METHODS FOR AUTOMATED IDENTIFICATION OF ROADWAY DRAINAGE-RELATED FEATURES FROM MOBILE LIDAR DATA

A Thesis

by

SAURAV RAJ NEUPANE

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Nasir G. Gharaibeh
Committee Members,	Francisco Olivera
	Fouad Jaber
Head of Department,	Robin Autenrieth

December 2017

Major Subject: Civil Engineering

ABSTRACT

Light Detection and Ranging (LiDAR) systems have been increasingly used in project planning, project development, construction, operations, maintenance, and asset management. Typical data collected by a LiDAR system include slant distance, incidence angle, and reflectivity measurements. This research focuses on mobile LiDAR systems (MLSs).

Processing of large amounts of data collected by MLSs remains tedious and time-consuming. For MLSs to be used efficiently in roadway drainage inventory and condition assessment, automated methods are needed to identify key features that affect drainage. The aim of this research is to develop computational methods for automated identification of such features from data collected through MLSs. The specific objectives of this research are to a) detect pavement surface type, b) detect the presence of driveways and underlying pipes and extract count, width, elevation difference and material cover and c) detect roadside features such as grass-cover area, curb location, and curb height based on the data collected using a SICK LMS-5XX series LiDAR scanner and hardware and software by Road Doctor.

Reflectivity, measured as a logarithmic index of power level called received signal strength indicator (RSSI), is used to develop an algorithm to detect surface type based on statistical analysis of RSSI. Cross-sectional geometry, along with material identification, is used to identify driveways and underlying pipes. RSSI distribution and material identification techniques are used to detect roadside grass areas. Elevation distribution and

filter template technique are used to detect curbs. Each method was tested and validated using data from actual road sections in Texas. The ability to detect aforementioned features reliably using automated means is an initial step to further the cause of MLS acceptance and implementation.

Generally, the accuracies of pavement and grass detection methods were at least 83%. The effect of reflectivity attenuation is pronounced for roadside. Therefore, in order to develop a reliable grass detection method, attenuation correction is required. It is possible to detect driveways and distinguish them from topographical features using a combination of elevation cross sections, material detection, and surface smoothness. It is possible to identify curbs using filter template technique.

DEDICATION

To my late father – Daivagya Raj Neupane.

ACKNOWLEDGEMENTS

I would like to express my gratitude to my committee chair, Dr. Gharaibeh, and my committee members Dr. Olivera and Dr. Jaber, for their guidance and support throughout the course of this research. Dr. Gharaibeh provided invaluable insights into condition assessment, asset inventory and provided a direction for this research, guiding me along the way. This research would not have been a reality without his timely counsel and patience. I would like to thank Charles Gurganus for providing me with his invaluable time, support, guidance and encouragement.

I would like to thank my friends and colleagues and the department faculty and staff for making my time at Texas A&M University a great experience. Thanks also go to all my teachers and friends throughout my life who have contributed, directly and indirectly, in shaping my knowledge and views so that I am able to undertake this study. Many thanks to my mother and my wife for their encouragement, patience and love. Finally, I would like to thank all my family members for their belief and endless support.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a thesis committee consisting of Dr. Nasir Gharaibeh (Chair) and Dr. Fransisco Olivera (Member) of the Department of Civil Engineering (Member), and Dr. Fouad Jaber (Member) of the Department of Biological and Agricultural Engineering. The data used for this study was collected and provided by Charles Gurganus.

All work for the thesis was completed independently by the student.

Funding Sources

Graduate study was supported in part by funding from the Texas Department of Transportation.

NOMENCLATURE

DMI	Displacement Measuring Interferometer
EM	Electromagnetic
FM	Farm to Market (Roads)
GPS	Global Positioning System
IMU	Inertial Measurement Unit
MLS	Mobile LiDAR System
NIR	Near-Infrared
LiDAR	Light Detecting and Ranging
RGB	Red, Green and Blue
RSSI	Received Signal Strength Indicator
SVM	Support Vector Machines
TTI	Texas Transportation Institute
TxDOT	Texas Department of Transportation

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

1 INTRODUCTION

1.1 Background

Light Detection and Ranging (LiDAR) systems have been increasingly used in project planning, project development, construction, operations, maintenance, and asset management. Typical, data collected by a LiDAR system include incidence angle, slant distance, and reflectivity intensity. Intensity measurement vary from one manufacture to another (Kashani, et al., 2015). Generally, the reflectivity value is measured as a logarithmic index of power level called Received Signal Strength Indicator (RSSI). Intensity is normalized between 0 and 1 or scaled to 8, 12 or 16 bit dynamic range (Kashani, et al., 2015). Some units record accelerometer and displacement data as well. Generally, LiDAR systems operate on fixed platforms (terrestrial LiDAR), aerial platforms, satellite platforms, or mobile platforms. This research focuses on mobile LiDAR systems (MLSs). In MLSs, a video camera mounted on the vehicle provides additional imagery information. In airborne system, the height of flight is significantly greater than the changes in ground elevation; hence higher accuracy can be achieved by a MLS (Williams, et al., 2013; Large, et al., 2009).

There are numerous ongoing research efforts to find useful and creative applications of LiDAR systems in civil engineering, such as project planning and development for creating CAD models and visualization, topographic mapping, construction automation (e.g., machine guidance), as-built surveys, post construction quality control, autonomous navigation, landslide assessment, monitoring of coastal

changes. In the infrastructure management area (which is the focus of this research), MLSs are increasingly being used for automated collection of roadway condition and inventory data. These applications are motivated primarily by improving safety and increasing efficiency in field data collection. Data collection can be carried out at traffic speed without any obstruction to traffic flow. Usage of LiDAR data in infrastructure asset management include inventory, mapping, condition inspection, and automated or semi-automated extraction of features. (Williams, et al., 2013)

1.2 Problem Statement and Research Questions

While MLSs collect large amounts of data (e.g., elevation and reflectivity measurements), the processing of these data remains tedious and time-consuming. Specifically, for MLSs to be used efficiently in roadway drainage inventory and condition assessment, automated methods are needed to identify key features such as pavement surface type, the presence of driveways and underlying pipes, presence of sidewalks, and type of roadside cover.

Pavement surface type and condition affect surface drainage and flow characteristics (e.g., manning coefficient). Driveways and underlying pipes, sidewalks, curbs, and roadside cover (e.g., dirt, gravel, and grass) are key roadway drainage features. The identification of these features will feed into drainage condition assessments and asset management systems. At the same time, the ability to detect the presence of these features using automated means will improve the application of MLSs in roadway inventory and drainage condition assessment.

1.3 Research Objectives

The aim of this research is to develop methods for automated identification of roadway drainage related features from data collected through MLSs. The specific objectives of this research are to:

1. Detect pavement surface type based on statistical analysis of reflectivity intensities.
2. Detect the presence of driveways and underlying pipes and extract relevant parameters, such as count, width, elevation difference and material cover based on reflectivity intensities and road geometry.
3. Detect roadside features including grass, dirt-cover, or side-walk, and detect grass-cover area, curb height and drainage features based on reflectivity and elevation distribution.

Since MLSs record surface reflectivity intensity values and geometric measurements (e.g., elevation), this study hypothesizes that it is possible to identify pavement surface types, driveways and underlying pipes, and roadside material cover using MLS-collected data, without the use of other specialized equipment or manual measurements.

1.4 Data Collection Equipment

The data analyzed in this study were collected using a Mobile LiDAR System, owned by the Texas Transportation Institute (TTI) and purchased from RoadScanner Oy of Finland. This equipment includes a planar SICK LMS-5XX series LiDAR scanner, Road Doctor CamLink camera, GPS, IMU, Road Doctor Camlink 7.0 in-vehicle software and Road Doctor 3 post-processing software. LiDAR scanner is mounted at the back of a

truck and a forward-facing video camera is mounted on top of the cab. The scanner operates on a multi-echo, pulse time method. The pulse time method calculates the distance between the sensor and an object by capturing the time interval between the last laser pulse leaving the sensor and being received back. The scanner has a field of view of 190^{o} and a range of up to 180m (590 ft) and operates on 100 Hz frequency. It sends 285 shots every second and records one data every 0.667^{o} increment in angle of incidence. Resolution and range of data collected is dependent on elevation of scanner and vehicle speed in addition to scanner frequency.

1.5 Research Task and Thesis Organization

The thesis consists of six main sections, as follows:

Section 1: Introduction and general background

This section provides a general background of the research topic, the problem to be studied throughout the research, and the research objectives.

Section 2: Literature review

This section provides a review of the literature on various relevant topics. It contains discussions of automated data processing techniques, LiDAR technology, and application of mobile LiDAR in infrastructure management.

Section 3: Development of surface type identification method

This section deals with the formulation and implementation of a pavement surface type and grass identification methods for data gathered through MLSs. Important statistical parameters allowing successful classification of these features are identified. Detection technique is developed and fine-tuned. Effect of data size on identification accuracy is studied. Results of pavement surface type identification tests on actual roadway sections is presented. Variation in accuracy of grass identification for different pavement surface type is presented.

Section 4: Development of driveway and underlying pipe identification method

In this section of the thesis, parameters for the detection of driveways and underlying pipes from MLSs data is investigated, followed by the development and testing of the identification method. Cross-section geometry together with surface type identification is used to formulate the identification method. Finally, results of detection tests on actual roadways are presented.

Section 5: Development of roadside features identification method

This section contains discussion of the development and testing of roadside features identification algorithm, followed by results of tests carried out. These roadside features include presence, offset and height of curb and the area and condition of roadside grass.

Section 6: Summary of research efforts, conclusions, and recommendations

This section provides a summary of the overall research efforts, the conclusions of the study, and recommendations for future studies.

2 LITERATURE REVIEW

2.1 LiDAR Technology

Light Detection And Ranging System (LiDAR) uses shorter coherent and monochromatic wavelength of electromagnetic (EM) spectrum. Laser ranging was developed in the 1960s. Early systems used 'single beam' profiling devices and were used for bathymetry. With the development of more accurate geo-referencing, LiDAR terrain mapping began in 1970s. Development of global positioning systems and inertial measurement systems improved accuracy of LiDAR systems for various applications.

Components of a typical Mobile LiDAR system (MLS) include camera, laser scanner, Global Positioning System (GPS), Inertial Measurement Unit (IMU), data storage and management systems. Complex MLS use multiple GPS receivers, an IMU and a Displacement Measuring Interferometer (DMI) for improved positioning. IMU measures body's specific force, angular rate using a combination of accelerometers and gyroscopes. The GPS/IMU system works together continually to report the best possible position. When satellite coverage is poor, the IMU fills the gap and corrects with GPS observation. Typical GPS receivers report positioning information at the rate of 1 - 10 Hz. IMU typically records positional information at 100-200 Hz. This improves the accuracy of LiDAR point cloud at higher speed. Yoo, et al. (2006) showed that scanner orientation on mobile platform can have drastic effects on the quality of data captured.

Mirror in the scanner spins to project laser pulses to the surface and measures the angle at which each pulse was fired. Scanner also receives reflected pulse from surface. This information is supplemented by location information for the scanner at the time of measurement, thus enabling calculation of coordinates of each scanned object. Typically scanners operate in line scan (or planar) mode. Generally, reflectivity value is measured as a logarithmic index of power level called Received Signal Strength Indicator (RSSI). Intensity measurement vary from one manufacture to another. Intensity is normalized between 0 and 1 or scaled to 8, 12 or 16 bit dynamic range (Kashani, et al., 2015).

Frequency of emitted signal typically lies in the range of 50 kHz to 200 kHz (Large, et al., 2009). Signal with various wavelength is used based on usage such as meteorology, terrestrial mapping and bathymetry. Time delay between transmitted and reflected signals and intensity of reflected signals are measured. Using constant speed of light, delay is converted to slant distance. Knowing positon and orientation of sensor, 3D coordinates of the reflective object is calculated.

Scanning modes are classified into two groups - phase scanning and pulse scanning. Phase scanning uses continuous wave laser scanners, where laser is constantly emitting light. Typically, this type has higher resolution and low range and is capable of measuring Doppler shifts. As such, it is more suitable for high velocity measurement. Data collection rate is also higher; in the range of 250-500 kHz (Large, et al., 2009). Pulse scanning devices emit single pulse or train of pulses. This type of system generally has long range but low resolution and cannot account for Doppler effects. Typical data collection rate is 100 - 10,000 Hz (Large, et al., 2009).

Intensity information from Mobile LiDAR system has been used in data registration, feature extraction, classification, damage detection, surface analysis and segmentation. A major application of LIDAR intensity that has been widely studied is to classify natural and urban surface covers such as asphalt roads, grass, trees and house roof (Kashani, et al., 2015). Intensity has also been used to discriminate snow covered areas from bare ice in a glacier, aging lava flows, rock properties, coastal land cover, flood modeling and wetland hydrology (Kashani, et al., 2015). Furthermore, LiDAR intensity has been used in conjunction with other measurements to improve the accuracy of results.

LiDAR intensity is influenced by surface reflectance, roughness and other surface characteristics. In addition there are confounding variables related to intensity measurement such as range, angle of incidence, transmittal power, atmospheric transmittance, scanning environment and sensors. To minimize the effect of these variables and produce values that are more closely related to true surface characteristics, several intensity processing techniques have been developed and implemented. These processes have been classified as intensity correction, intensity normalization and rigorous radiometric correction and calibration. Intensity correction refers to adjustment made to the intensity values to reduce variation caused by range, angle of incidence etc. Intensity normalization processes increase contrast to facilitate distinction. Radiometric correction is a more detailed process based on calibration and mechanistic models. This process is most effective in improving consistency in collected data. For example, using these techniques, studies have shown an improvement of 9% to 31% in LiDAR-based canopy classification results. (Kashani, et al., 2015).

