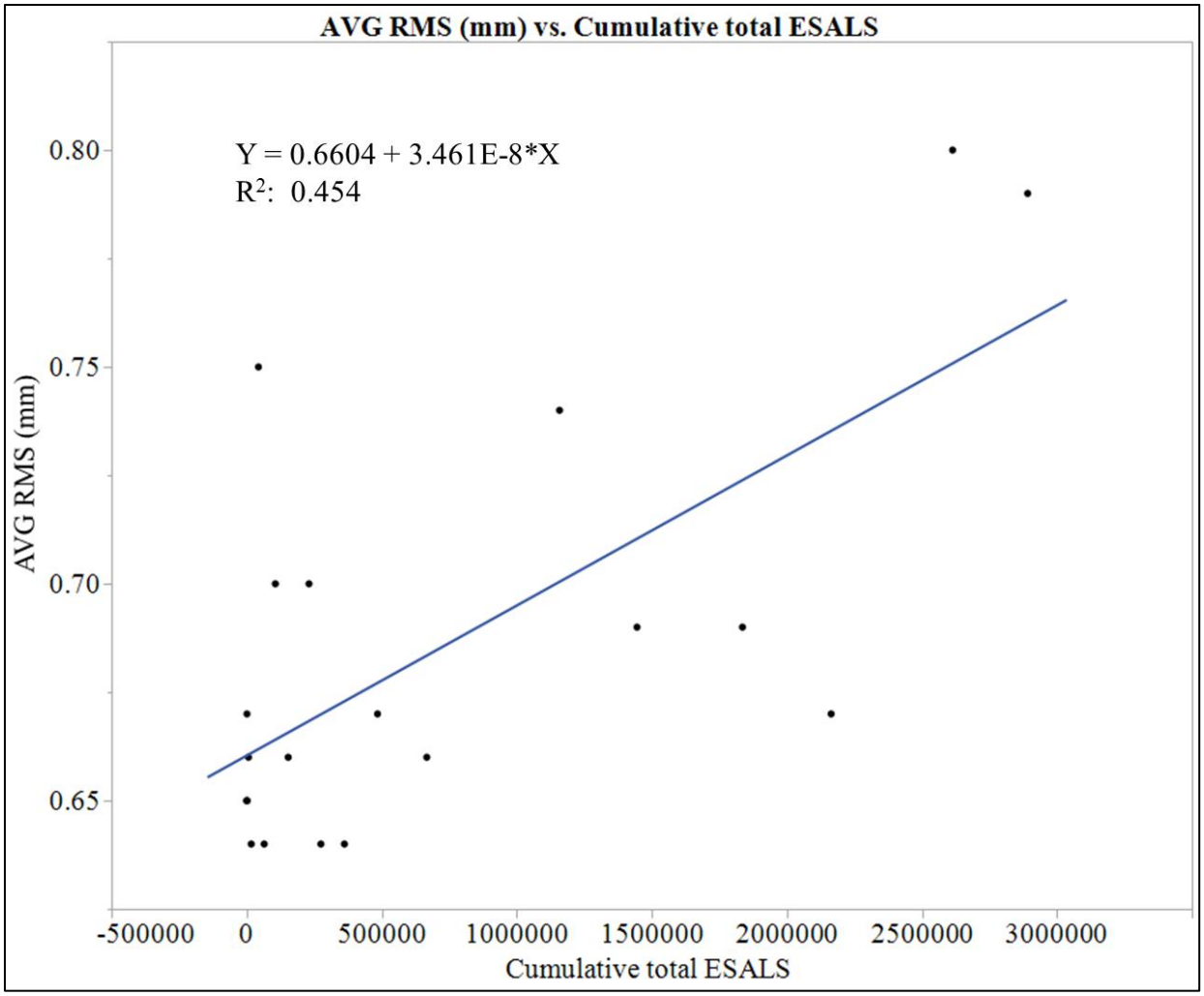


Figure 28 - ESALs vs RMS for Position 4R.

Another interesting observation was that in some areas, RMS and MPD showed opposite trends as bleeding occurred. Position 2R was the most pronounced example of this. The photographs in Figure 29 show the development of bleeding and accumulation of material on the surface forming ridges similar to position 4R. However, these ridges appear to be somewhat wider. Interestingly, the macrotexture measured in MPD shows an increase with traffic loading while the trend of RMS was a decrease of similar magnitude (Figure 30). It appears that as the buildup of material occurred, the ridges of buildup created peaks in the CT Meter profile that affected the device's calculation of MPD, which is based on peaks, as previously described. Because bleeding would be expected to lead to a decrease in friction, and friction is correlated

with macrotexture, these results indicate that RMS is a more appropriate macrotexture measure when bleeding results in surface binder accumulation, which would likely decrease friction.

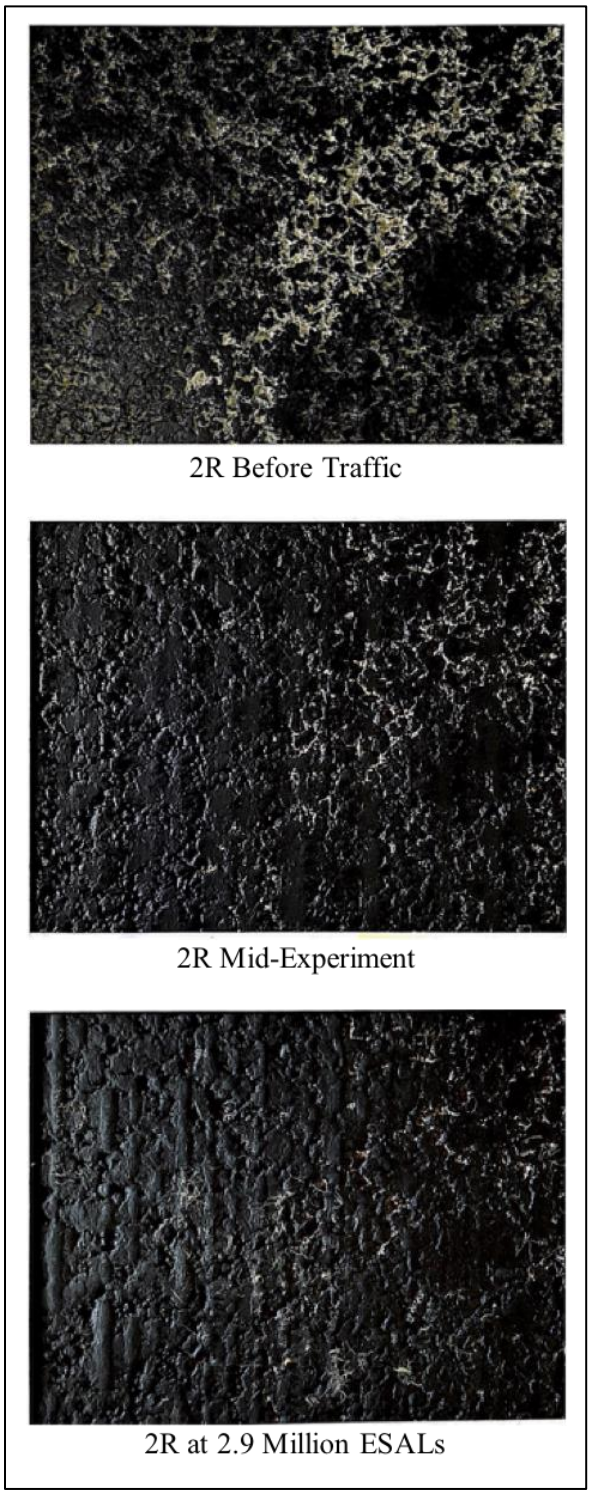


Figure 29 - Photographs from position 2R.

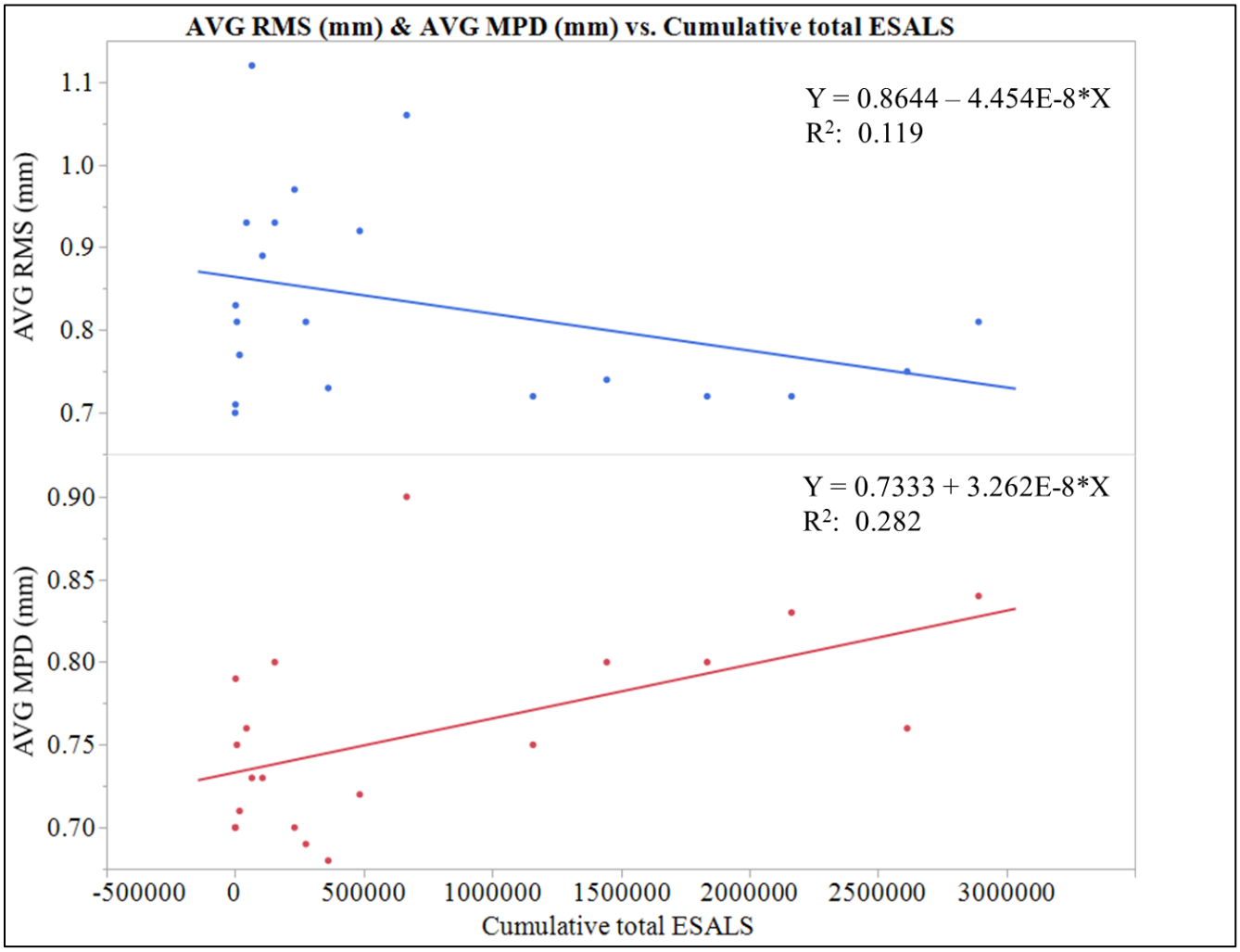


Figure 30 - Comparison of RMS and MPD for position 2R.