2.2 Surface Type Detection Techniques

Gavilán et al. (2011) used a vehicle equipped with line scan cameras, laser illumination and acquisition to collect digital images. That study proposed a linear multi-class support vector machines (SVM)-based classifier able to distinguish between up to 10 different types of pavement (7 bituminous and 3 concrete). SVM are classifiers based on the concept of decision planes used to distinguish data points. Multi-class problem is reduced into multiple binary classification problems and optimization is carried out to find out best hyper-plane for overall classification. Gavilán et al. (2011) suggested that the use of pre-processing steps involving detection of non-crack features reduces the impact of false positives in identifying pavement types.

Omer and Fu (2010) investigated the feasibility of winter road surface classification using low-cost RGB camera and a trained Support Vector Machine (SVM). Classification groups selected in the study are bare, wheel track bare, and fully snow covered. Like other imagery based methods, change of light intensity, shadows and noise are found to be main challenges in developing a robust classification system.

Mohammadi (2012) developed a method for classification of road surface materials using hyperspectral data. Mean and standard deviation are found to be suitable spectral functions for distinguishing between asphalt, concrete and gravel (See Figure 1). Good, intermediate, and bad asphalt condition are differentiated using mean and image ratio. Usefulness of reliable reference spectra in classification of spectrally similar road surface material is identified. The multidimensional data help to reliably identify various materials under consideration due to differences in absorption and reflectivity for different

spectral bands. That study provides an insight into variation in reflectance for asphalt, concrete, and gravel against different incident wavelengths. Herold and Roberts (2005); Noronha et al. (2002) have shown that it is possible to map road surface condition and distress using hyperspectral imagery.

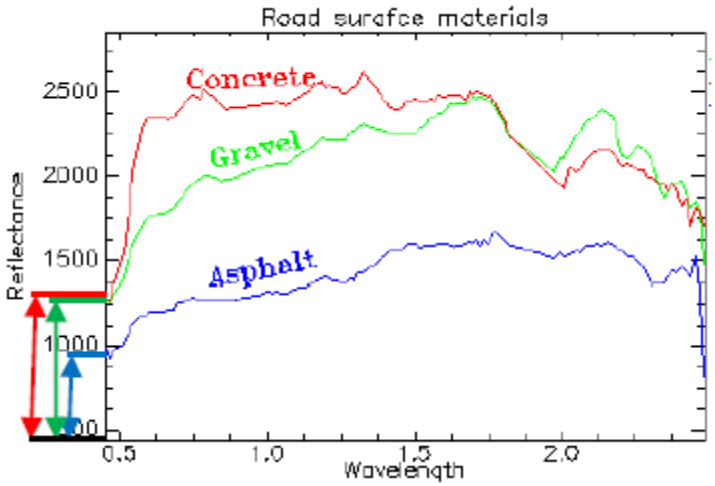


Figure 1: Asphalt, concrete and gravel spectra (Reprinted from Mohammadi, 2012)

Jonsson et al. (2015) used near infrared (NIR) camera images to develop a method that can reliably distinguish between dry, wet, icy or snow covered road surface areas. The study experimented with and compares various classification algorithms, namely K-Nearest Neighbour (KNN), Neural Networks (NN), Support Vector Machines (SVM), Discriminant Analysis (DA), and Partial Least Squares (PLS). Different spectral response is obtained from laboratory and field tests for different surface types. Along with successful use of NIR camera system for surface status detection, that study identified SVM algorithm to be the most accurate type of classification.

2.3 Methods for Feature Extraction from Images

Classification of features based on automated field measurement is an active area of research. For such classification, statistical calculations like mean, deviation, correlation, kurtosis, skewness, distribution etc. can be used solely or in association with numerical techniques like pattern recognition (e.g., classification, clustering), Neural Networks, K-Nearest Neighbor, Support Vector Machines, Discriminant Analysis and Partial Least Squares. The following paragraphs review the application of these techniques for extracting pavement features from camera-based and LiDAR-based images. Principle benefit of using these automated means for collecting pavement condition data are safe data collection at traffic speeds and increasingly consistent and reliable results (Ong, et al., 2010; Timm & McQueen, 2004).

Pavement types differ in granulation size, coloration and distribution. Previous classification attempts included implementation of multi-class support vector machines (SVM) on gray-scale images of road surface using four statics (average, deviation, skewness and kurtosis) and Fast Fourier Transform, hyperspectral imagery using spectral angle mapper approach and spectral functions such as mean, standard deviation and image ratio, and use of support vector machines (SVM) on RGB image using Gaussian filter (Mohammadi, 2012; Omer & Fu, 2010; Jonsson, et al., 2015; Noronha, et al., 2002).

Teomete, et al. (2005) studied digital image processing for pavement distress analysis. Pixel wise sum of 8-bit grayscale image was used without the use of any filtering algorithms to identify cracks with different orientations. Rajab, et al. (2008) estimated areas of pothole, alligator cracking and rutting based on image measurements. The result

is compared with traditional methods and showed close agreement. Mustaffara, et al. (2008) worked for development of a digital image processing and photogrammetry program able to classify longitudinal, transverse and alligator cracking and record intensity. The results obtained is found to be in 90% agreement with visual manual methods. Ahmed et al. (2011) used close range photogrammetric techniques for automatic pavement distress surveying with an aim of developing a low-cost solution. Kaseko and Ritchie (1993); Bray et al. (2006) presented integration of artificial neural network models with conventional image processing techniques and demonstrated its potential to further study in this area. Bray et al. (2006) classified pavement surface into cracks and non-cracks and suggested that their results were promising to continue research in this field. Nguyen et al. (2009) introduced a method to detect cracks along with joint and bridged gap from collected imagery using anisotropy measure. This method detects cracks based on its color (relative darkness), continuity, and dominant orientation.

Images collected by automated means have non-uniform background due to varying lighting condition, wetness, dirt, shadows and obstacles. This will cause difficulty for pavement image segmentation and pavement distress identification by use of imagery. A non-uniform background removal algorithm based on multi-scale wavelet transform presented in Sun and Qjan (2016) which is effective in removing non-uniform background and has an advantage for the extraction of tiny cracks compared to median filter algorithm and morphological closing algorithm.

Yu et al. (2014) used intensity information in point clouds generated by Mobile LiDAR system to identify and recreate 3D crack skeleton. Crack candidates are extracted

by applying Otsu thresholding algorithm, then a spatial density filter is used to remove outliers. Crack points are grouped into crack-lines using Euclidean distance clustering and finally crack skeletons are extracted based on medial skeleton extraction method. The method is found to be promising with high density point clouds. Guan et al. (2015) used curb-based road extraction, georeferenced feature image generation and iterative tensor voting-based crack extraction from high-density point clouds collected by a mobile laser scanning system. Use of iterative tensor voting - a continuous grouping method - is found to be more powerful for low contrast georeferenced feature images containing cracks with non-uniform intensity and low signal-to-noise ratio. Requirement of intensive computation is a limitation of this method.

Gavilán et al. (2011) smoothed the texture and enhanced linear feature in the pre-processing phase of LiDAR data followed by non-crack feature detection to remove the areas of images with joints, sealed cracks, and white painting. For this task, Gavilán et al. (2011) proposed a seed-based approach with Multiple Directional Non-Minimum Suppression (MDNMS) with symmetry checks. MDNMS method defines linear feature as a sequence of points where image has a minimum in the direction of largest variance, gradient or surface curvature after performing directional pixel search.

According to Moussa and Hussain (2011), different types of distresses, complex texture and color of the pavement surface are some of the challenges in developing a reliable and accurate automated system for detection and evaluation of pavement distresses. To overcome the limitation of image-based automated systems, Moussa and Hussain (2011) presented an automated pavement assessment system based on image

processing and machine learning. Their method consists of four main stages - segmentation, feature extraction, classification and parameters quantification. In that study, Support Vector Machine (SVM) is used for classification. Crack length and crack width are computed in the quantification stage, followed by crack type identification.

Lewis (1995) developed a fast template matching methodology based on normalized cross correlation method and image processing techniques to efficiently identify patterns in images. That study was later expanded into fast-normalized cross-correlation method by the author. Correlation-based methods have been used for object recognition, face detection, and motion analysis, etc. Studies like Tsai and Lin (2003) have proposed faster normalized cross-correlation methods for applications such a defect detection. Recently, Shen and Bao (2014) proposed a normalized cross-correlation method with invariant feature transform to develop a more efficient algorithm for application in remote sensing images. These developments have been paving the way for wider applications for such filter template techniques.

2.4 Application of Laser-based Data in Roadway Infrastructure Management

In roadway infrastructure management, laser-based measurements are most popular for calculation of the International roughness index (IRI), rut-depth measurements followed by joint-fault measurements (Timm & McQueen, 2004).

Classification of road surface type is an essential step for developing efficient automated methods for roadway condition and inventory assessment. Different road surface types have different effect on surface drainage. Pavement surface characteristics

such as texture, granulation size, coloration, and porosity affect laser reflectivity; and thus lead to different RSSI values.

Hans et al., (2003) qualitatively analyzed LiDAR-based elevation data for highway drainage analysis by comparing against standard USGS-based elevation data for watershed and drainage pattern delineation along a section of highway Iowa 1. The study used flow-modeling tools from Hydrologic Engineering Center (HEC) and GIS in conjunction with terrain obtained from LiDAR data and USGS Digital Elevation Model (DEM). The study did not find significant benefit due to additional detail from aerial LiDAR data in terms of highway hydrology in the area studied. Use of Mobile LiDAR for highway corridors significantly improves level of details captured, and as such can capture important drainage related features and inventory.

2.5 Data Quality Control

Quality control and assurance is carried out in most states by means of computer diagnostics, visual verification of video images and inclusive field checks for diverse road types. To promote and standardize quality control in automated pavement condition data collection, Ong et al. (2010) investigated the inherent variability of automated pavement roughness and pavement surface distress data collection processes. A set of guidelines is proposed for pre-project, data collection, and post-processing phases based on accuracy and reliability of data collection processes studied. Chief findings include necessity of equipment vendor to test and certify for accuracy and precision before data collection as well as perform back-end checks for completeness and accuracy during the post-

processing phase. Regular tests for consistency and quality is recommended for data collection phase (Ong, et al., 2010).

This thesis seeks to extend the data and image analysis and processing methods discussed earlier to data gathered through LiDAR. Specifically, the thesis focuses on method for processing and analyzing reflectivity intensity and elevation data.

3 DEVELOPMENT OF SURFACE TYPE IDENTIFICATION METHOD

The MLS unit was used for collecting data on 38.1 miles of road from different parts of Texas with concrete, dense graded, open graded, and seal coated surfaces. These sections are summarized in Table 1. The number of reflectivity readings represents the number of data points obtained within about 3 ft on each side of the MLS centerline and throughout the length of the roadway section. The locations of these sections are shown on the map presented in Figure 2. This data was used for developing and testing the pavement surface type detection method. Data was collected for additional sections within the City of College Station for the sole purpose of testing the developed method.

Table 1: Road sections considered in development of pavement surface type identification method.

Road Name	**Pavement Type**	**Length**	**Filtered No. of Reflectivity Readings**
George Bush Drive[1]	Concrete	4.5 miles (7.3 kms)	1,091,721
Penberthy Road[1]	Concrete	2.7 miles (4.3 kms)	643,068
University Drive[1]	Dense Graded	4.4 miles (7 kms)	868,329
Texas 6[1]	Open Graded	11.2 miles (18 kms)	1,702,827
FM95[4]	Seal Coated	7.7 miles (12.5 kms)	1,483,840
FM320[2]	Seal Coated	7.3 miles (11.7 kms)	1,388,874
FM2661[3]	Seal Coated	0.2 miles (0.35 kms)	47,483

Figure 2: Location of road sections considered in development of pavement surface type identification method (numbers correspond to sites in Table 1)

3.1 Pre- and Post-Processing of Reflectivity Data

Road Doctor post-processing tool was used for extraction, filtering, and formatting of data to be used in the analysis. In addition, Road Doctor was found to be a useful tool for visualization of reflectivity, elevation, and video data, facilitating development of the detection algorithm. Firstly, measurement error and discrepancies related to section lengths and synchronization was corrected using a built-in tool within Road Doctor. Next, bits of data containing noise (especially at the start and end locations of each data collection section) were identified and noted for removal. Additionally, data collected while the vehicle is slowing down or standing still were removed to impart uniformity in resolution and reduce error. Finally, data is extracted into ASCII files for further analysis using tools like Python and MATLAB.

The reflectivity data in the ASCII files revealed differences in distribution of measured reflectivity between driven lane and side lane. For driven lane, distribution has smaller deviation and consequently smaller band-width (Figure 3).

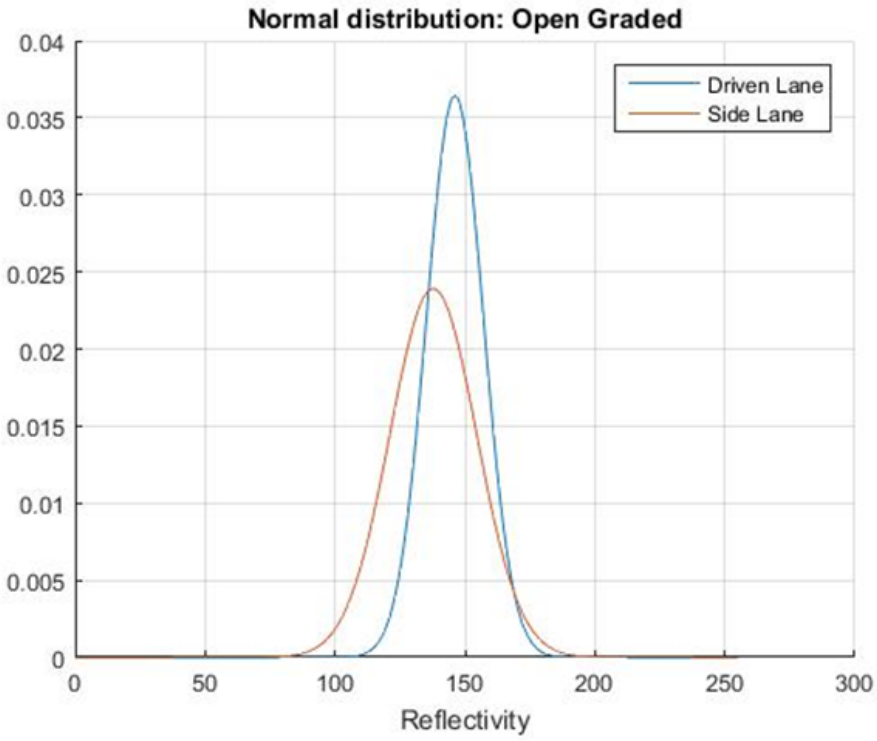


Figure 3: Distribution of reflectivity in driven lane and side lane

Normalized reflectivity distribution for driven lane is compared with distribution for whole road (Appendix A). Much higher spread (deviation) is observed in case of the later. This is due to distortion by white stripes and attenuation of intensity with angle of incidence, range, and environmental factor; as discussed by Kashani, et al. (2015).

Only reflectivity data from the driven lane were used in model development due to a) concise reflectivity distribution in driven lane, b) since the angle of incidence is

nearly vertical, there is no need for correcting reflectivity intensity, and c) driven lane does not require filtering to remove the effect of passing vehicles.

Difference in reflectivity distribution between tined and un-tined concrete was found to be minimal. Thus, these two types were combined in the analysis of concrete pavement sections.

3.2 Development of Pavement Surface Type Detection Method

3.2.1 Formulation

This method identifies pavement surface type in two ways:

- Asphalt vs. concrete surfaces (Figure 4)
- Dense graded, open graded, seal coated and concrete surfaces (Figure 5)

For both reflectivity distributions (Figure 4 and 5), the abscissa represents 8-bit RSSI values, ranging from 0-255 as measured by Mobile LiDAR System used in this study. The following can be observed from these distributions:

- Reflectivity distribution for open graded and dense graded asphalt are very similar.
- Reflectivity distribution for concrete section is more uniform with less kurtosis.
- Seal coated sections exhibit reflectivity in between concrete and open graded surface.
- Compared to concrete surfaces, asphalt surfaces exhibit higher variation.

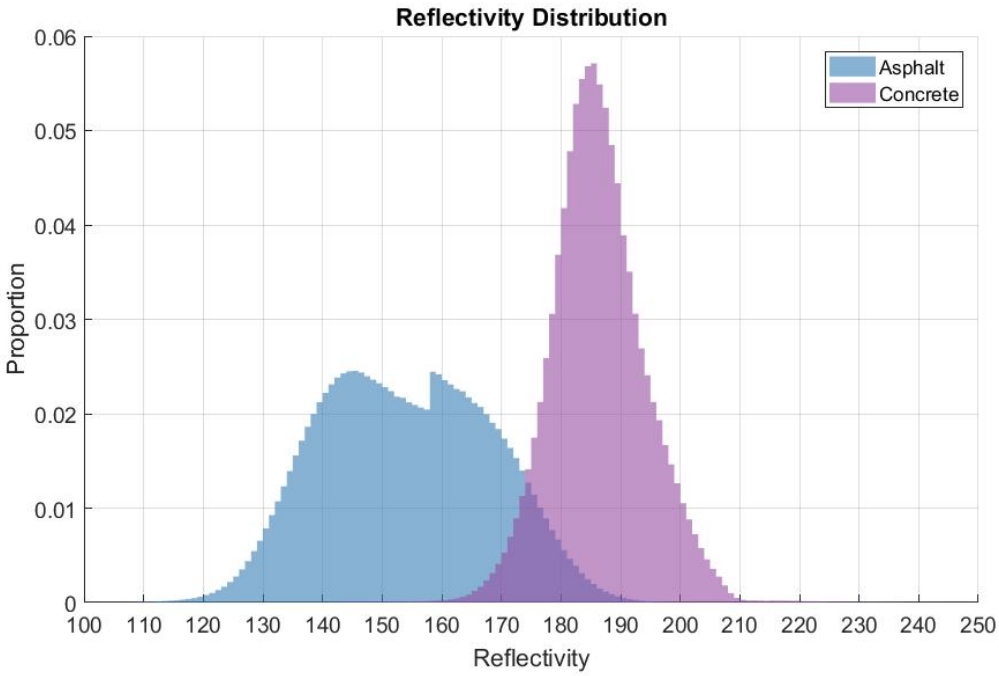


Figure 4: Distribution of reflectivity intensities for asphalt and concrete

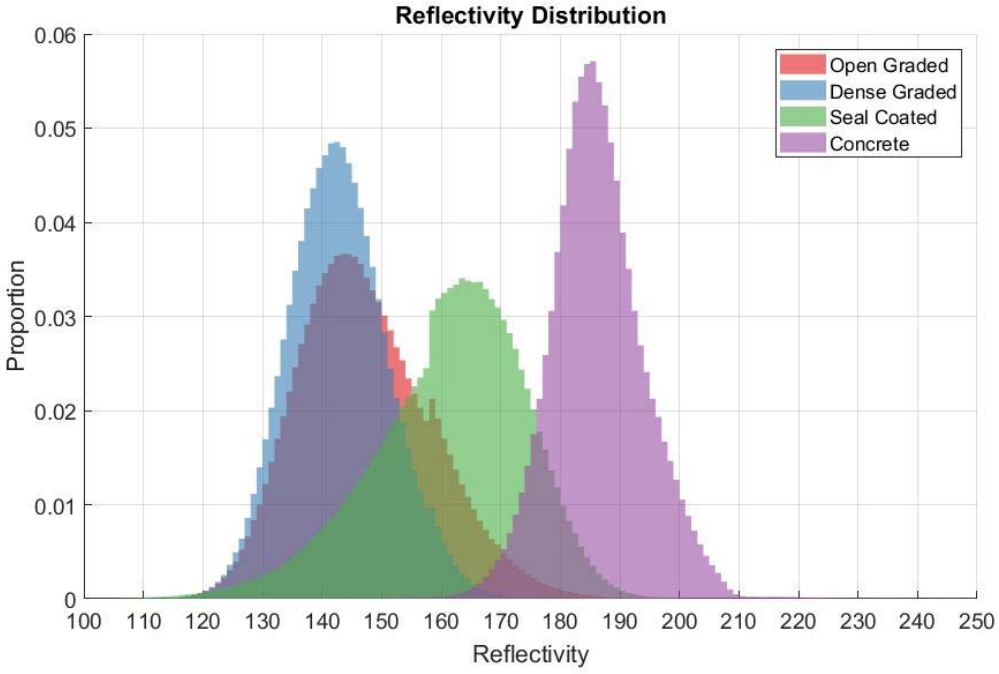


Figure 5: Distribution of reflectivity intensities for different pavement types

To discover discernible parameters for use in the surface identification algorithm, various statistical metrics were studied. These metrics include Mean, Variance, Skewness, Kurtosis, Spectral analysis and check for periodicity are statistical metrics studied for reflectivity values. As can be seen from Figure 6, the mean of reflectivity relatively appears to be the most promising metric for distinguishing between the pavement types considered in the research. However, skewness is quite distinct for seal coated sections, and thus could enhance the accuracy of the developed algorithm when considered along with the mean. As presented in Table 2, skewness for seal coated section was consistently less than that for other pavement types.

Table 2: Distribution of skewness for different pavement types

Pavement Type	**Skewness**
Open Graded	0.776
Dense Graded	3.569
Seal Coated	-0.550
Concrete	0.825

To understand the effect of the section length on these statistical metrics, the analysis was repeated for multiple section lengths: $1/10^{th}$ of a mile (528', 160.9m), $1/20^{th}$ of a mile (264', 80.7m), $1/40^{th}$ of a mile (132', 40.2m) and $1/80^{th}$ of a mile (66', 20.1m). The pattern observed for the 0.05-mile section length was observed for all other section lengths.

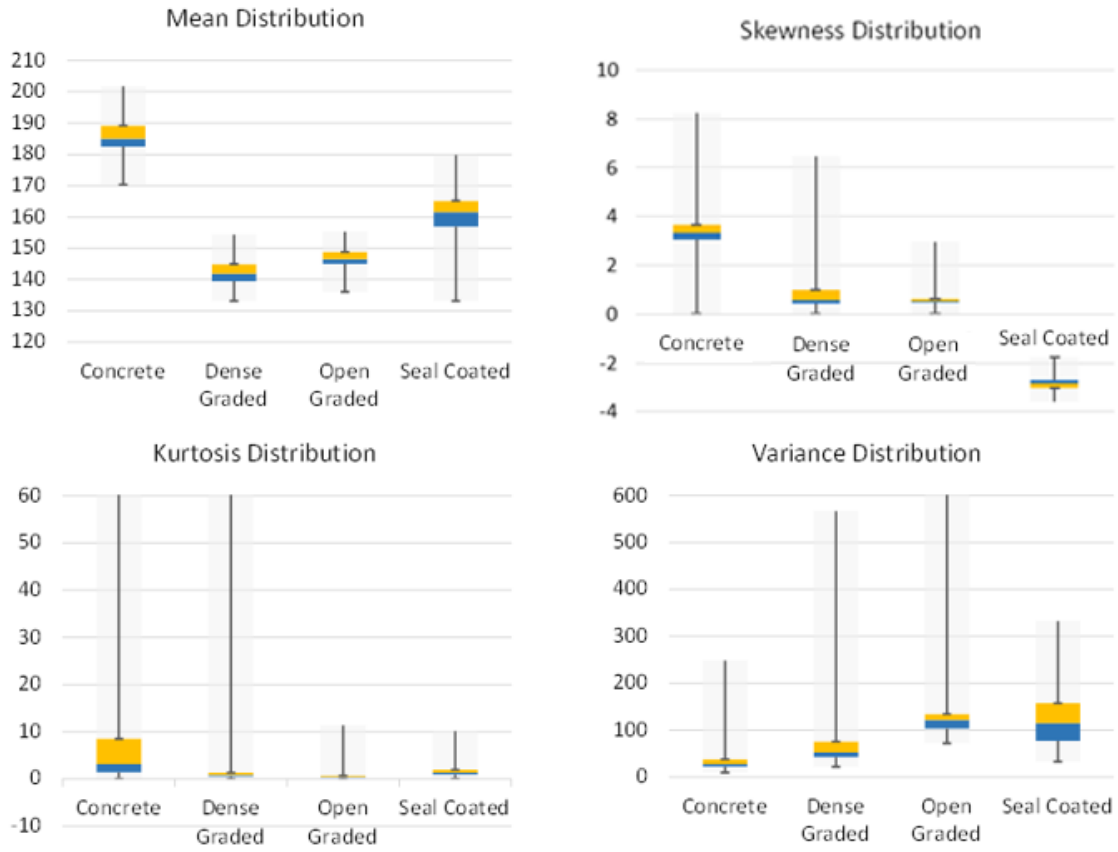


Figure 6: Variation in statistical metrics for pavement surfaces considered in this study. Concrete sample size (n) = 144, dense graded n = 88, open graded n = 223, seal coated n = 301. Each sample unit is 0.05 miles in length.

Figure 7 depicts spectral density of reflectivity values. Seal coated section exhibits multiple peaking for selected non-equispaced fast fourier transform value but no discernable pattern. Lack of any significant spatial periodicity can be observed from the figure. Thus, spectral distribution of reflectivity intensities was excluded from further consideration.