A final observation is that the way in which material accumulated on the pavement's surface is likely a source of the considerable variability in the data collected. In one instance near the midpoint of the study, wearing of the surface, which exposed aggregates, began to occur while binder simultaneously began to accumulate nearby (Figure 31). There is also indications that the high load used in the HVS caused rubber for the tires to accumulate on the pavement surface as shown by the wearing on the outside edge of the tires (Figure 32).

This type of varied response to the traffic load made modeling the overall change in macrotexture difficult, which is reflected in the spread of the data points and the lower coefficient of determination for the centerline position regression model.

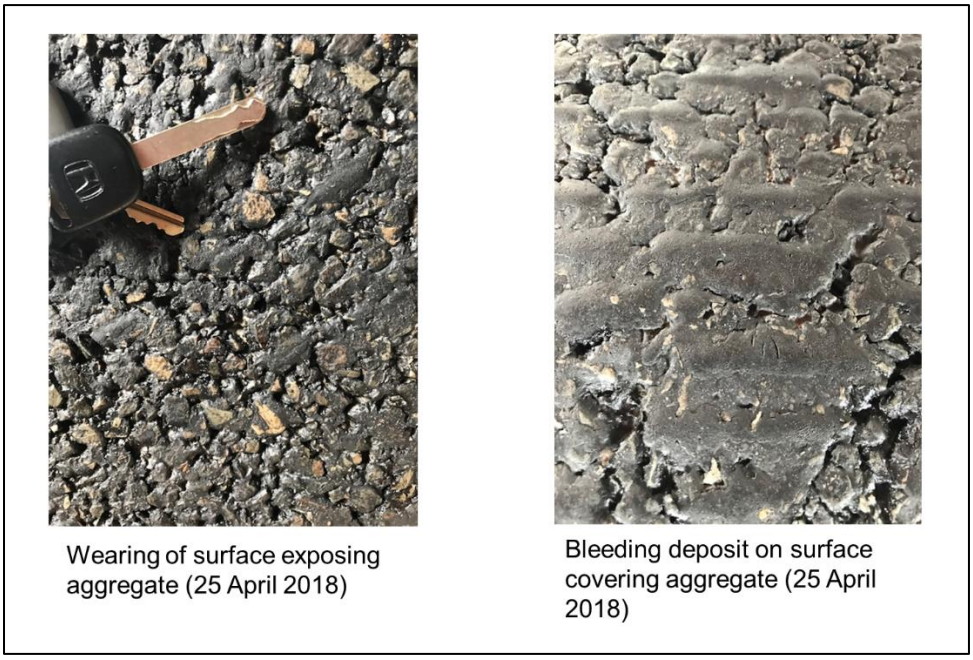


Figure 31 - Simultaneous wear and buildup of binder.

Figure 32 - Wear of HVS Tires

Summary and Conclusions

This study measured and modeled the development of bleeding and its effect on macrotexture under simulated traffic loading from an HVS. More research is needed to determine the relationship between traffic-induced bleeding and macrotexture, preferably by testing various binder contents and severities of bleeding and correlating resulting surfaces with friction. While this study showed that macrotexture decreases with bleeding due to traffic load, there are instances where it can increase due to the buildup of material from bleeding. The results also indicate that RMS is a more sensitive than MPD to macrotexture changes due to buildup of material. This is because the peaks created by the buildup may artificially inflate the readings when using MPD and would, therefore, be less likely to correlate with friction.

The possibility that the observed phenomena are unique to the experimental conditions created by the HVS and not representative of real-world observations must be considered. These experimental conditions include the temperature control, high loads and channelized nature of the HVS passes, which could have resulted in rubber accumulation from the tires contributing to the buildup of material on the surface.

While data collected outside of the wheel path was useful for modeling the bleeding observed along the edge of the pavement, the variability in the data was higher for readings taken outside of the center of the wheel path. The center of the wheel path is the measurement location typical for network level macrotexture collection and this study indicates that the measurements within the wheel path are more predictable as load is applied and are subject to less variability in measurements than other areas on the pavement surface.

Acknowledgement

The authors would like to express our gratitude to Brian Diefenderfer for his oversight of accelerated pavement testing at Virginia Tech Transportation Institute.

Author Contributions

The authors confirm contribution to the paper as follows: study conception and design: K. Maeger, V. Bongioanni, Gerardo Flintsch; data collection: K. Maeger, V. Bongioanni; analysis and interpretation of results: K. Maeger, V. Bongioanni; draft manuscript preparation: K.

Maeger, V. Bongioanni. All authors reviewed the results and approved the final version of the manuscript.

References:

[1] Wambold, J., C. Antle, J. Henry, Z. Rado, G. Descornet, U. Sandberg, M. Gothié, and S. Huschek. International PIARC Experiment to Compare and Harmonize Skid Resistance and Texture Measurements (Paris: PIARC) Publication n 01.04.In, PIARC, Paris, France, 1995.

[2] Weir, D., J. Strange, and R. Heffley. Reduction Of Adverse Aerodynamic Effects Of Large Trucks Volume I: Technical Report.In, 1978.

[3] Parry, A., and H. Viner. Accidents And The Skidding Resistance Standard For Strategic Roads In England.In *TRL report*, TRL Limited, Wokingham, Berkshire, UK., 2005.

[4] Roe, P. G., A. R. Parry, and H. E. Viner. High and low speed skidding resistance: the influence of texture depth.In, TRL Limited, Crowthorne, Berkshire, U.K., 1998. p. 22p.

[5] Sandberg, U., A. Bergiers, J. A. Ejsmont, L. Goubert, R. Karlsson, and M. Zöller. Road surface influence on tyre/road rolling resistance. *Models for Rolling Resistance in Road Infrastructure Asset Management Systems (MIRIAM),(http://miriam-co2. net/Publications/MIRIAM_SP1_Road-Surf-Infl_Report,* Vol. 20111231, 2011.

[6] Descornet, G., B. Faure, J. Hamet, X. Kestemont, M. Luminari, L. Quaresma, and D. Sandulli. Traffic noise and road surfaces: state of the art. *Belgian Road Research Centre, Brussels*, 2000.

[7] PIARC. *Road Dictionary*. http://www.piarc.org/en/Terminology-Dictionaries-Road-Transport-Roads/. Accessed 16 December 2016.

[8] Hall, J. W., K. L. Smith, L. Titus-Glover, J. C. Wambold, T. J. Yager, and Z. Rado. Guide for Pavement Friction. 2009, p. 257p.

[9] ASTM. E1845 Standard Practie for Calculating Pavement Macrotexture Mean Profile Depth. *ASTM International*, 2015.

[10] ISO. ISO 13473-2 Characterization of pavement texture by use of surface profiles. Vol. Part 2 Terminology and basic requirements related to pavement texture profile analysis.

[11] Aktaş, B., M. Karaşahin, and M. Tığdemir. Developing a macrotexture prediction model for chip seals. *Construction and Building Materials,* Vol. 41, 2013, pp. 784-789.

[12] Roque, R., D. Anderson, and M. Thompson. Effect of Material, Design, and Construction Variables on Seal-Coat Performance. *Transportation Research Record No. 1300, Asphalt Pavement and Surface Treatments: Construction and Performance 1991*, 1991.

[13] Riemer, C., and D. Pittenger. Modeling Pavemnt Texture Deterioration as a Pavement Preservation Management System Tool. *TRB*, 2012.

[14] Lin, C., and W. Tongjing. Effect of fine aggregate angularity on skid-resistance of asphalt pavement using accelerated pavement testing. *Construction and Building Materials,* Vol. 168, 2018, pp. 41-46.

[15] Powell, R. B., and S. Buchanan. Long Term Performance of a Thin Asphalt Overlay on the NCAT Pavement Test Track. *TRB*, 2012.

[16] Woodward, D., P. Millar, C. Lantieri, C. Sangiorgi, and V. Vignali. The wear of Stone Mastic Asphalt due to slow speed high stress simulated laboratory trafficking. *Construction and Building Materials,* Vol. 110, 2016, pp. 270-277.

[17] Virginia Department of Transportation. Pavement Design Guide for Subdivision and Secondary Roads In Virginia. 2014.

[18] Huang, Y. H. *Pavement Analysis and Design, 2nd Ed.* Pearson Prentice Hall, 2004.

[19] D. Jones, R. J. H. Reflective Cracking Study: First-LEvel Report on HVS Testing on Section 586RF -- 45 mm MB15-G Overlay.In, University of California Pavement Research Center UC Davis, UC Berkeley, 2006.

[20] Nippo Sangyo Co. *Circular Track Meter Product Guide*. http://www.nippou.com/en/products/ct.html. Accessed 1 June 2018, 2018.

CHAPTER 4: REPEATABILITY AND AGREEMENT OF VARIOUS HIGH-SPEED MACROTEXTURE MEASUREMENT DEVICES[2]

Abstract

Five high-speed macrotexture measurement devices were tested on a variety of asphalt and Portland Cement Concrete pavement surfaces to evaluate their repeatability and their pairwise agreement. Experiments were run under three speed conditions: highway speed, varying constant speeds, and various acceleration and deceleration profiles. Data were processed and reduced per current industry standards. A novel approach for dropout and outlier removal from line laser devices was developed. The pairwise device agreement was evaluated using a Limits of Agreement analysis. The results demonstrate good repeatability for each of the devices tested. The agreement analysis showed that not all high-speed distance triangulation devices can be used interchangeably for all pavement surfaces. Data acquisition speed was found to be a factor in macrotexture parameter calculation for two of the five devices. The effect of speed was found to be worse on randomly textured surfaces than on transversely textured surfaces. Finally, acceleration was shown to have an effect on the parameters produced by one of the devices, giving further credence to the fact that care should be taken to gather high-quality datasets for the critical pavement characteristic of macrotexture.

Keywords: Macrotexture, Limits of Agreement, Repeatability, MPD, Surface

[2] This manuscript was coauthored with Vincent Bongioanni, Dr. Samer Katicha, Dr. Edgar D. de León Izeppi and Dr. Gerardo Flintsch. It was submitted to the Transportation Research Board and was selected for presentation and is being considered for publication.

Introduction

Background

Various US and international standards exist to calculate parameters that describe a road surface's macrotexture. However, the US currently has no standard for gathering network-level macrotexture measurements. Several high-speed devices were tested on the Virginia Smart Road to assess their ability to characterize the macrotexture of various road surfaces.

Typical means of data comparison, such as regression analysis with associated correlation coefficients, can be misleading in comparing datasets from two or more devices. Correlation only shows how strongly the data are related (*1*), and depends on both the measurement errors and the true variability of macrotexture. For data to be well-correlated, it just needs to fall along a straight line or follow any selected function. Therefore, correlation is not appropriate to determine the repeatability (i.e., its ability to produce the same results under the same conditions) of a given device.