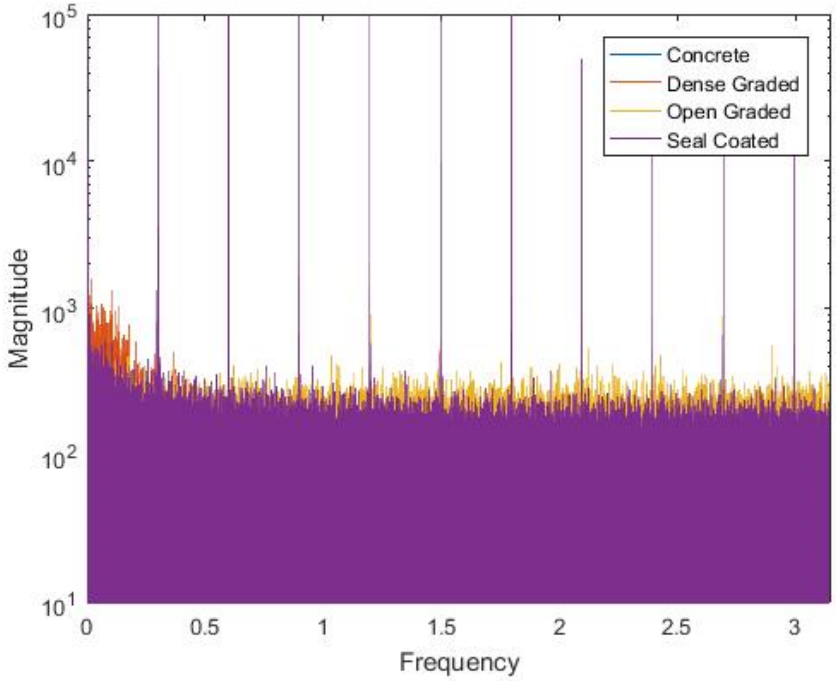


Figure 7: Periodogram power spectral density (PSD) estimate of reflectivity values

In lack of periodicity and any observable patterns and similarity in variance and kurtosis distribution, mean and skewness were deemed the most suitable statistical metrics for developing the pavement surface type detection algorithm.

3.2.1.1 Identification Based on Skewness and Mean

In this method, surface type is identified based on skewness and closeness to mean of reference distributions for known surface types (see Figure 8). These properties are computed as follows:

Skewness:

$$S_k = E\left[\left(\frac{X - \mu}{\sigma}\right)^3\right] = \frac{E\left[(X - \mu)^3\right]}{\sigma^3}$$

Where, S_k = *Skewness (Pearson's moment coefficient of skewness)*

X = *Random variable*

μ = *Mean*

σ = *Standard Deviation*

Closeness to mean:

$$\Delta\mu_{Ri} = |\, \mu_{R\,Test} - \mu_{Ri}\,|$$

$$I = i\ for\ Min(\Delta\mu_{Ri})$$

Where, $\mu_{R\,Test}$ = *Mean of reflectivity distribution for test section*

μ_{Ri} = *Mean reflectivity distribution for reference distribution i*

$\Delta\mu_{Ri}$ = *Absolute difference in means* $\mu_{R\,Test}$ *and* μ_{Ri}

$I = i$ *corresponding to identified reference section*

When mean of reflectivity values for a test section is relatively equidistant from mean of seal coated and concrete, or mean of seal coated and dense graded reference data, skewness is used to distinguish seal coated section from other surface types. Otherwise, type corresponding to the closest reference mean is identified as surface type for the test section (see Figure 9).

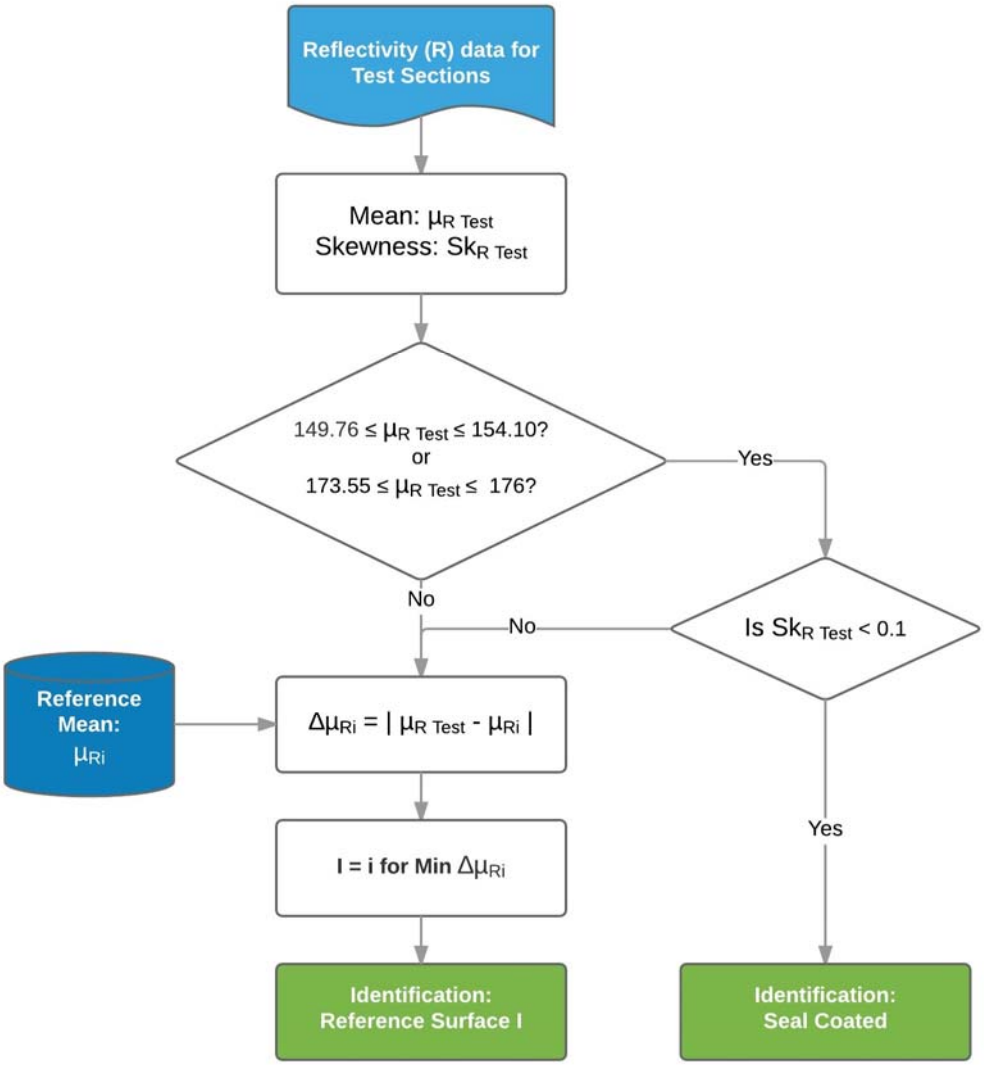


Figure 8: Steps involved in identification based on skewness and mean

The thresholds for mean range and skewness used in this algorithm were determined based numerous tests and iterative sensitivity analysis.

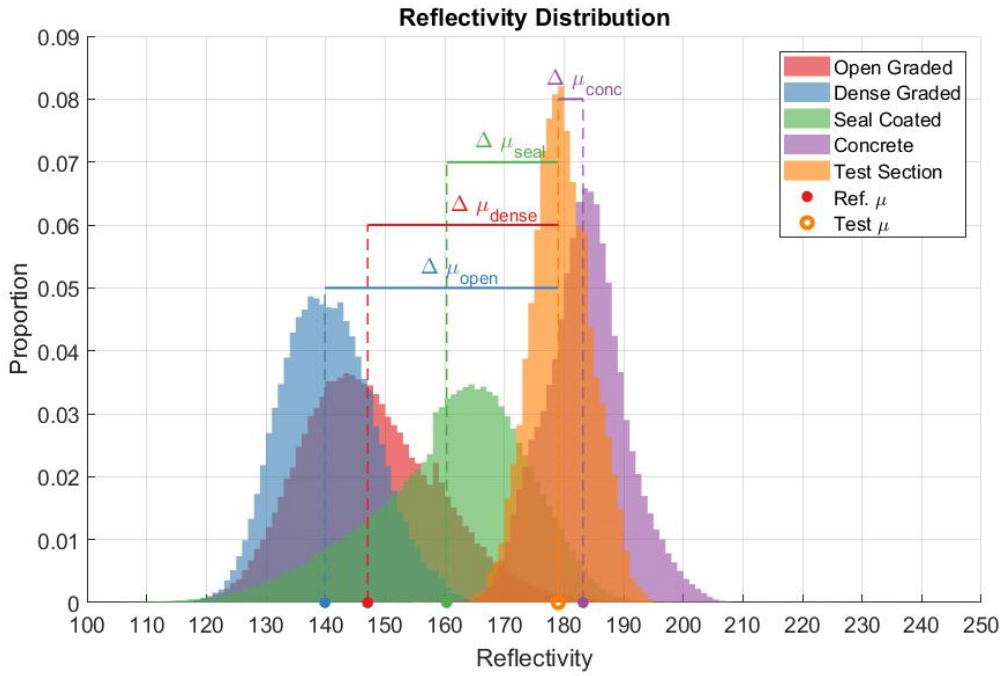


Figure 9: Graphical depiction of detection of pavement surface type based on closeness to mean

A reference data set was created for each surface type to establish reference distributions for these surface types. Each reference data set represents the reflectivity values extracted from driven lane of road sections with known surface type. The surface types are:

- Concrete surface
- Dense graded asphalt surface
- Open graded asphalt surface
- Seal coated asphalt surface
- Asphalt surfaces (includes dense graded asphalt, open graded asphalt, and seal coated)

The size of each reference data set was varied to determine the effect of number of sections on the reference distribution. In each case, a portion of reflectivity data from each road surface type was used as known data (reference data) and remaining portion was used as test data. Reference data and test data were always selected at random from available pool of data. This will remove any inherent biases by user selection.

A computer code was developed to carry out this task automatically based on user-defined instructions. Required set of instruction include list of filenames containing data to be used, location range of data to be extracted, width of road section to consider relative to vehicle position, and other parameters such as data description and output location. These instructions are followed by length of test sections to use and percentage of data to be used as reference data. The code allows the user to specify the number of data points to use for each set and whether to sample randomly or sequentially. The program then divides the total road length into sections of user-defined lengths. Specified percentage of these sections are randomly selected as reference data. These sections are grouped based on surface types and saved as reference data sets.

3.2.2 Results

The charts presented in Figure 10 show variation in accuracy for classification into four categories – concrete, dense graded, open graded and seal coated pavements. Reference data size and test length are varied (from $1/80^{th}$ of a mile to $1/10^{th}$ of a mile) for each case. Accuracy is computed as:

$$Accuracy = \frac{TP}{N} * 100$$

Where TP = True Positives (number of correctly identified sections); N = Total number of sections in the test data set.

While identifying between four types – Concrete, Dense Graded, Open Graded and Seal Coated pavement surfaces, accuracy remains high for concrete sections. Identification of dense graded vs open graded section is found to be most difficult. This is due to the fact that reflectivity distribution for these two pavement types are very close to each other. It is observed that overall accuracy is only slightly increased when the size of the reference data set increases many folds. For example, in case of test length of 1/80th of a mile, increasing reference data size from around half a million points to 6.5 million points, the increase in accuracy is less than 3.5%. The dotted trend line illustrates trend of overall accuracy as the reference size and test lengths are increased. Change in accuracy stayed within a range of about 6% for all cases.

Similar analysis was carried out for asphalt vs concrete identification. Overall accuracy of more than 95% was achieved, as shown in Figure 11, for different reference data size and test lengths. Change in accuracy stayed within a range of about 5.5%.

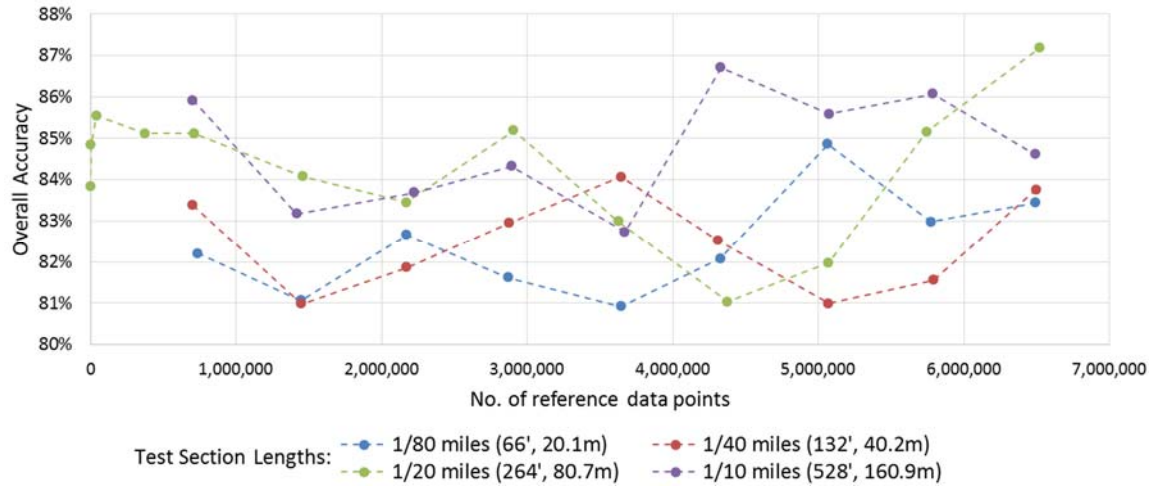


Figure 10: Overall accuracy – 4 category classification

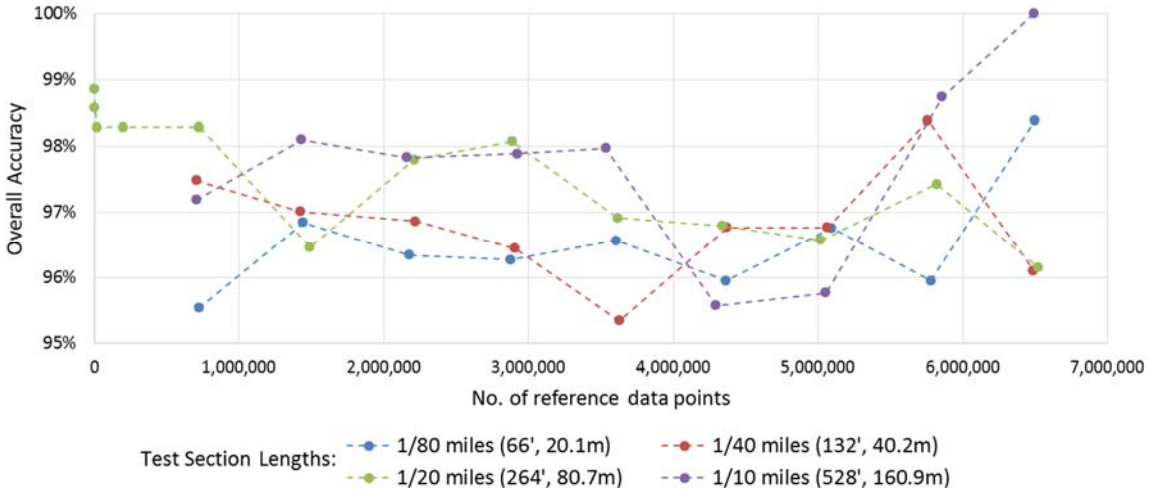


Figure 11: Overall accuracy - asphalt vs concrete identification

As an example, taking a closer look for a case where test length is 1/20th of a mile (i.e. 0.05 miles), Figure 12, different reference data size was used starting from 400 values (100 for each pavement type) to about 6.5 million values. Change in accuracy was very

small and remained within a small range of about 3% for asphalt vs concrete test and about 6% for 4 categories test.

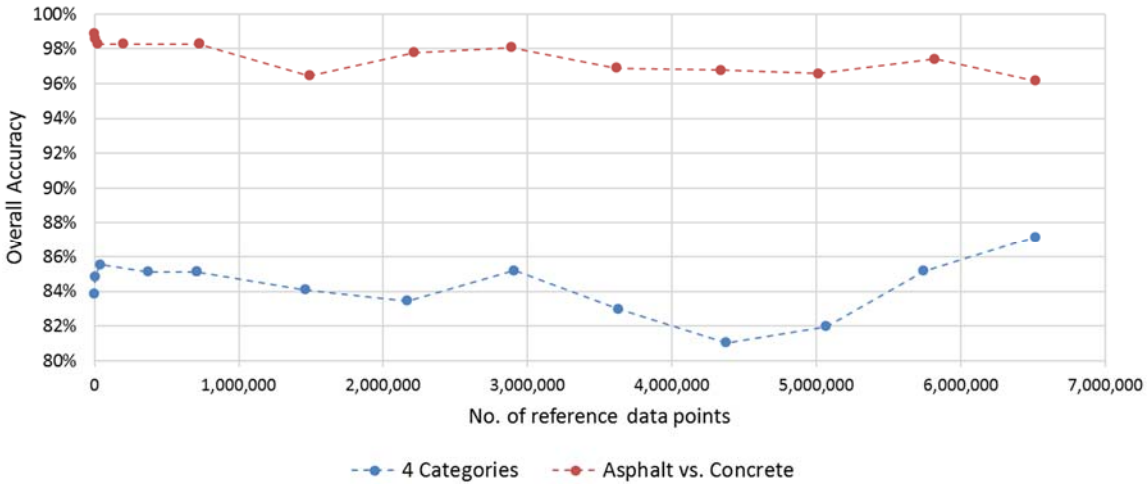


Figure 12: Reference data size and overall accuracy for test length of 0.05 miles

Based on Appendix B, identification accuracy for sections not represented in reference data was low. Reflectivity is found to be dependent on pavement wetness, surface asphalt concentration, polishing of aggregates, and other distresses. Identification accuracy for such sections was found to be poor.

The graph below (Figure 13) depicts accuracy of prediction based on mean and skewness for 4 - categories identification (blue bars) and asphalt vs concrete identification (green bars). In this graph, accuracy is calculated for each pavement surface type, separately. In this case, 20% of randomly selected data is used as a reference data and the remainder as test data. The highest accuracy of detection was achieved for concrete surface (95.39%) and the lowest accuracy of detection was for dense graded surface (63.33%).

For open graded sections, it is found to be 88.44% and in case of seal coated sections it is 92.58%. For distinguishing between asphalt and concrete, accuracy is found to be 98.08% for asphalt surface and 95.39% for concrete surface.

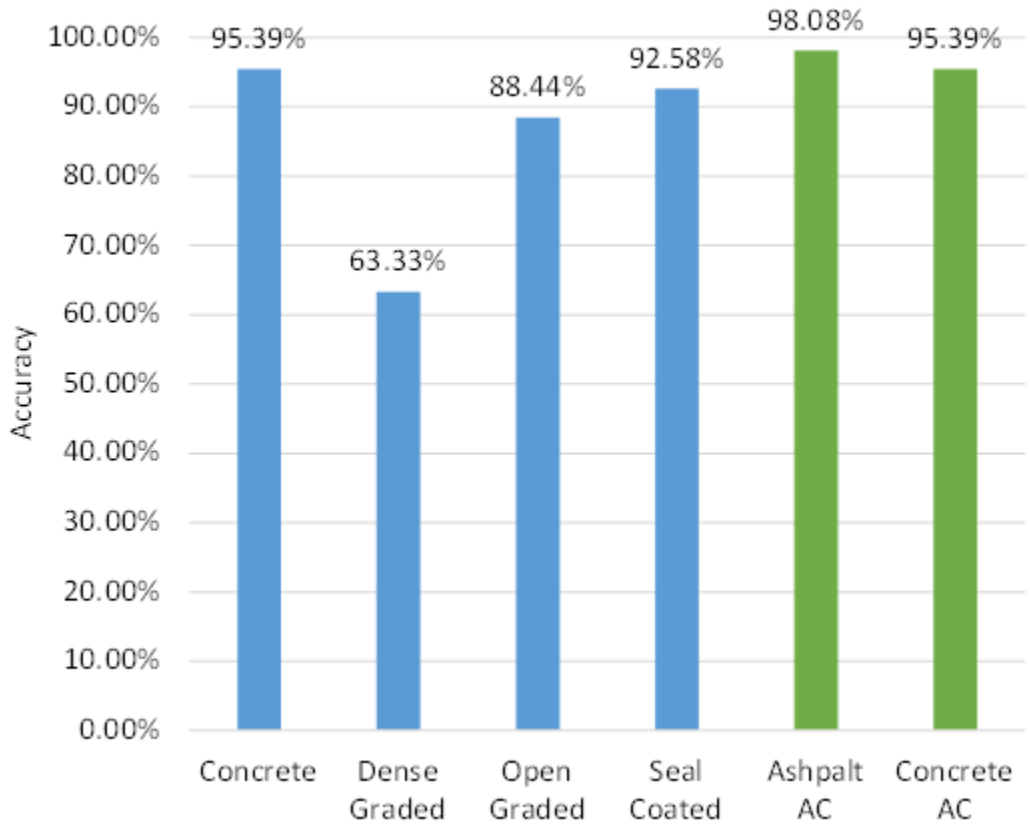


Figure 13: Accuracy of prediction based on mean and skewness for 0.05 mile (264', 80.47m) test sections

3.2.3 Adjustment Based on Adjacent Section

For network-level studies and applications, pavement surface type is likely to extend for a long segment of the road (e.g., 10 miles). This information is likely to improve the accuracy of detection at the cost of reduced ability to detect small changes in surface

type (e.g., patchwork). Additional steps were added to the surface type detection algorithm to account for this knowledge. These additional steps are depicted in Figure 14 and are described as follows:

- Current identification is compared with identification of previous adjacent five sections and subsequent adjacent five sections.
- If three out of five identifications in either direction are the same, identification for current test section is set to that identification. Algorithm is stopped at this point.
- If there is a difference, the next 11 identifications are compared.
- If 6 out of 11 of those is the same, identification is set to the common value. Algorithm is stopped at this point.
- If no agreement is achieved, no change is made to identification postulated by the skewness and mean algorithm is kept.

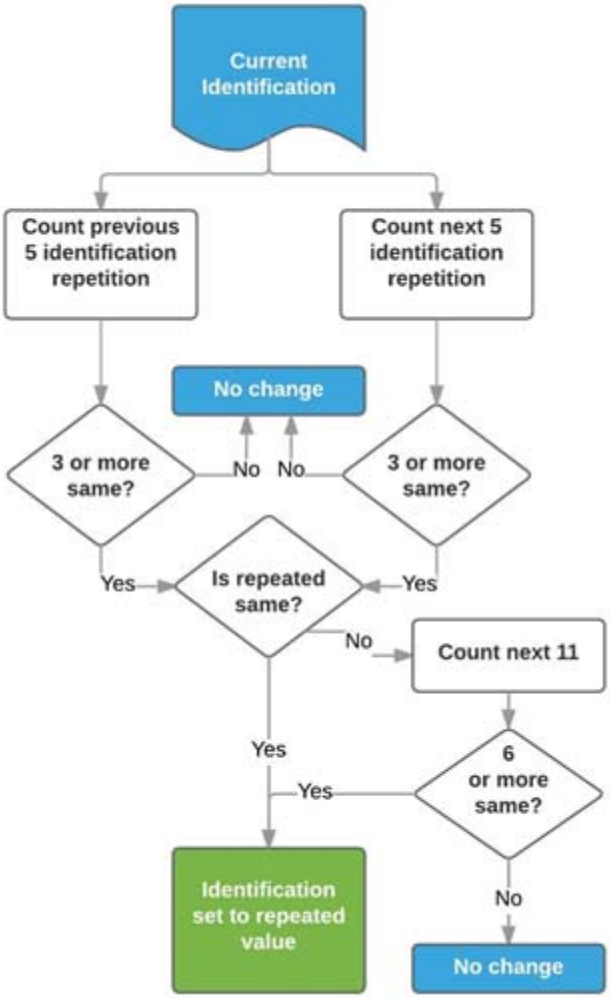


Figure 14: Steps involved in adjustment of pavement surface type based on adjacent section

To arrive at the thresholds used in this algorithm (3 out of 5 and 6 out of 11), extensive tests were carried out comparing accuracies for test sections used in the study.

Based on the Adjustment by Adjacent Section, the accuracy for concrete and asphalt sections increased to 100%. Figure 15 reveals that the accuracy in identifying

dense graded sections has increased to 73.33%. Similarly, the accuracy of open graded and seal coated sections increased to 89.33% and 97.42%, respectively.

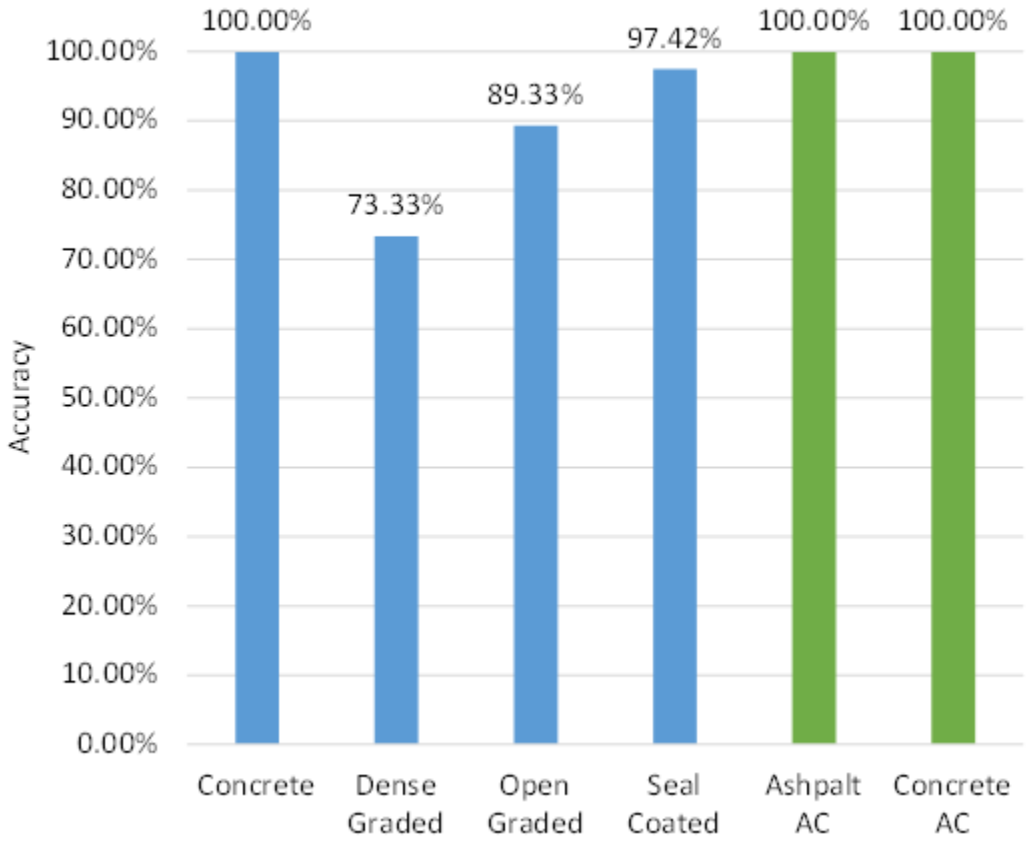


Figure 15: Accuracy of prediction with adjustment based on adjacent section for 0.05 mile (264', 80.47m) test sections

3.3 Grass Detection

3.3.1 Formulation

Figure 16 depicts histograms for reflectivity distribution in different pavement material along with distribution for pure roadside grass. Pure roadside refers to continuous roadside grass areas (i.e., after removal of driveways and dirt areas). Reflectivity distribution for grass is quite distinct from asphalt pavements. When comparing with concrete pavement, the distinction is not as much. This translates to reduced accuracy when grass section is adjacent to a concrete section. This difficulty was observed during development of pavement surface type identification method in Section 3.2. The implications of this is further discussed in subsequent sections. Reflectivity distributions in Figure 16 consider reflectivity values for pavement within 1m from center of MLS vehicle while reflectivity values for grass section is aggregated between edge of pavement to 4m into the roadside. Different studies discussed in Section 2 (e.g., Kashani et al. 2015) observed attenuation of reflectivity with range and angle among other environmental factors. Same phenomena was observed during development of pavement surface type identification method (Section 3.2). Attenuation is further discussed in Section 3.3.1.1. Figure 17 shows such aggregated reflectivity distribution between all of roadside and pure grass sections. Reflectivity for both cases is aggregated from edge of pavement to distance of 4m from MLS vehicle center into the roadside. Since area occupied by driveways and discontinuities is much less compared to grass areas, distribution observed is almost identical.