In the past, the typical approach was to perform harmonization studies to allow the results of pavement surface interaction characteristics from different devices (*2-5*) to be compared. However, the true value of the parameter measured is unknown for the entire surface. One could try to use a mechanical stylus or perform sand patch tests on an entire wheel path, but manual tests have been known to be susceptible to operator error (*6*).

Another issue with harmonization studies is that results are often difficult to replicate (*7-9*). Given the shortcomings of harmonization between macrotexture measurement devices, another method is needed. De Leon, Flintsch and McGhee (*10*) showed that valid comparisons between various friction measuring devices can be made in the absence of a true value via agreement methods first used in the biopharmaceutical field (*11-15*). These methods assess the agreement or closeness between readings of the devices evaluated (*11*). Katicha et al. (*16)* used this method to compare Falling Weight Deflectometer (FWD) and Traffic Speed Deflectometer (TSD) measurements. However, no similar such study has been carried out for many macrotexture measurement devices.

Problem Statement

No guidance exists for network-level macrotexture data collection in the US. Comparison of devices to measure macrotexture can be difficult as the true macrotexture is unknown for the entire roadway. Furthermore, correlation comparisons can be misleading for the purposes of device comparisons (*15*).

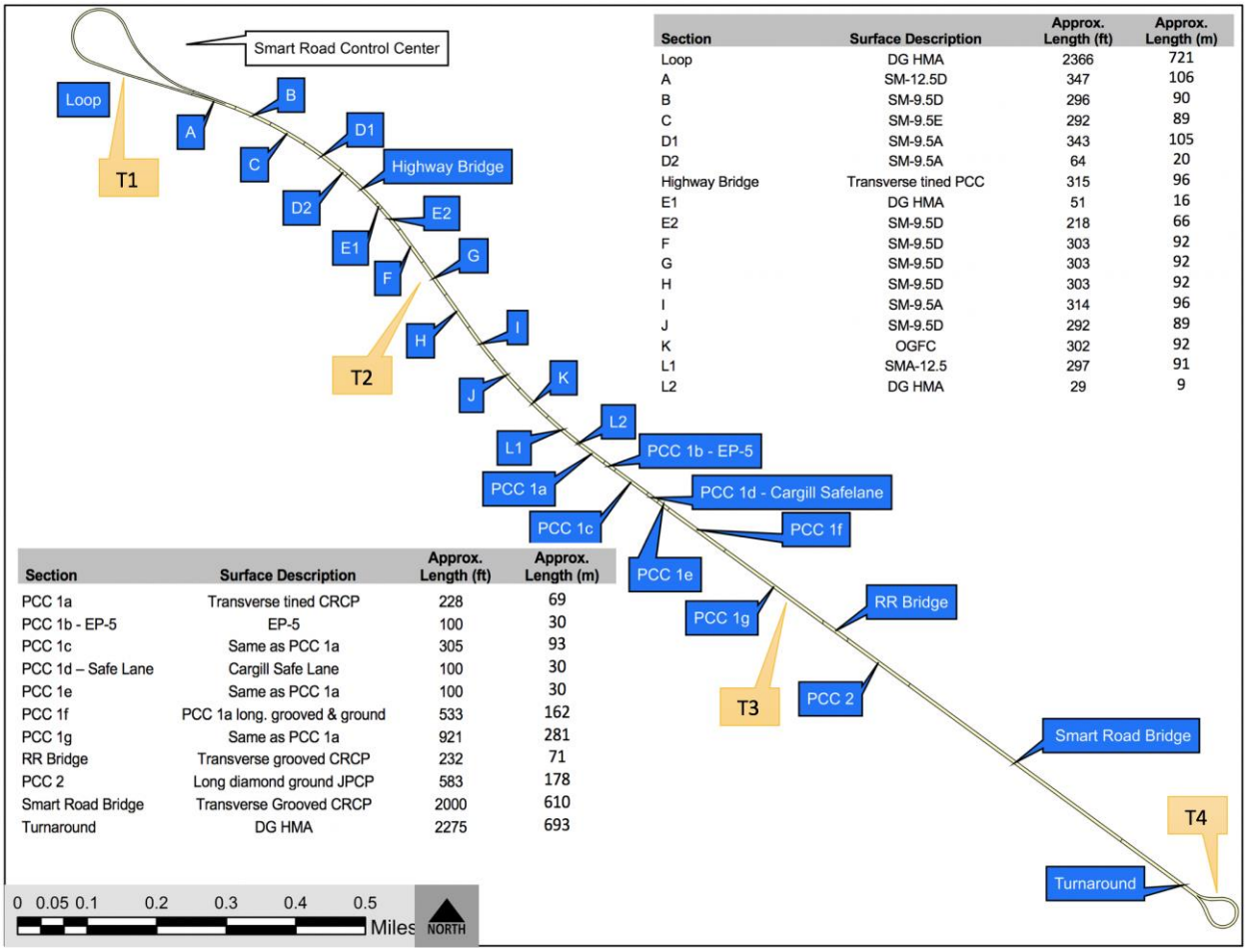
Objective

The objective of this study is to evaluate the ability of various macrotexture measurement devices to measure typical pavement surfaces to determine if they can be used interchangeably within an acceptable range. As network-level measurements are envisioned to be made on a reoccurring basis, the ability of a device to produce consistent results on given surfaces (its repeatability) is studied. Finally, since data collection does not occur in an ideal world, the effects of operational factors such as speed, acceleration, and deceleration on the measurements are also evaluated.

Methodology

Experiment Setup

Vehicles with high-speed macrotexture scanning equipment were used to make measurements on the Virginia Smart Road. The Smart Road is a 2.2-mile test track dedicated to research located in Blacksburg, Virginia. It offers a variety of surfaces, including Dense Graded Hot Mix Asphalt (DGHMA), an Open-Graded Friction Course (OGFC), proprietary High-Friction Surface Treatments (HFST), as well as grooved, tined, or ground Portland Cement Concrete (PCC) sections as shown in Figure 33.



Section	Surface Description	Approx. Length (ft)	Approx. Length (m)
Loop	DG HMA	2366	721
A	SM-12.5D	347	106
B	SM-9.5D	296	90
C	SM-9.5E	292	89
D1	SM-9.5A	343	105
D2	SM-9.5A	64	20
Highway Bridge	Transverse tined PCC	315	96
E1	DG HMA	51	16
E2	SM-9.5D	218	66
F	SM-9.5D	303	92
G	SM-9.5D	303	92
H	SM-9.5D	303	92
I	SM-9.5A	314	96
J	SM-9.5D	292	89
K	OGFC	302	92
L1	SMA-12.5	297	91
L2	DG HMA	29	9

Section	Surface Description	Approx. Length (ft)	Approx. Length (m)
PCC 1a	Transverse tined CRCP	228	69
PCC 1b - EP-5	EP-5	100	30
PCC 1c	Same as PCC 1a	305	93
PCC 1d – Safe Lane	Cargill Safe Lane	100	30
PCC 1e	Same as PCC 1a	100	30
PCC 1f	PCC 1a long. grooved & ground	533	162
PCC 1g	Same as PCC 1a	921	281
RR Bridge	Transverse grooved CRCP	232	71
PCC 2	Long diamond ground JPCP	583	178
Smart Road Bridge	Transverse Grooved CRCP	2000	610
Turnaround	DG HMA	2275	693

Figure 33 - Virginia Smart Road sections and surface types.

Data Collection

Data were collected using five vehicles, each fitted with laser triangulation sensors capable of collecting pavement macrotexture information at the rates shown in Table 4. All data were gathered in the left wheelpath of the lane that runs from east to west. Data were received in the raw spatial form from the operators. Height measurements were given with the fixed sampling interval shown in Table 4. In the case of Device 3, the authors interpolated time-domain data into spatial-domain data according to the maximum interval observed for the speeds used. The operators performed no filtering or outlier correction. Raw data were requested for an objective comparison of all data by applying the same analysis methods to remove variables such as filtering, outlier correction, and method of parameter calculation.

Table 4: Data Collection Equipment Information

Device ID	Laser Type	Sampling Frequency	Raw Data Spatial Interval
1	Single Spot	100 kHz	0.25 mm
2	Single Spot	32 kHz	1 mm
3	Single Spot	32 kHz	0.9 mm
4	Single Spot	100 kHz	0.5 mm
5	Line Laser	5 kHz	0.3 mm (transverse) 25 mm (longitudinal)

The Virginia Smart Road was prepared by placing rubber strips at or near the pavement section transitions of the surfaces tested. These were used to accurately identify the boundaries between the different surfaces. Traffic cones with a strip of reflective tape along the entire height were also placed on the road's shoulder to trigger photocells to mark where each transition occurred. Operators calibrated their distance measurement equipment at 105 km/h using survey points placed on the road for that purpose.

Three sets of tests were completed at high speed and are summarized in Table 5. The first test, High Speed, aimed at studying the overall repeatability and agreement of the various devices. The high-speed test commenced with the test vehicle on the Virginia Smart Road turnaround (labeled "T4" in Figure 33). The vehicle subsequently accelerated to the test speed and remained at this speed (typically through the use of cruise control) across all the surfaces tested. To study the possible effect of speed on a device, the second test, Constant Speed, started at T3 and ended at T2. This test involved a PCC section and an AC section. The vehicle commenced at T3, accelerated to the test speed and remained at that speed for the duration of the test. The speeds listed in Table 5 were used in the constant speed test. The third test, Variable Speed, evaluated the effect of acceleration on measurements. It commenced at Section H and ran through Section F. As shown in Figure 33, the surface type is the same for these sections, so they can be considered as one section. Four speed profiles were used, and drivers were instructed to achieve the initial speed from T3 and then perform the test action (acceleration or deceleration) by the mid-point of the test section, which was marked by a cone. The vehicles then proceeded through the remainder of the test section at that speed or accelerated to the next required speed by the end of the test section.

Table 5: Summary of Tests Performed

Test Name	Speed	Number of Runs	Section Tested
High Speed (HS)	55 mph (89 km/h)	5	Smart Road Bridge through Section C
Constant Speed (CS)	15 mph to 65 mph (24 km/h to 105 km/h) in 10 mph (16 km/h) increments	4	PCC 1a, Section L
Variable Speed (VS)	25 mph (40 km/h) speed up to 50 mph (80 km/h) 50 mph (80 km/h), slow to 25 mph (40 km/h) 50 mph (80 km/h), full stop, 50 mph (80 km/h) 50 mph (80 km/h), slow to 25 mph (40 km/h), speed up to 50 mph (80 km/h)	4	Section H through F

Data Preparation

For the high-speed runs, data were collected without interruption along the length of the road. To break these datasets into individual sections (i.e., various surface types), reflective cones were placed to trigger event markers using the various devices' photocell triggers, and bump strips of known dimensions were placed across the width of the lane. The reflective cones did not always trigger an event marker for all devices, so the bump strips were used to locate the begging and end of each section definitively. These bumps were meticulously located in the data for each of five runs of the 18 sections of interest for all five devices. This ensured that each section's set of data had a common starting point across devices and runs to minimize the effect of vehicle wander on profile lengths compared to simply using the overall test start and stop points. The bumps were similarly located for the constant speed and variable speed tests data.