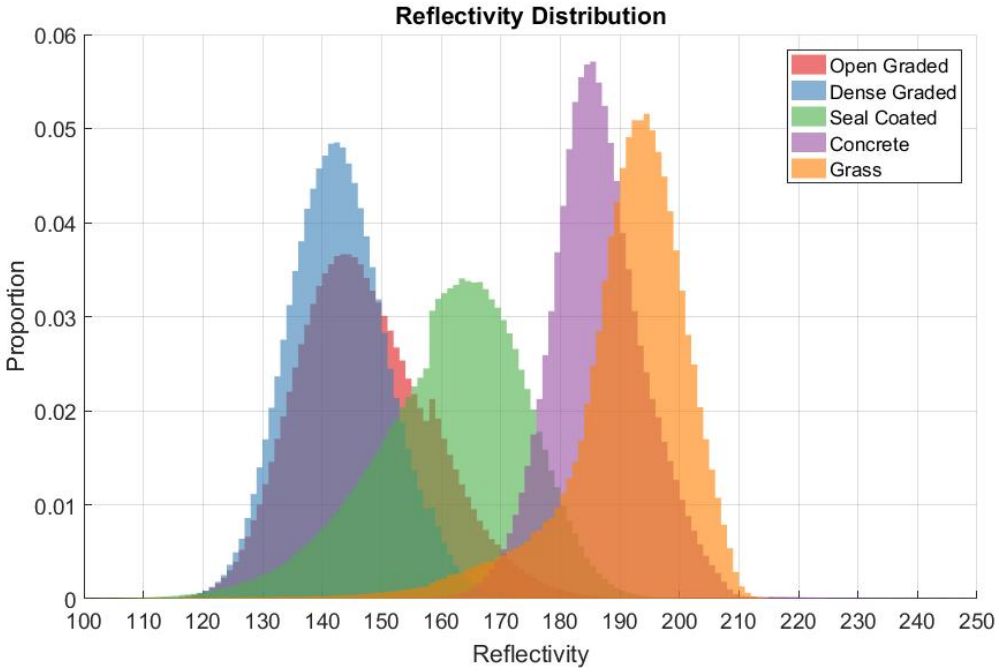


Figure 16: Reflectivity distribution for different pavement material and grass.

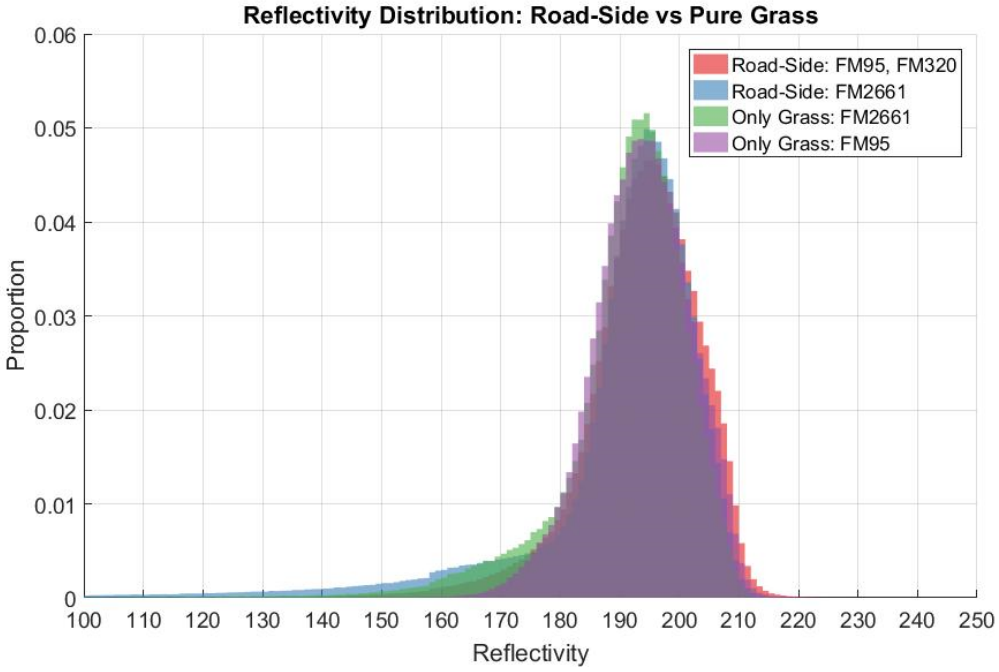


Figure 17: Reflectivity distribution comparison between roadside and pure grass

Reflectivity distribution in dirt section is observed to be slightly higher than for grass section. Even though it is inseparable by statistical means developed thus far, this observation can be useful in identifying dirt patches in grassy areas. A portion of available dirt section (from LP79N road) contained asphalt and grass impurities. Hence double peaking is observed in its distribution.

3.3.1.1 Attenuation of Reflectivity Values

Figure 18 shows reflectivity histograms for pure grass areas as the LiDAR readings move away from the MLS vehicle center. Abscissa denote reflectivity values; ordinate on the left denotes proportion for histograms; and ordinate on the right denotes distance from center of MLS vehicle. When the mean reflectivity is plotted on this chart, a clear attenuation trend is observed. Further, Figure 19 highlights this attenuation with mean values and standard deviation of reflectivity.

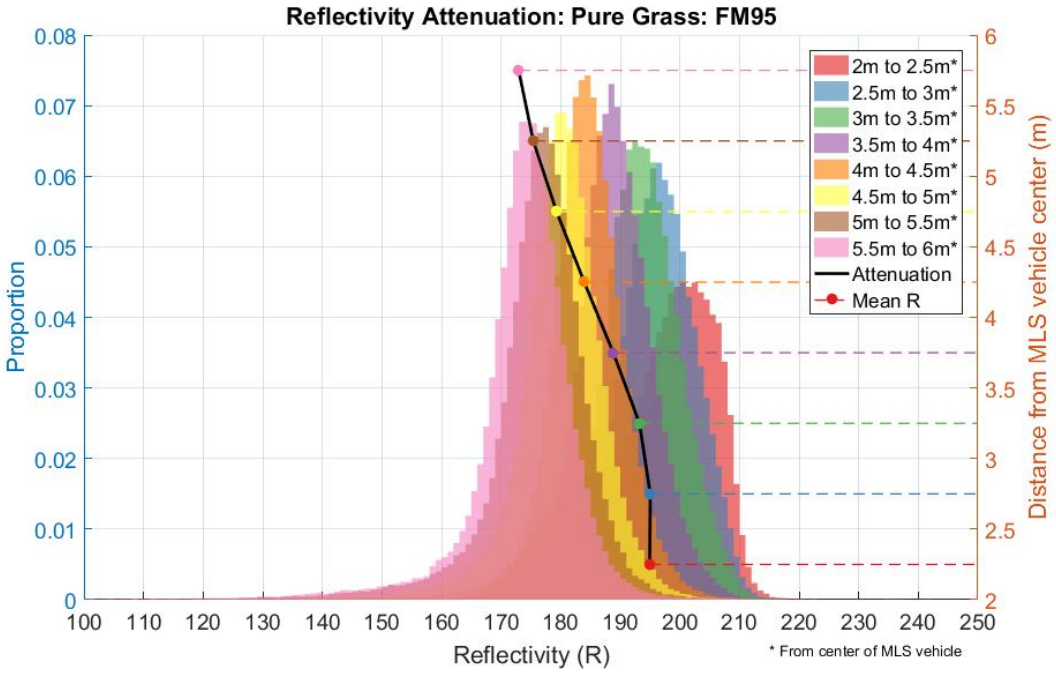


Figure 18: Attenuation of reflectivity for grass sections

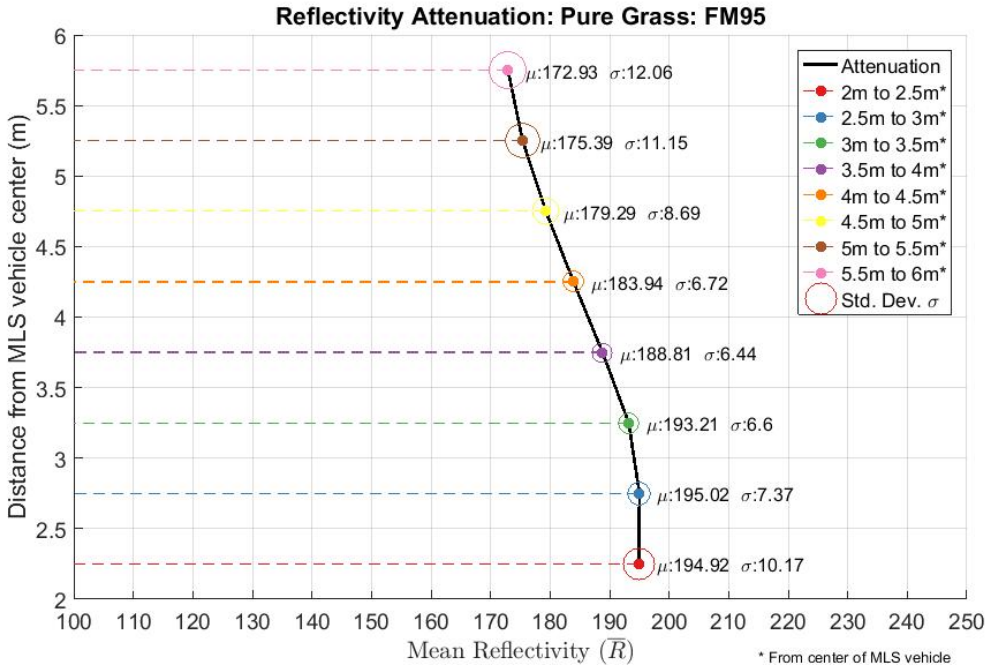


Figure 19: Variation of mean and standard deviation of reflectivity values for grass (transverse direction)

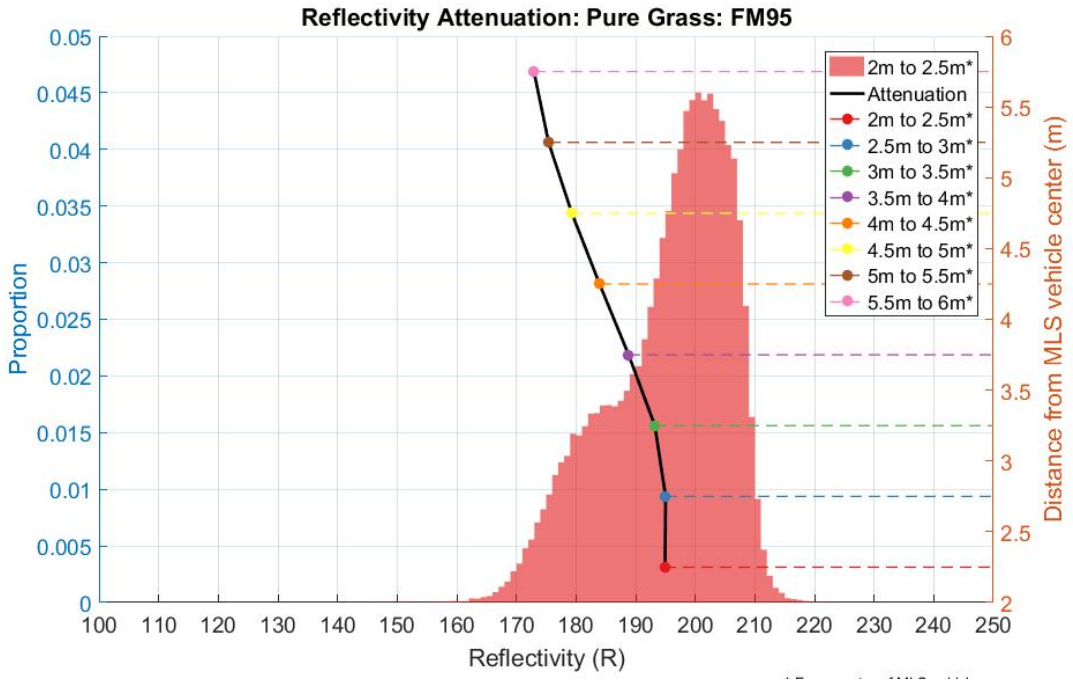


Figure 20: Reflectivity distribution in grass (first distance range)