Once sections were parceled out, 3 meters were cut off the beginning and end of all data. This was done to ensure that the bump strip would not affect the data analysis. The length of 3 meters was selected also to minimize the effect of inconsistencies in the road surface. For example, steel bridge expansion joints, areas of PCC that could not receive grinding/grooving to match subsequent pavement, and abrupt transitions between PCC/AC or surface treatments were all removed in this manner. This provides a more consistent surface profile from which to calculate macrotexture parameters.

The mean profile depth (MPD) was selected as the parameter to represent the macrotexture. This is because it is widely used in both the US and abroad. An initial analysis of variance (ANOVA) of repeated runs of the same devices showed that aggregating MPD to a 1-meter or greater reporting length resulted in failure to reject the null hypothesis, indicating that means are equal for each run. For this reason, MPDs aggregated to 1 meter are used for the entirety of the analysis in this work. Aggregation reduces the negative impacts of slight vehicle wander, operator experience, and misalignment of individual runs.

Outlier Removal

As noted in various standards for calculating MPD (*17; 18*), outliers in the data must first be removed. This is especially important for MPD calculations as MPD is sensitive to large data variations. All data from single-spot laser devices were treated with the same adaptive outlier removal routine as described by Katicha et al. (*19*). Figure 34 provides an example of profiles before and after outlier removal.

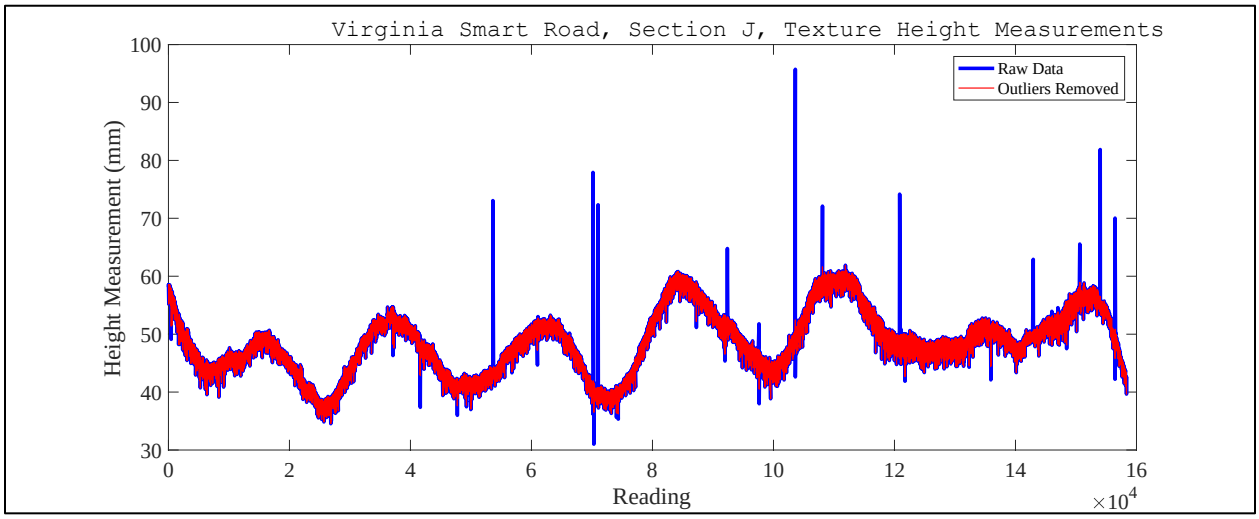


Figure 34 - Example outlier removal of single-spot lasers.

The data from Device 5 presented a challenge that inspired a novel approach to dropout and outlier removal from wide-footprint data sets. Relatively large quantities of dropouts (data points not recorded by the device at the specified sampling time) were found in the raw data

from the line laser. These dropouts occurred most typically along the edges of the laser footprint or, at times, across the entire width of the laser for one or more samples. To overcome these missing data and proceed to subsequent outlier removal, transverse rows of data were completely removed if more than 10%, as suggested in ASTM E1845 (*17)*, of the data points were dropouts. Next, columns of data outside of the central 100-mm footprint (the base length for our MPD calculation) were removed, which eliminated the majority of remaining dropouts. Then, for the remaining rows, the mean value of valid measurements was calculated, and the dropouts were then replaced with this mean value. The dataset was now ready for outlier removal. The authors wanted to use a consistent approach to outlier removal for all datasets. The method (*19*) we employed is a statistics-based approach that adapts to the given dataset to set an appropriate threshold that guarantees that no more than a certain percentage (10% in our case) of outliers are incorrectly identified. Since the quality of this test increases with the size of the dataset, consecutive rows were concatenated according to equation 1 to create a dataset most similar in size to those of the single-spot lasers for their outlier elimination.

$$\begin{bmatrix} \mathbf{h_{1,1}} & \mathbf{h_{1,2}} & \dots & \mathbf{h_{1,n}} \\ \dots & & & \\ \mathbf{h_{m,1}} & \mathbf{h_{m,2}} & \dots & \mathbf{h_{m,n}} \end{bmatrix} \xrightarrow{\textbf{yields}} [\mathbf{h_{1,1}} \quad \dots \quad \mathbf{h_{m,n}}] \tag{1}$$

To minimize edge effects in the concatenation, a simple linear regression subtracted from the row of interest suppressed the slope and set the mean of the row to zero. After outlier removal, the dataset is reconstructed to the original matrix dimensions. The result (Figure 35) was a 3-D profile (of non-homogeneous spatial distances) with outliers removed whose profile was untouched in areas not containing outliers.

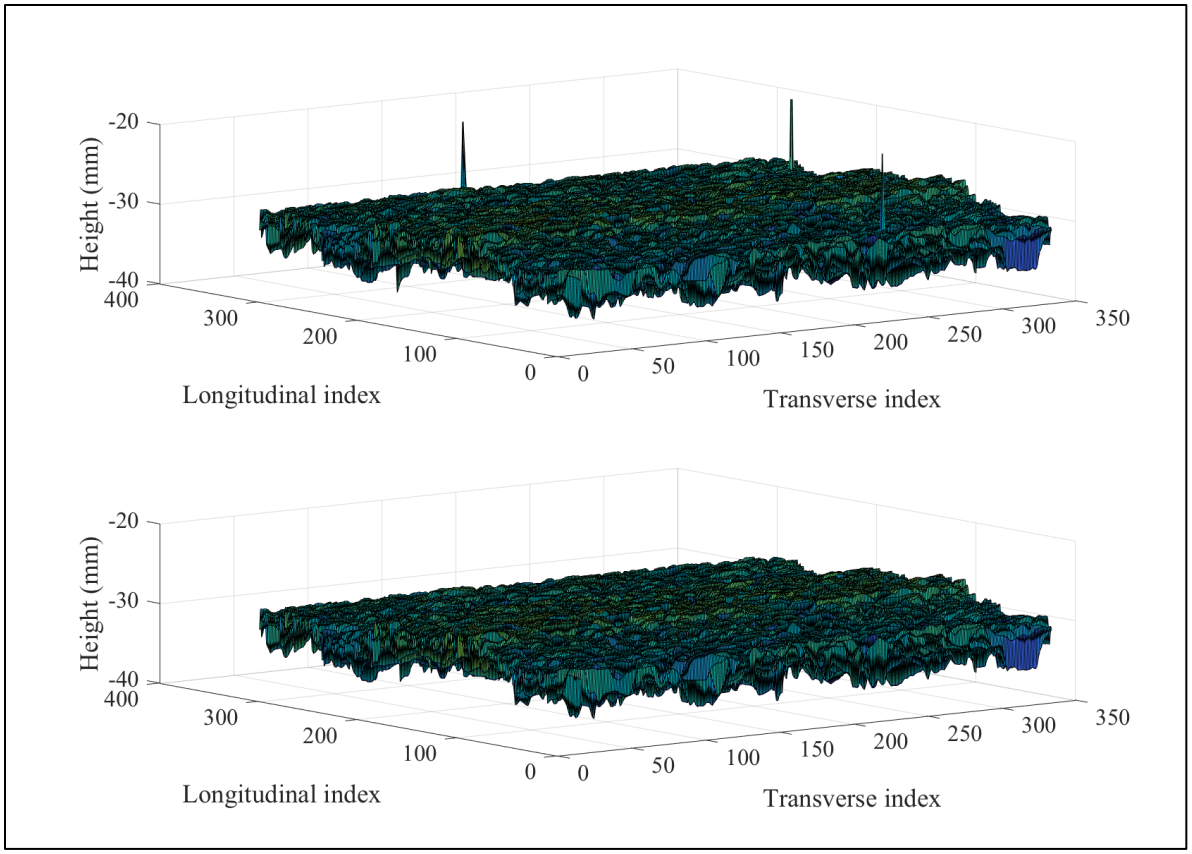


Figure 35 - Example outlier removal of line laser.

After outlier removal, all data were filtered to ensure only the wavelengths of interest for macrotexture are evaluated. A low-pass, lowest-order Butterworth Infinite Impulse Response filter was designed for each device based on the sampling interval used according to the guidance in ASTM E1845 (*17*). Finally, MPDs were calculated for every 100 mm of longitudinal travel in the left wheelpath according to ASTM E1845. For Device 5, MPD values were calculated in the transverse direction and then aggregated to 100 mm of longitudinal travel in the left wheelpath by averaging the 4 MPD values taken within the 100mm at the longitudinal spacing of 25 mm.

Device Comparison

The repeatability of each device and the agreement (bias) in measurement among the devices were evaluated. Conducting an ANOVA can show if means are equal for multiple runs of a device, where a high *p*-value corresponds to a failure to reject the null hypothesis, meaning that the means of repeated measurements taken are equal. However, this approach is limited

when comparing devices to each other because a device with high variance in collected data tolerates a more substantial difference in means before rejecting the null hypothesis that means are equal than a device with lower variance.

Repeatability

Repeatability determines the extent to which a device can reproduce previous results of the same device. This is accomplished quantitatively by calculating a device's repeatability coefficient. The repeatability coefficient is derived from the device's mean square error (MSE) for several runs over the same pavement section, where MSE is essentially the variance of the device. The repeatability coefficient is derived from the within-device standard deviation as proposed by Bartlett and Frost (*20)*:

$$c_r = 1.96 * \sqrt{2} * SD \quad (2)$$

As multiple runs were made with each device, SD (the within-device standard deviation) in equation 2 is calculated by taking the square root of the MSE after performing an ANOVA. Two measurements made on a subject by a device should differ by no more than the repeatability coefficient 95% of the time assuming a normal distribution of differences between measurements (*20*). For each device, the ANOVA was conducted for the High-Speed data with pavement section as the input and MPD as the model effect. For all the analyses, 1-meter aggregated data collected over the pavement sections were averaged to a single average MPD for each pavement section for each of the five runs.

Device Agreement

The second analysis technique employed was Limits of Agreement (LOA), which is presented by Bland and Altman (*12*) as a superior device comparison method for correlation coefficients because a strong correlation between devices does not necessarily guarantee strong agreement between them. Five runs of High-Speed data were collected over the 18 pavement sections as shown in Table 5. The LOA method used assumes that the true value measured does not change and that measurements are made in quick succession to minimize effects caused by changes in experimental conditions. To ensure that the true value of the pavement macrotexture does not change, the standard deviation of runs for each section should be constant and not related to the magnitude of the measurements. This is checked by plotting the standard deviation

against the mean for each device (*14*). Figure 36 is an example of this check for Device 4 and appears to show no clear relation between standard deviation and magnitude, which suggests the constant standard deviation assumption can be used.

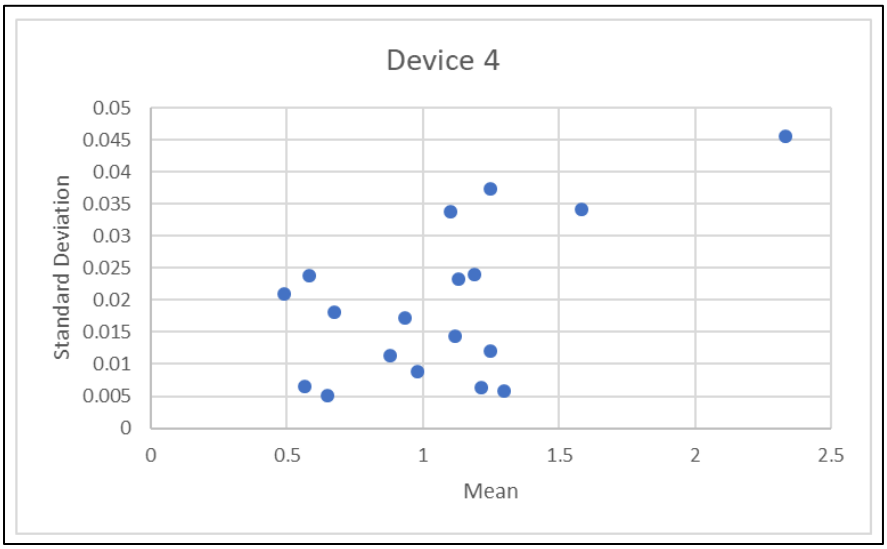


Figure 36 - Determination of constant means for Device 4.

The 95% confidence LOA are calculated as prescribed by Bland and Altman (*14*):

$$LOA = 1.96 * S_c \quad (3)$$

where:

S_c is the corrected standard deviation of differences = $\sqrt{S_D^2 + f_1 \cdot S_1^2 + f_2 \cdot S_2^2}$ (4)

S_D = standard deviation of the difference between the mean of the runs for each road section by two devices compared;

f_1 and $f_2 = 1 - \frac{1}{m}$ (m is the number of runs for each section) (5)

Each section had 5 runs; therefore, $f_1 = f_2 = 0.8$.

S_1 and S_2 are the variances of the devices (MSE determined from an ANOVA of the device with the pavement section as the model input and average MPD as the response).

The boundaries for LOA are the mean of the differences for the devices over all the sections ± the LOA. The LOA are plotted with the mean of the two devices' measurements on the *x*-axis and the difference between the devices on the *y*-axis.

Operational Factors Test

For the Constant-Speed tests, an ANOVA was conducted for both pavement sections with speed as a continuous variable model input and MPD as the response. Using a significance level of 0.05, a failure to reject the null hypothesis demonstrates no effect of speed, and rejection of the null hypothesis demonstrates an effect of speed.

For the Variable-Speed tests, an ANOVA was conducted with the four various acceleration conditions as the model input and MPD as the response. Using a significance level of 0.05, a failure to reject the null hypothesis demonstrates no effect of variable speeds, and rejection of the null hypothesis demonstrates an effect of the variable speeds.

Results

To begin the analysis, a two-way factorial ANOVA was performed on the entire data set of 55 mph runs of all devices. In this analysis, the device, pavement section, and interaction term of device*section are all used as model effects with average MPD for the section as the response. The resulting *p*-values (all < 0.0001) indicated that all model effects were significant to the analysis. The effects test of the device was used to determine if any of the devices differ. From the two-way ANOVA, a Tukey's Honest Significant Difference (HSD) test with connecting letters report from JMP software was used to make a preliminary evaluation of device agreement. It should be noted that Tukey's HSD is a very sensitive test and small differences may show as disagreement of an entire device, as shown in Table 6. Further light is shed in subsequent pairwise comparisons in this work.

Table 6: Summary of Tukey HSD Comparison

Device ID				Least Square Mean of MPD
5	A*			1.088
1		B*		1.073
4			C	1.059
2			C	1.057
3			D*	0.944

**Devices not connected by the same letter are significantly different according to the Tukey's HSD test*

Repeatability

The calculated coefficients of repeatability are found in Table 7 for the data analyzed according to the procedures described above. C_r for all devices tested was in a similar range (from 0.063 to 0.088 mm), meaning that measurements on pavement types similar to those tested by the devices used will differ by no more than the repeatability coefficient (i.e., 0.063 mm) on 95% of occasions. This is thought to be sufficient since the repeatability of sand patch tests has been found to vary by approximately 24% of the mean value of each test location in another study (*21*), indicating that these high-speed, non-contact means are an improvement to manual means.

Table 7: Summary of Device Repeatability

	Device 1	**Device 2**	**Device 3**	**Device 4**	**Device 5**
MSE	7.3 E-4	6.7 E-4	10.07 E-4	5.27 E-4	5.2 E-4
$\boldsymbol{c_r}$	0.075	0.072	0.088	0.064	0.063

Device Agreement

The mean of all 1-meter MPDs for each run of a particular pavement section was calculated. The results are plotted in Figure 37 for each of the devices tested. From the figure, it is easy to discern that the results were similar for most of the surfaces tested. For example, PCC1e had relatively low MPDs, and asphalt section K (an open-graded friction course) had relatively high MPDs reported by all devices. It is also apparent that PCC1f (a longitudinally ground and grooved PCC section) yielded much higher results than the other four devices tested

for each of its runs. PCC2 shows a similar but less extreme trend. The next task is to quantify the agreement (or lack thereof) between devices.

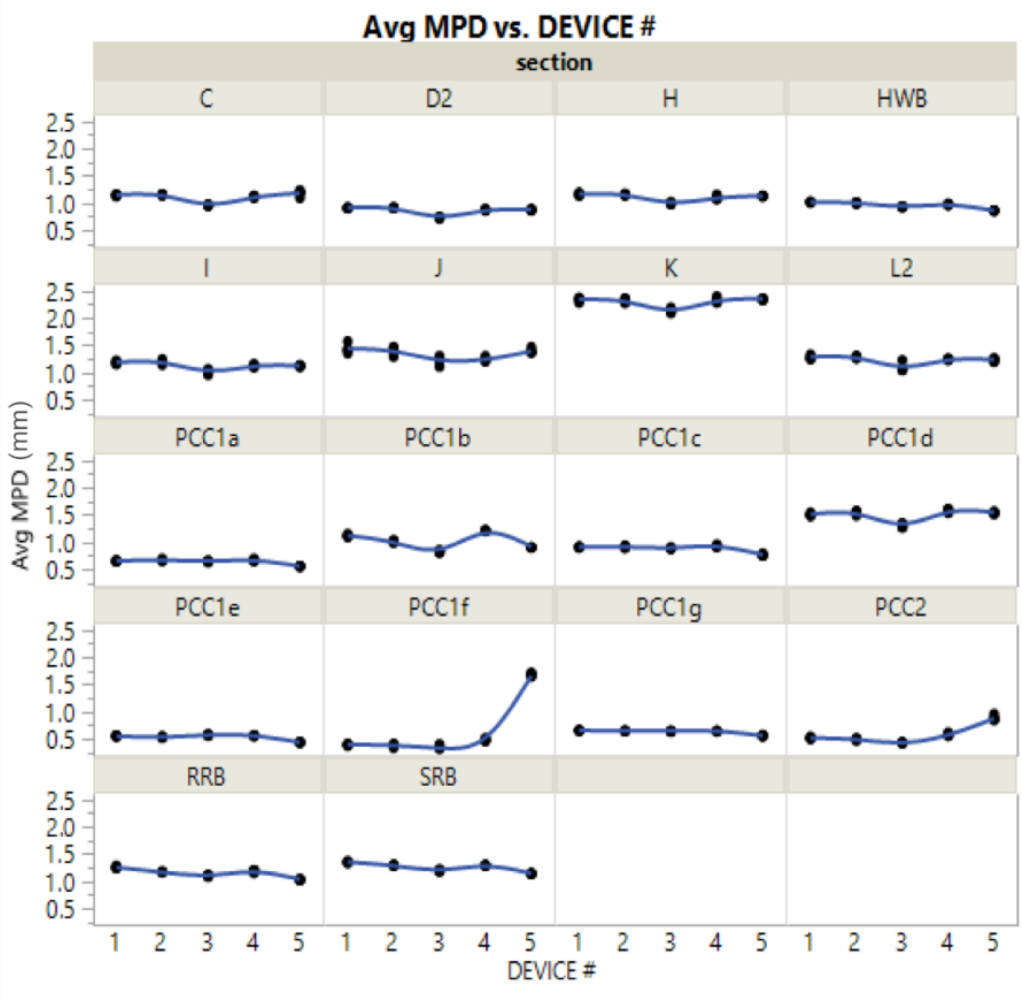
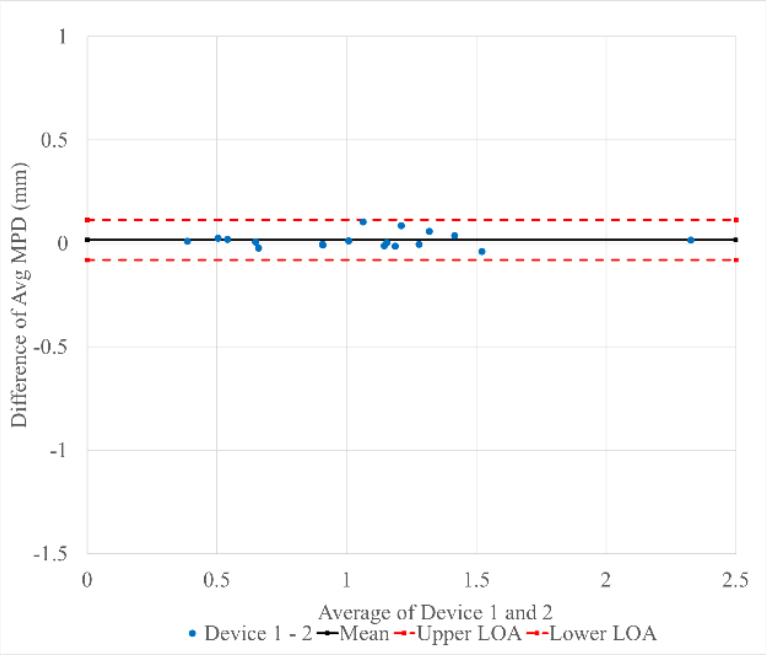


Figure 37 - Mean MPD results for each section tested.