At a distance of 2 to 2.5m from the center of the MLS, the surface contains asphalt and dirt impurities, as made evident by the shape of the histogram and change in attenuation trend for the first distance range (Figure 20). Therefore, a correction in reflectivity distribution was applied for this range to remove the effect of asphalt impurities and a new mean was calculated. Reflectivity correction is the difference between mean reflectivity at a certain distance from MLS center and reflectivity at a distance of 2.5m from MLS towards the roadside (distance at which roadside grass started for the test section). Reflectivity difference is computed as follows:

$$\Delta r = \mu_{r@d} - \mu_{r@2.25m}$$

where, $\mu_{r@d}$ = *Mean of reflectivity distribution at a transverse distance d*

$\mu_{r@2.25m}$ = *Mean of reflectivity distribution at a transverse distance of 2.25m*

A regression analysis was conducted (see Figure 21) to determine the reflectivity correction at any given point as a function of transverse distance from the MLS center towards the roadside (*d*). The best fit equation is as follows:

$$\bar{r} = 8.03d - 17.82$$

where, $\bar{r}$ = *Reflectivity correction*

d = Transverse distance from the MLS center (2.25m < d < 6m)

Attenuation-corrected reflectivity is computed as follows:

$$r = r_i - \bar{r}$$

where, r_i = *Initial reflectivity before correction*

As we can observe, there is miniscule difference in linear and quadratic attenuation equations and attenuation behavior can be approximated as a linear function of distance from MLS center for efficiency and simplicity. Figure 22 depicts attenuation corrected reflectivity for pure grass at different distance and Figure 23 depicts reflectivity distribution in different pavement types with attenuation correction applied to grass reflectivity. Mean and standard deviation is presented in Table 3. In comparison with Figure 16, distribution of reflectivity in grass is more distinct from pavement surface types after applying the attenuation correction.

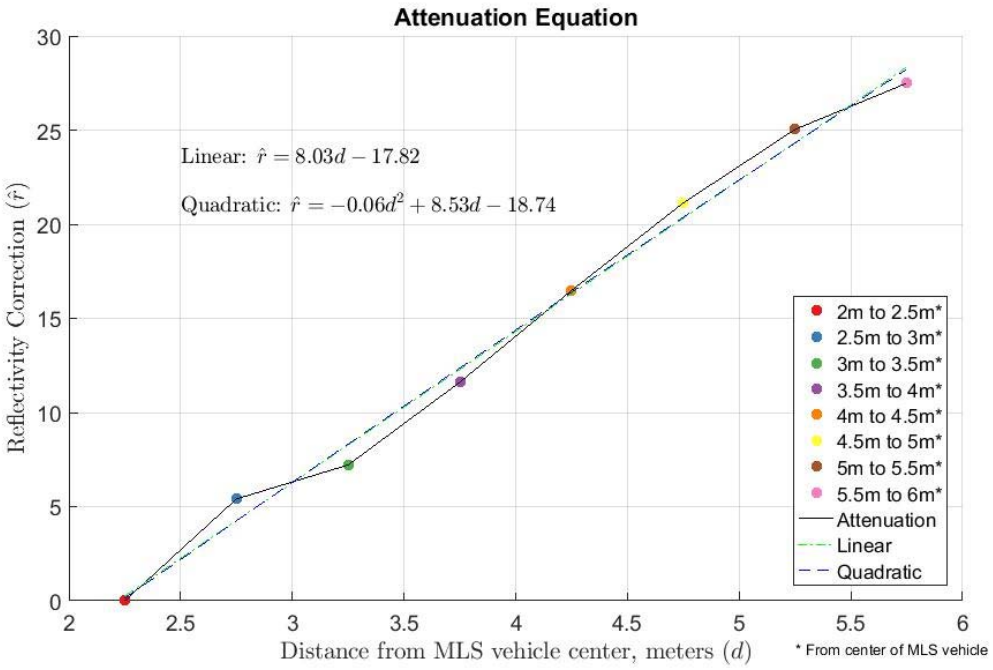


Figure 21: Attenuation correction equation

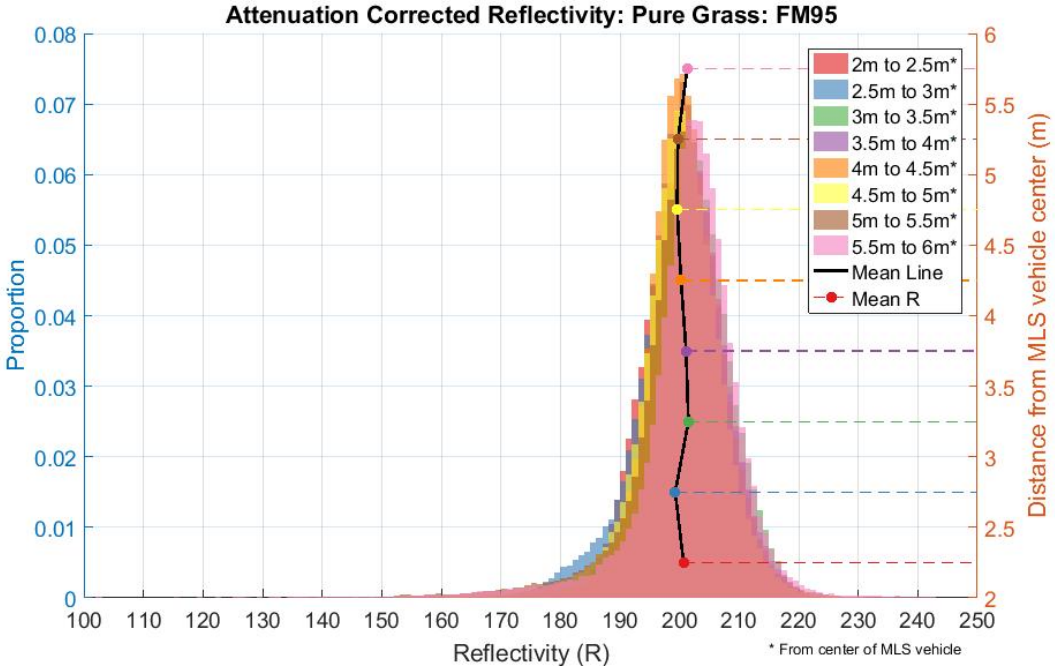


Figure 22: Attenuation corrected reflectivity distribution for pure grass.

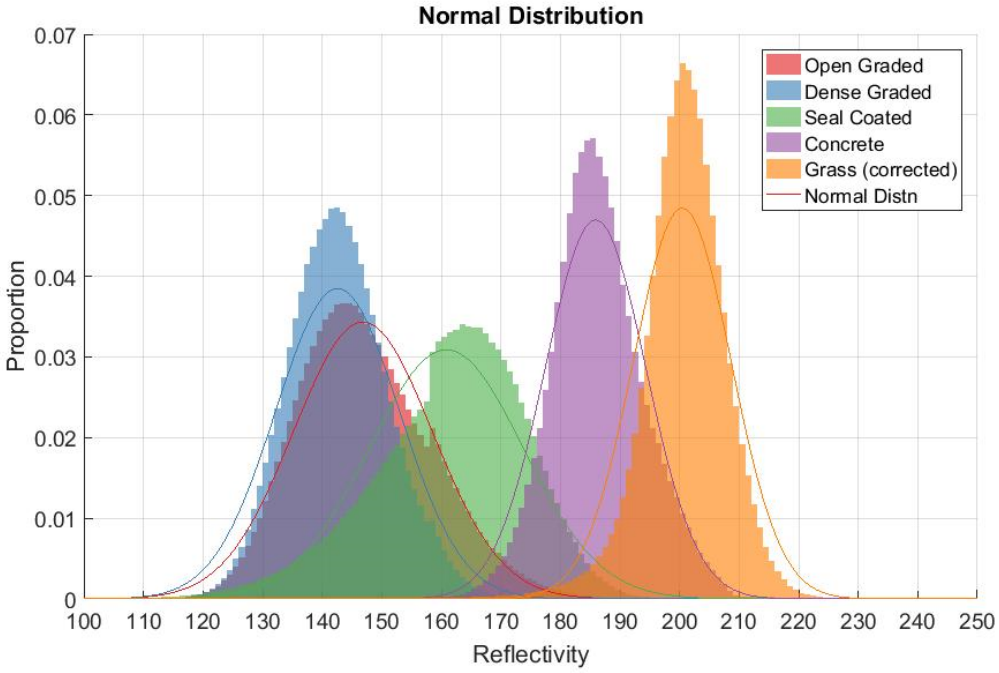


Figure 23: Reflectivity distributions with attenuated corrected pure grass and fitted normal distributions

Table 3: Mean and standard deviation for reflectivity distributions with attenuated corrected pure grass

	Open Graded	Dense Graded	Seal Coated	Concrete	Grass
Mean	146.84	142.59	160.86	185.91	200.45
Standard Deviation	11.62	10.37	12.91	8.48	8.23

3.3.1.2 Identification of Grass

Normal distribution is fitted to the reflectivity distributions as shown in Figure 23. The choice of normal distribution was made based on histograms of reflectivity distribution (bin size 1 for 8-bit integer reflectivity representation), and simplicity in estimation of probability for the detected material type. Method for identification of grass

is postulated with the aim of distinguishing grass from pavement material. As such, identification between grass vs asphalt and grass vs concrete are the two problems to be solved in this section (Figure 24 and Figure 25). Parameters required for normal distribution (mean and standard deviation) are presented in Table 4.

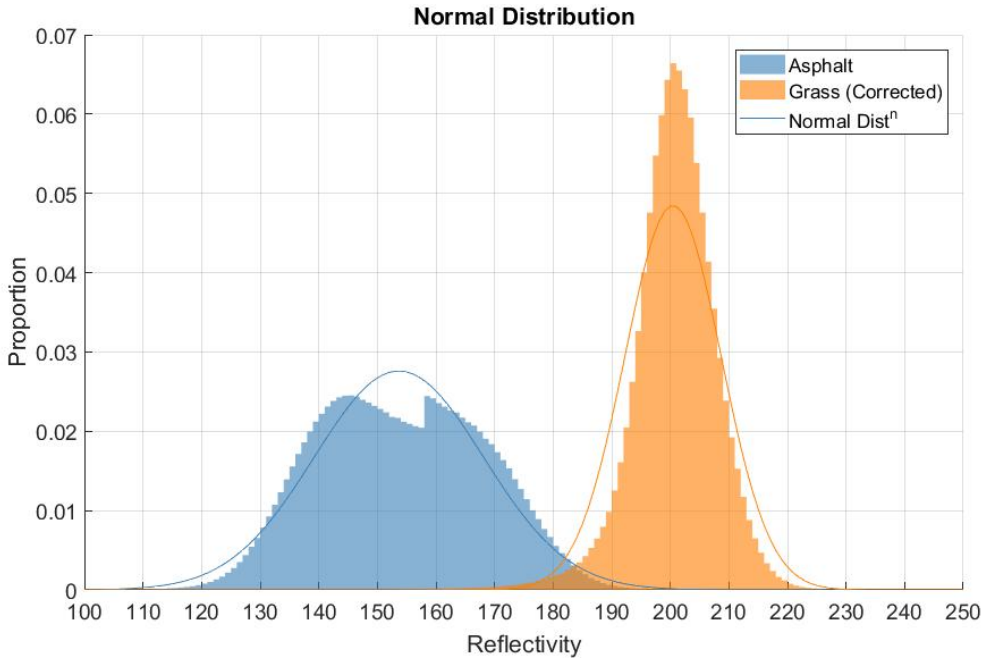


Figure 24: Reflectivity distributions for asphalt and attenuated corrected pure grass with fitted normal distributions

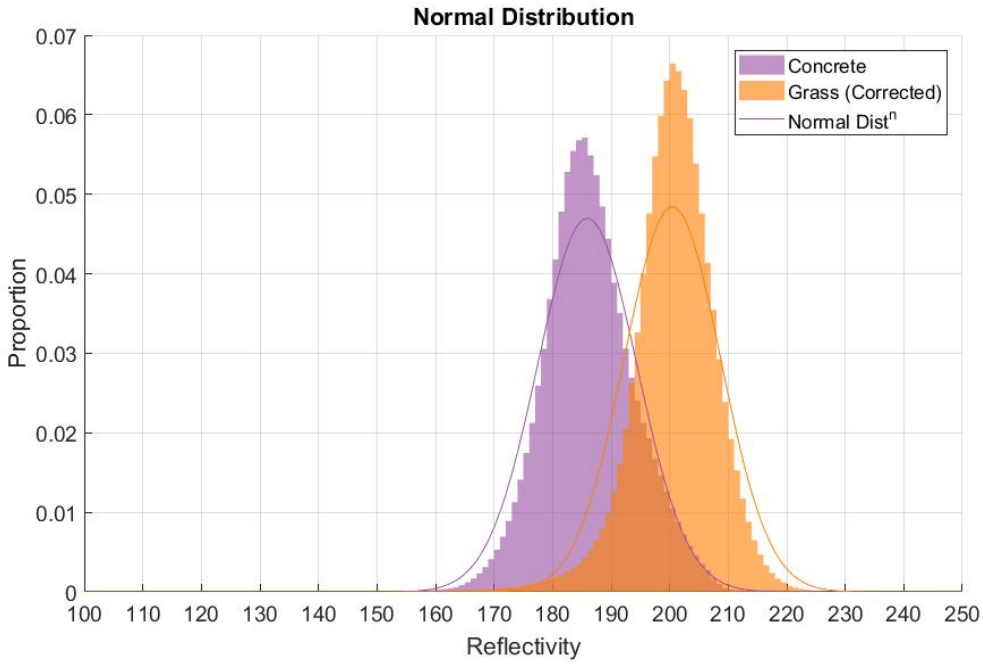


Figure 25: Reflectivity distributions for concrete and attenuated corrected pure grass with fitted normal distributions