The LOAs were initially calculated for pairs of all devices and all sections of pavement. However, the agreement of all devices with the transversely mounted line laser (Device 5) is poor compared to the pairwise comparisons of the other devices. The difference in MPD for Device 5 is seen in Figure 37; note all five runs are plotted for each device but are typically too close to distinguish at this scale. The plots in Figure 37 show that there is a difference for the longitudinally texture surfaces, and the plots in Figure 38 show how much the single-spot lasers disagree with the line laser system. Note the narrow band of agreement of two single-spot laser

devices (Figure 38a) and a wider range for the single-spot laser compared to the line laser (Figure 38b). This does not mean that the readings made by either of the devices are necessarily poor; however, it demonstrates that the devices are not interchangeable for certain pavement types.

(a)



(b)

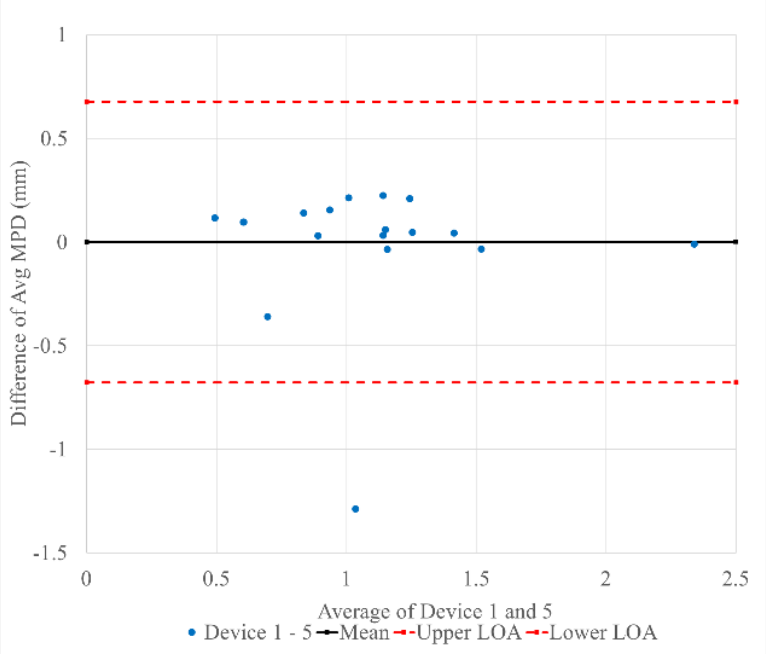


Figure 38 - Bland-Altman plots for all sections tested device pair s(a) 1-2 and (b) 1-5.

This prompted us to evaluate the pavement sections with no directional texturing (i.e., asphalt, and PCC1b and PCCC1d, epoxy-based surface treatments which are similar to asphalt), hence the four pavement groups shown in Table 8. The "Random" texture columns show better agreement for the line laser (Device 5) with all other devices when compared to the LOA for all sections. In fact, the LOA drops by an average of 74%. Since it was apparent that the engineered directional textures (i.e., grooves and tine marks) created poor agreement between the single-spot and line laser devices, LOA analysis was then carried out on the PCC sections, first those with longitudinal texture (PCC2 and PCC1f, "longitudinal only" in Table 8), next those with transverse grooves or tine marks (the remaining PCC-surfaced sections, "transverse only" in Table 8).

Table 8: Summary of LOAs

	Texture Type							
	All		Random		Longitudinal Only		Transverse only	
Device Pair	Mean of Difference	LOA	Mean of Difference	LOA	Mean of Difference	LOA	Mean of Difference	LOA
1,2	0.02	0.10	0.01	0.12	0.02	0.05	0.02	0.08
1,3	0.13	0.19	0.20	0.12	0.07	0.08	0.06	0.14
1,4	0.01	0.15	0.03	0.18	-0.07	0.07	0.02	0.08
1,5	0.00	**0.68**	0.04	0.17	-0.82	**1.29**	0.15	0.10
2,3	0.11	0.18	0.19	0.10	0.05	0.08	0.04	0.09
2,4	0.00	0.16	0.02	0.20	-0.09	0.06	0.00	0.05
2,5	-0.03	**0.67**	0.03	0.12	-0.84	**1.27**	0.13	0.06
3,4	-0.12	0.21	-0.17	0.23	-0.14	0.06	-0.04	0.08
3,5	-0.14	**0.66**	-0.16	0.14	-0.89	**1.22**	0.09	0.06
4,5	-0.03	**0.63**	0.01	0.25	-0.75	**1.24**	0.13	0.06

Note the line laser again showed poor agreement with all devices (as expected) for the "Longitudinal Only" sections (the line laser can capture this texture and the other devices cannot). However, it is interesting to note that the best agreement between all devices was for the PCC sections with transverse texturing with LOA ranging from 0.05 mm to 0.14 mm. A transverse-mounted line laser theoretically should fall either into a peak or a valley for each reading. The close agreement, however, is likely due to the fact that the line laser does not always run parallel to the transverse groove or tine and, therefore, captures peak and valley information similar to the single-spot laser systems. Table 8 shows that the two technologies (single-spot and line lasers) cannot be used interchangeably for all pavement types. It is also noted that Device 3 has a consistent bias of approximately 0.1mm (see the "Mean of Difference" column in Table 8) when compared to all other devices. This is likely due to aggressive hardware filtering of the signal before it is output as a raw pavement profile.

Operational Factors Tests

Constant Speed

To analyze the effect of speed (24 km/h to 105 km/h in 16-km/h increments) on the devices tested, an ANOVA was performed for each device on the two sections tested. A significance level of 0.05 was selected, above which we fail to reject the null hypothesis that speed does not affect the mean 1-meter MPD calculated. The results are presented in Table 9. Recall that section L2 is a dense-graded hot mix asphalt and PCC1a is a transversely tined PCC section.

Table 9: Summary of ANOVA Results for the Constant-Speed Experiment

Device	Section	*p*-value
1	L2	0.3854
2	L2	0.5966
3	L2	**0.0009**
4	L2	0.3977
5	L2	**0.0001**
1	PCC1a	0.0607
2	PCC1a	0.2206
3	PCC1a	**0.0001**
4	PCC1a	0.4951

5	PCC1a	**<0.0001**

As seen in Table 9, Devices 3 and 5 reject the null hypothesis that speed has no effect on MPD calculated. Recall that Device 3 and 5 have lower sampling frequencies: 32 kHz and 5 kHz, respectively. However, Device 5 is a line laser and captures over 300 points transversely. Figure 39 shows the average of 1-meter MPDs calculated for Device 3 for each of the test speeds. Here, a clear linear trend is seen with MPD negatively correlated to speed. The range of the data, however, is entirely different for the two pavement types tested. For Section L2 (DGHMA) the range of MPDs through the speeds was from 1.24 down to 0.99 mm for a spread of 0.25 mm. For the transverse-tined section PCC1a, the data ranged from 0.68 down to 0.61 (a spread of 0.07 mm), which is within the range of good repeatability for the device. The range of MPD for the linear regression line from 15 to 65 mph is 0.11 mm and 0.04 mm for sections L2 and PCC1a respectively. In reviewing the data for Devices 2 and 3 (both equipped with 32kHz sensors), it appears Device 3 had some hardware filtering applied before the data was manipulated with any software filtering by the authors. This demonstrated by the relatively smoother profile illustrated in Figure 40.

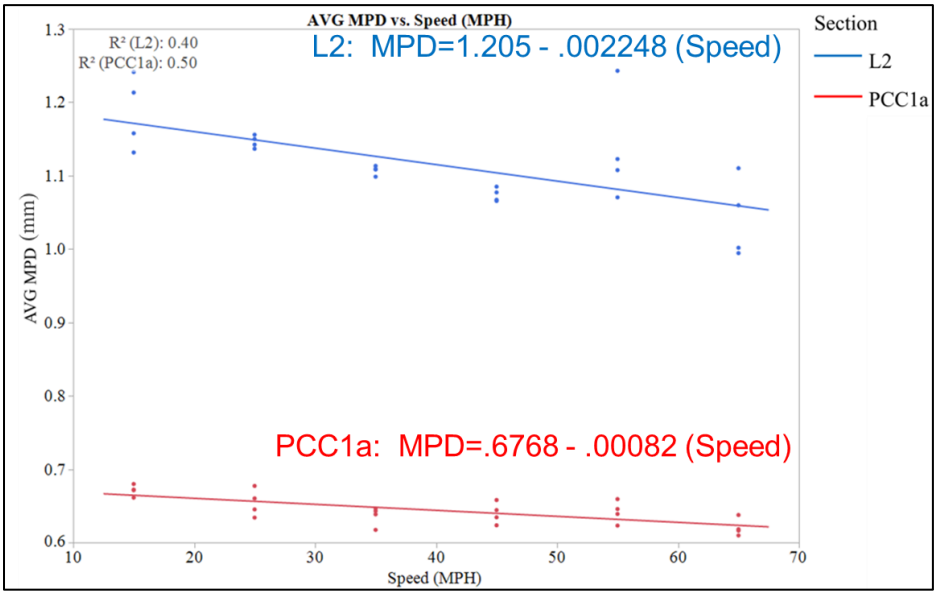


Figure 39 - MPDs for all runs of Device 3 in the constant speed experiment

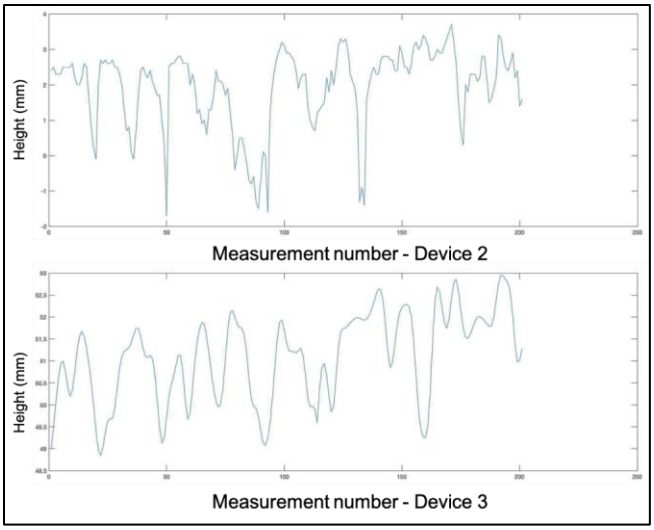


Figure 40 - Comparison of Device 2 and 3 Profiles

Device 5 also had a negative correlation as shown in Figure 41. It also shared similar range characteristics with MPDs from 1.32 down to 1.20 mm for L2 (a spread of 0.129 mm) and a spread of only 0.054 mm (MPDs from 0.60 down to 0.54 mm) for the PCC1a section. The range of MPD for the linear regression line from 15 to 65 mph is 0.08 mm and 0.03 mm for sections L2 and PCC1a respectively This indicates speed has an influence this line laser device, but the effect is not significant to the device's calculation of MPD for the transverse-tined surface but has a minor effect for the DGHMA surface. When the raw profiles were investigated, it was noted that the profiles were smoother and more closely resembled a flat line as speed increased. This could be due to an averaging effect on the signal with increased vehicle speed and constant exposure time on the device's sensor. Users can adjust the 'exposure time' or, more accurately, the amount of time photons are collected on the device's charged-couple device (CCD) before the CCD is cleared for the next reading. Devices are also capable of adjusting this exposure time to compensate for the surface being measured. Lighter surfaces reflect more light and require less exposure time. Darker surfaces reflect less light and, therefore, more exposure time allows more of these scare photons to be collected. With longer exposure time, more features of the surface scanned are collected on one single line of data, averaging peaks and valleys and forming flatter surfaces as can be observed in Figure 42.

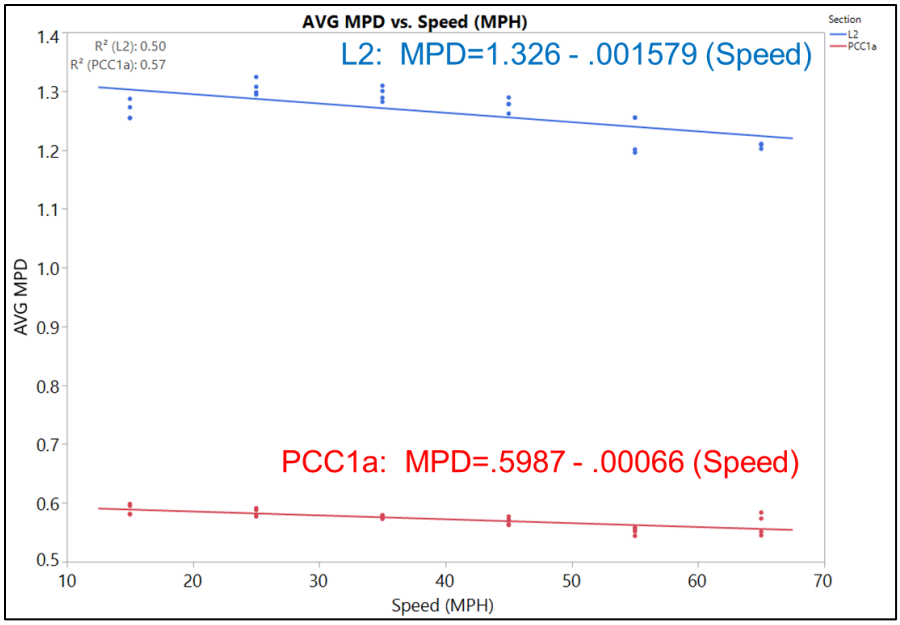


Figure 41 - MPDs for all runs of Device 5 in the constant speed experiment device 5

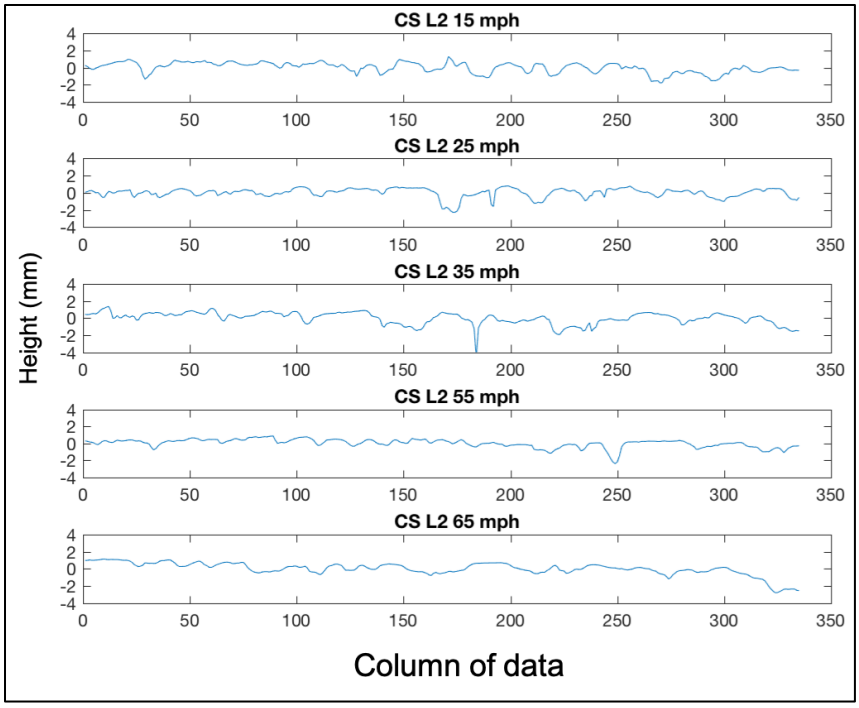


Figure 42 - Smoothing of Line Laser with Speed

Variable Speed

The effect of acceleration (both speeding up and braking) was evaluated similarly to the constant speed experiment. An ANOVA analysis with the four various acceleration conditions as the model input and MPD as the response was completed. Results are summarized in Table 10.

Table 10: Summary of ANOVA Analysis for Variable Speed Experiment

Device	*p*-Value
1	0.4653
2	0.6184
3	0.4306
4	0.8423
5	**0.0051**

Only the device with a line laser rejected the null hypothesis that acceleration does not affect the MPD calculated. Box plots of the data collected by this device are shown in Figure 43. Although the null hypothesis was rejected for Device 5, it should be noted that the data ranged from 1.23 down to 1.17 mm (a spread of only 0.06 mm), which is within the range of good repeatability for the device. Because we fail to reject the null hypothesis for the other devices, there is statistical evidence that the variable speeds tested do not affect the reading of average MPD for single-spot devices. However, the good repeatability shown for the line laser indicates that the ANOVA may be penalizing the line laser data due to its low variance. In other words, with low variability, a small deviation from the mean causes a rejection of the null hypothesis that all means are equal. Since the range of 1-meter MPDs (0.06 mm) is within the device's coefficient of repeatability, acceleration may not affect readings taken with a line laser.

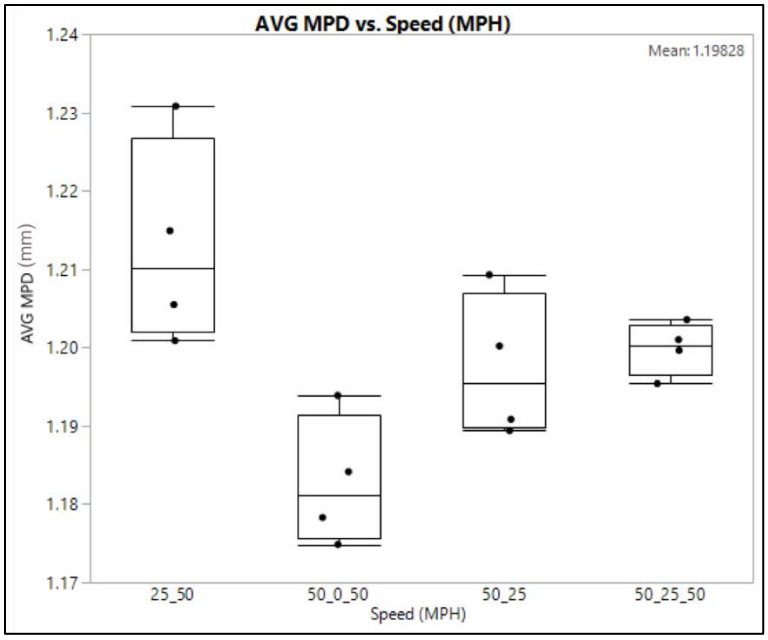


Figure 43 - Data collected by Device 5 during the variable speed experiment.

Summary and Conclusions

This study compared 5 macrotexture measurement devices to determine repeatability, reproducibility and the effect of various operational factors expected for data collection at traffic speed.

- A novel approach to removing dropouts and outliers from line laser data was presented. The approach is adaptable to the given dataset and guarantees that the rate of falsely discovered outliers is limited to a pre-determined threshold.
- All devices tested showed good repeatability for all surfaces tested with coefficients of repeatability ranging from 0.063 to 0.088 mm.
- Poor agreement was found for the line laser when compared to all other single-spot laser devices on all sections in an LOA analysis. This does not necessarily mean that the measurements taken by this type of technology are bad. The wide band of agreement is likely because the line laser measures the peaks and valleys of longitudinal texturing whereas single spot lasers cannot. All pairwise comparisons of the line laser and the other devices yielded similar results when only transverse grooved sections were considered.

- Agreement improved significantly (lack of agreement fell by more than half) when considering only surfaces similar to asphalt (i.e., SMA, OGFC, and HFST) in the LOA analysis.
- Device 3, a single-spot laser with apparent hardware filtering of the raw data signal appears to have a bias in calculated MPD of approximately 0.1mm when compared to all other devices.
- Speed is a significant factor in determining a pavement's macrotexture for device 3 which is attributed to hardware filtering. Surface type tends to either magnify or reduce this effect.
- Acceleration is not a significant factor for devices equipped with single-spot lasers. ANOVA revealed that, statistically, acceleration is a significant factor for the line-laser-equipped device. However, the range of data gathered by this device during the acceleration testing is relatively low and within the device's coefficient of repeatability, so it may not necessarily be the case that acceleration produces unacceptable results for devices equipped with line lasers.
- Surprisingly, the best agreement between all devices comes when considering only PCC surfaces with transverse texturing. This is likely due to the fact that the line laser does not always run parallel to the transverse groove or tine and, therefore, captures peak and valley information similar to the single-spot laser systems.
- Readings from single-spot and transversely mounted line lasers should not be used interchangeably on longitudinally textured surfaces such as tined, grooved, or ground PCC surfaces.

Acknowledgement

The authors would like to extend our sincere appreciation to those who performed the macrotexture measurements at the Virginia Smart Road.

Author Contributions

The authors confirm contribution to the paper as follows: study conception and design: V. Bongioanni, E. de León Izeppi, Gerardo Flintsch; data collection: V. Bongioanni, K. Maeger; analysis and interpretation of results: V. Bongioanni, K. Maeger; draft manuscript preparation:

V. Bongioanni, K. Maeger. All authors reviewed the results and approved the final version of the manuscript.

References

[1] Altman, D. G., and J. M. Bland. Measurement in medicine: the analysis of method comparison studies. *The statistician*, 1983, pp. 307-317.

[2] PIARC. Technical Committee Report On Surface Characteristics --Piarc XVIII World Road Congress, Brussels, Belgium, September 13-19, 1987.In, World Road Association, Permanent International Assoc of Road Congresses (PIARC), 1987. p. 108 p.

[3] Descornet, G. The HERMES project.In *Symposium on Pavement Surface Characteristics [of Roads and Airports], 5th, 2004, Toronto, Ontario, Canada*, 2004.

[4] Sandberg, U., A. Bergiers, J. A. Ejsmont, L. Goubert, R. Karlsson, and M. Zöller. Road surface influence on tyre/road rolling resistance. *Models for Rolling Resistance in Road Infrastructure Asset Management Systems (MIRIAM),(http://miriam-co2. net/Publications/MIRIAM_SP1_Road-Surf-Infl_Report,* Vol. 20111231, 2011.

[5] Haider, M., M. Conter, M. Green, B. Schmidt, and U. Sandberg. Status of the EU-project ROSANNE. *Transportation Research Procedia*, Vol. 14, 2016, pp. 2946-2955.

[6] Mokarem, D. W. Use of the Digital Surface Roughness Meter in Virginia.In, Virginia Transportation Research Council, 2006.

[7] Flintsch, G., E. de León Izeppi, K. McGhee, and J. Roa. Evaluation of international friction index coefficients for various devices. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2094, 2009, pp. 136-143.

[8] Fuentes, L., and M. Gunaratne. Evaluation of the Speed Constant and Its Effect on the Calibration of Friction-Measuring Devices. *Transportation Research Record: Journal of the Transportation Research Board,* Vol. 2155, 2010, pp. 134-144.

[9] Vos, E., and J. Groenendijk. Report on Analysis of Previous Skid Resistance Harmonization Research Projects. FEHRL, Brussels, Belgium, 2009.In.

[10] Izeppi, E. d. L., G. Flintsch, and K. McGhee. Limits of Agreement Method for Comparison of Pavement Friction Measurement. *Transportation Research Record: Journal of the Transportation Research Board,* Vol. 2306, 2012, pp. 188-195.

[11] Barnhart, H. X., M. J. Haber, and L. I. Lin. An overview on assessing agreement with continuous measurements. *Journal of biopharmaceutical statistics,* Vol. 17, No. 4, 2007, pp. 529-569.

[12] Martin Bland, J., and D. Altman. Statistical Methods for Assessing Agreement Between Two Methods of Clinical Measurement. *The Lancet,* Vol. 327, No. 8476, 1986, pp. 307-310.

[13] Bland, J. M., and D. G. Altman. Measuring agreement in method comparison studies. *Statistical methods in medical research,* Vol. 8, No. 2, 1999, pp. 135-160.

[14] Bland, J. M., and D. G. Altman. Agreement between methods of measurement with multiple observations per individual. *Journal of biopharmaceutical statistics,* Vol. 17, No. 4, 2007, pp. 571-582.

[15] Carstensen, B., J. Simpson, and L. C. Gurrin. Statistical models for assessing agreement in method comparison studies with replicate measurements. *The international journal of biostatistics,* Vol. 4, No. 1, 2008.

[16] Katicha, S. W., G. W. Flintsch, B. Ferne, and J. Bryce. Limits of agreement (LOA) method for comparing TSD and FWD measurements. *Journal of Pavement Engineering,* Vol. 15, 2014, p. 6.

[17] ASTM E1845. Standard Practice for Calculating Pavement Macrotexture Mean Profile Depth.In, ASTM International, 2015.

[18] ISO 13473-1. Characterization of pavement texture by use of surface profiles - Part 1: Determination of Mean Profile Depth.In *Part 1: Determination of Mean Profile Depth*, International Organization for Standardization, 1997.

[19] Katicha, S. W., D. E. Mogrovejo, G. W. Flintsch, and E. D. d. L. Izeppi. Adaptive Spike Removal Method for High-Speed Pavement Macrotexture Measurements by Controlling the False Discovery Rate. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2525, 2015, pp. pp 100–110.

[20] Bartlett, J., and C. Frost. Reliability, repeatability and reproducibility: analysis of measurement errors in continuous variables. *Ultrasound in Obstetrics and Gynecology: The Official Journal of the International Society of Ultrasound in Obstetrics and Gynecology*, Vol. 31, No. 4, 2008, pp. 466-475.

[21] Doty, R. N. Study of the sand patch and outflow meter methods of pavement surface texture measurement.In *Surface Texture Versus Skidding: Measurements, Frictional Aspects, and Safety Features of Tire-Pavement Interactions*, ASTM International, 1975.

CHAPTER 5: SUMMARY, FINDINGS, CONCLUSIONS AND RECOMMENDATIONS

This thesis sought to enhance the network level collection and deterioration modeling of pavement macrotexture for use in pavement management. Reliable collection and accurate modeling of parameters are necessary to use PMS for optimized network and project level decisions. This was accomplished through two experiments, each of which produced a manuscript.

The first experiment measured pavement macrotexture as simulated traffic was applied using an HVS. The pavement surface was a SMA with a high binder content of 6.7%, which led to bleeding as the primary distress that developed as load was applied. The HVS applied the traffic load at a normal distribution to simulate the effect of vehicle wander. Macrotexture was measured at 4 stations along the road, and each station had 5 measurement locations transverse to the direction of traffic. Measurements were taken on a weekly basis over the course of 6 months as the cumulative load increased from zero to a final load of 2.9 Million ESALs. Photographs, ANOVA and linear regression where used to analyze the data and model the effect of traffic induced bleeding on macrotexture.

The second experiment compared various high-speed macrotexture measurement devices. The devices collected data the data under three operating conditions: high-speed, constant speed and variable speed. The high-speed condition collected macrotexture at 55 mph over 18 pavement sections with difference surface types. The constant speed condition collected macrotexture of two pavement sections at 15, 25, 35, 45, 55, and 65 mph to determine if speed effects the devices. The final experimental condition required various braking and accelerating scenarios to determine how realistic traffic would have an effect on the devices. ANOVA and Limits of Agreement were used to analyze the data to determine the repeatability and reproducibility of the measurement devices.

Summary of Findings

The deterioration modeling of this thesis was limited to observing the development of bleeding and its effect on macrotexture under simulated traffic loading from an HVS. The experiment showed that in general macrotexture will decrease under traffic induced bleeding,

however there were some instances in which buildup of material on the surface led to an increase in macrotexture.

The macrotexture measurement device comparison showed that all the devices tested showed good repeatability with coefficients of repeatability ranging from 0.063 to 0.088mm. This indicates that on 95% of occasions two measurements from the same device on the same subject will differ by no more than this coefficient. Generally, the devices showed good agreement except for comparing the line laser device to single spot lasers which showed poor agreement. This was primarily due the differences in how the devices measured longitudinally grooved pavement sections. When the longitudinally grooved pavement sections were removed from analysis the agreement greatly improved.

The results also showed speed to be a significant factor in determining a pavement's macrotexture for two devices. It is likely that built-in hardware filtering is the cause. This was evident in Device 3, in which the hardware filtering also appears to have a created a significant bias in calculated MPD of approximately 0.1mm lower when compared to the other devices. Speed also had an effect on the line laser (Device 5) showing a decrease in average measured macrotexture with increasing speed of data collection. It is likely that as speed increased, the data collected by the device was spread over a larger area for each data point collected leading to a blur effect at higher speeds. This caused an averaging of texture effect that smoothed the profile as speed increased thereby decreasing the measured macrotexture. This effect did not cause the range of measurements to be outside of the device's repeatability when measuring a tined PCC surface, however it was problematic when measuring a randomly textured surface.

The various acceleration and braking scenarios showed that these "real world" operational factors are not a significant for devices equipped with single-spot lasers. While ANOVA showed that acceleration is a significant factor for the line-laser-equipped device, the range of data gathered by this device during the acceleration testing is relatively low and within the device's coefficient of repeatability, so it may not necessarily be the case that acceleration produces unacceptable results for devices equipped with line lasers, but rather this demonstrates the limitations of ANOVA for devices with extremely low variance.

A final observation is that surfaces with transverse texturing yielded the best agreement among devices, which is surprising when observing the considerable differences in measurements of longitudinal textured surfaces.

Engineering and Scientific Contribution

The rate of change for macrotexture (RMS) in the wheelpath due to traffic induced bleeding was modeled as an effect of cumulative ESALs and is demonstrated by the equation:

$$RMS = 1.186 - 1.422 \times 10^{-7} \times ESALs \quad (14)$$

indicating that the RMS macrotexture decreases 0.142 mm per million ESALs. This enables an approximate prediction of when a pavement may reach a predetermined macrotexture trigger value for a maintenance when traffic induced bleeding is the primary distress that develops. Additionally, the practicality of measuring macrotexture in the center of the wheel path was demonstrated by showing that data collected outside of the wheel path was subject to much higher variability than measurements in the center of the wheel path. The experimental results also lead to the conclusion that RMS is a more sensitive than MPD to macrotexture changes due to bleeding because it is impacted less by artificial peaks created by bleeding in the macrotexture profile. It is possible that these peaks are related to the high loads and channelized traffic generated by the HVS and may not model real world bleeding phenomena, however more research is needed to confirm this. It is possible this observed behavior could be applicable with advent of autonomous trucking where highly channelized, high-load traffic will become more common.

The reproducibility and repeatability of five macrotexture measurement devices was evaluated. Each individual device showed low repeatability coefficients indicating they can be used for collection of network level macrotexture data for a PMS database with satisfactory repeatability. The reproducibility or interchangeability was acceptable for the devices tested with the exception of the line laser compared to single-spot lasers. The difference is due to the limitation of single spot lasers to accurately capture both peaks and troughs of longitudinal grooves, which is easily achieved with a transverse line laser. This finding indicates that single-spot and transversely mounted line lasers should not be used interchangeably on longitudinally textured surfaces for collecting network level macrotexture. Additionally, the effect of

operational factors such as speed, acceleration and braking on these devices was assessed. The results showed that speed had an effect on some device's measure of macrotexture with the effect attributed to hardware filtering and laser exposure time for the line laser device. This highlights the need to ensure line laser devices are properly calibrated for the surface type being measured and the operating speed to minimize the effect of exposure time blur. Finally, the results showed acceleration and braking did not have a substantial effect any of the device's measurement of macrotexture.

Conclusions

Overall, more research would benefit the integration of macrotexture into pavement management in the US. As stated previously, PMS relies on condition measurement and condition prediction to optimize decisions, however, the variety of surface types and distresses make the modeling of macrotexture changes under traffic difficult. This however should not preclude macrotexture from being utilized in PMS. Good data collection practices can yield accurate and repeatable results, as demonstrated by the equipment comparison experiment. As macrotexture remains an indicator of friction and splash and spray potential, this data can still inform pavement managers of the condition of the network and be used as a trigger for maintenance, particularly for surface treatments or other pavement preservation treatments.

Recommendations for Future Research

Expanding upon the work presented in Chapter 3, future experimental design should consider testing various binder contents ranging from below optimum to above optimum and observing the different severities of bleeding and how they affect macrotexture. Additionally, correlating the measured macrotexture with friction would enhance understanding of the impact of traffic induced bleeding on safety. Furthermore, utilizing a larger sample size than the single test lane used in this study would help identify sources of variability. Finally, other macrotexture parameters in addition to those presented in this research could be assessed and correlated with friction measurements. For example, assessing the skewness of profiles may indicate if there is a significance of positive vs. negative texture that is created by the buildup of material. Assessing additional parameters would enable a broader scope of recommendations that could be made for evaluating bleeding.

To expand upon the work described in Chapter 4, future experiments would benefit from comparing multiple line lasers among each other to determine interchangeability. The effect of the line lasers' orientation, (transverse, longitudinal, and 45 degrees from transverse) should be observed. Future experiments could also further explore the discovered issues of hardware filtering and the blur effect of line lasers. This could be accomplished by adjusting calibration for individual devices and comparing the results to find optimal settings to provide the best repeatability and agreement with other devices when operating at various speeds over different surface types. This study used raw spatial data that was similarly processed for each device, however, it is likely that agencies would use automated outputs provided by manufacturer software paired with devices as a matter of convenience. It would be advantageous to make comparisons of data provided by each individual device's automated macrotexture parameter outputs to determine how repeatable and reproducible they are and if they are different than the results presented in this research.