Table 4: Mean and standard deviation for reflectivity distributions in Asphalt, Concrete and Grass surface

	Asphalt	Concrete	Grass
Mean	153.656	185.91	200.45
Standard Deviation	14.45	8.48	8.23

Identification accuracy for each material type is calculated based on these probability distributions functions (Figure 24 and Figure 25). The underlying assumption is that the material (or rather reflectivity) being tested belongs to one of these types. Probability for a value to belong to each of the material type is calculated and rescaled such that the summation of probabilities is one. The confidence level for a test point to be

grass is depicted in Figure 26 for asphalt roads and in Figure 27 for concrete roads. The confidence level is computed as follows:

$$pdf_{Test\ i} = pdf(Reference\ Reflectivity\ \ i, \qquad Test\ Value)$$

$$c_i = \frac{pdf_{Test\ i}}{\sum_k^n pdf_{Test\ k}}$$

where, pdf = *probability density function*

c_i = *confidence level for test value to be material i*

n = *total number of material in consideration (two for asphalt vs grass or concrete vs grass)*

As evident in these figures, for a reflectivity value, confidence in grass identification for concrete road is lower than for asphalt road. This is due to the fact that reflectivity distribution for concrete is much closer to that of grass. Therefore, using the method developed, grass area extraction for asphalt road will be more accurate.

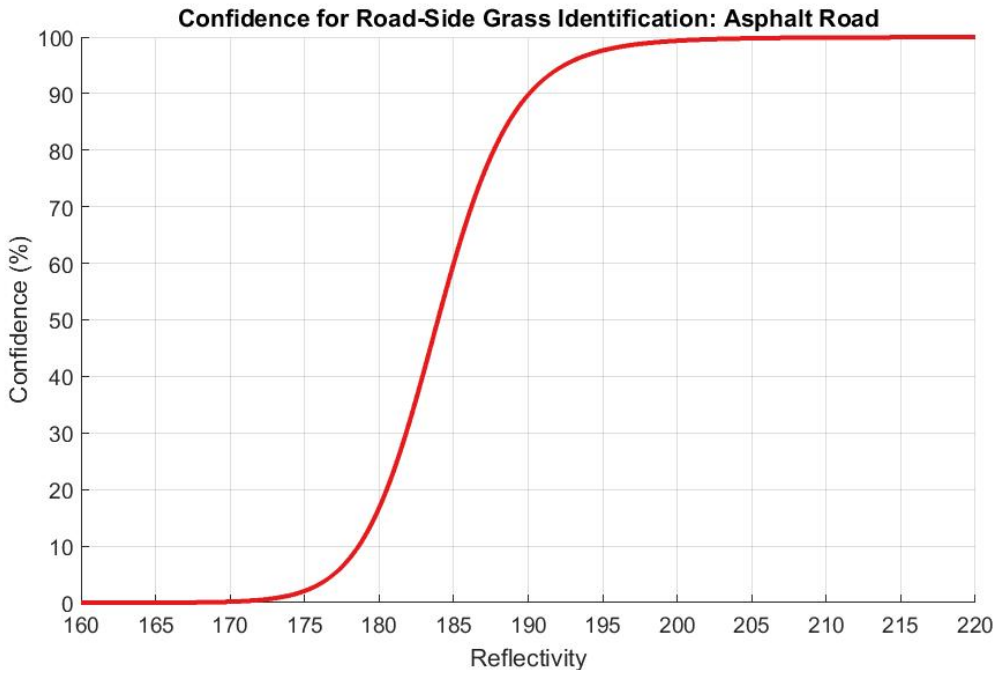


Figure 26: Confidence level for grass identification based on attenuation-corrected reflectivity (for asphalt roads only)

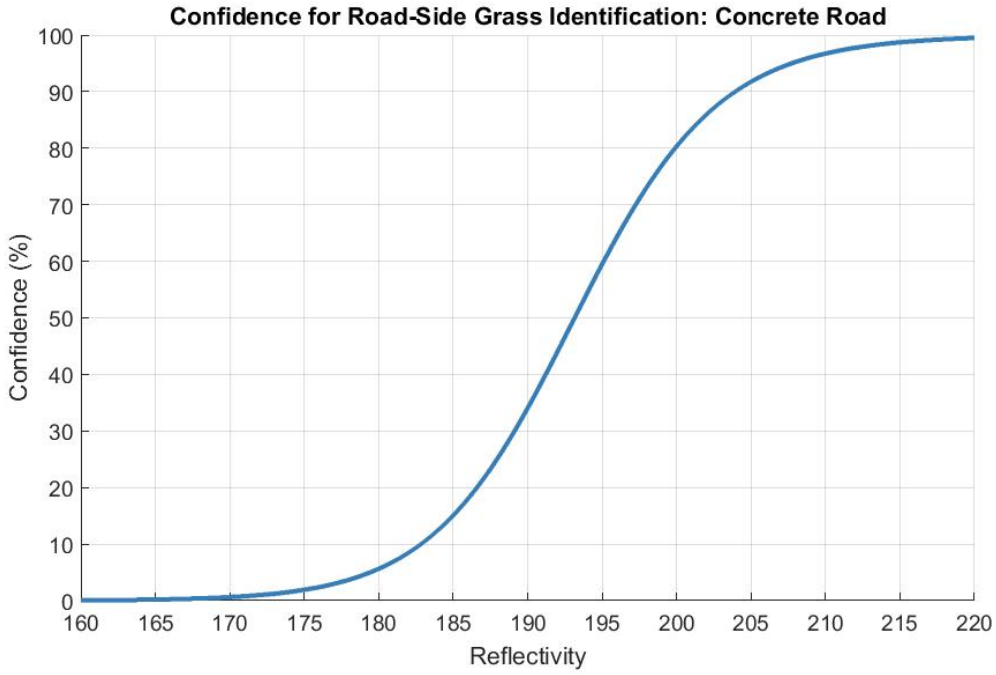


Figure 27: Confidence level for grass identification based on attenuation-corrected reflectivity (for concrete roads only)

3.3.2 Results

To test the proposed roadside grass detection method, one thousand random samples were extracted from each of the known surface types – open graded, dense graded, seal coated, concrete, and pure grass. The road sections used for this tests are presented earlier in Table 1. The reflectivity values for the samples were processed using the developed algorithm. Results obtained are presented in Table 5. When the pavement is asphalt (i.e., the identification is asphalt vs. grass), grass is identified accurately for 98.7% of the samples. When the pavement is concrete (i.e., the identification is concrete vs. grass), 90.3% of grass samples were correctly identified and 16.3% of concrete points were misidentified as grass. So clearly, the result is better for roadside grass detection in asphalt road as depicted in Figure 26.

Table 5: Grass detection test results

SN	Actual Type	Percent Reflectivity Points Identified As		
		Asphalt	Concrete	Grass
1	Grass	1.3	~	98.7
2	Open Graded	99.7	~	0.3
3	Dense Graded	99.6	~	0.4
4	Seal Coated	97.6	~	2.4
5	Grass	~	9.7	90.3
6	Concrete	~	83.7	16.3

4 DEVELOPMENT OF DRIVEWAYS AND UNDERLYING PIPES IDENTIFICATION METHOD

This section deals with development of a method for the identification of driveways and estimation of underlying stormwater pipe length. The method extracts location, width, elevation difference across driveways, and driveway surface material type. A section of FM320 road is used for development of the method. The two lane seal coated section is of 470m in length and contains roadside ditches on both sides. Table 6 presents driveway information collected on the section based on visual observation. For testing, two sections were used: another FM320 section 5000m in length and a 5300m section of FM2661. These sections also contain two lane – seal coated roadway with roadside ditch in most places. The FM2661 section contains paved and mostly flat driveways whereas the FM320 section (used for testing) contains mixed driveways (gravel, grass, deteriorated asphalt, etc.).

Table 6: Visual observation of FM320 driveways used for method development

SN	Location Start	Location End	Remark	Material
1	32	42	Driveway	Dirt
2	62	77	Driveway	Dirt + Asphalt
3	181	188	Driveway	Dirt
4	200	206	Driveway	Concrete
5	223	229	Driveway	Asphalt + Concrete
6	235	254	Driveway	Concrete

Table 6 Continued

SN	Location Start	Location End	Remark	Material
7	258	264	Driveway	Grass
8	265	276	Driveway	Asphalt
9	359	370	Driveway	Dirt
10	445	459	Adjoining road	Concrete

4.1 Formulation

In order to detect a driveway and estimate relevant metrics, the primary challenge lies in detecting changes in elevation distribution and preventing false positives for elevated areas lacking proper ditches that are not driveways. In theory, a properly designed section without sharp changes in ditch geometry, jump in ditch bottom elevation can be used to detect presence of driveways. However, in many practical scenarios, this is not sufficient as sudden jump in elevation might not be present all the time and there may be instances where a sudden jump detected is due to change in longitudinal profile of the ditch (or road section) rather than due to presence of a driveway. To overcome this issue, the study considers cross sectional geometry rather than longitudinal, in conjunction with additional criteria such as smoothness and material type to verify suspected driveways.

Figure 28 depicts the steps involved in the postulated method for driveway identification. This method applies to roadways with open ditches only. In other words, this method does not apply to roadways with curb and gutter. Elevation and reflectivity are extracted from mobile LiDAR data. The governing assumption for detecting underlying pipes and estimating pipe length is that every driveway over a ditch has an

underlying pipe. In all cases, the driveway width is assumed to be equal to the pipe length. This introduces a limitation in estimating driveway width when there is a protruded pipe. Additionally, start-end elevation difference (ΔE) is estimated using this method based on average of three consecutive cross sections on either side of the detected driveway, as follows:

$$\Delta E = \bar{E}_{start} - \bar{E}_{end}$$

where, $\bar{E}_{start}$ = *mean of three consecutive ditch-bottom elevation before a driveway starts*

$\bar{E}_{end}$ = *mean of three consecutive ditch-bottom elevation after a driveway ends*

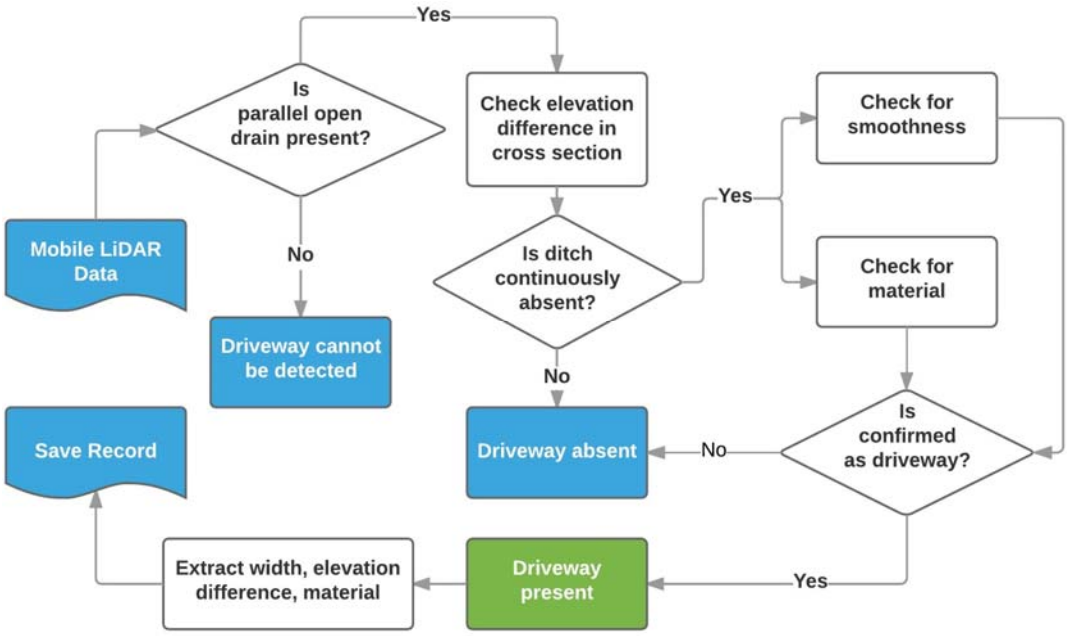


Figure 28: Steps involved in postulated method for driveway identification

4.1.1 Detection of Ditches Using Cross Section Geometry

The presence of roadside ditches is detected based on differences in cross sectional elevation, as presented in Figure 29. Elevation data is extracted from Mobile LiDAR Data. Minimum roadside elevation and average road surface elevation are calculated. This calculation is followed by ditch detection formulation. The criterion for the presence of a ditch is that the difference between the average road surface elevation and minimum roadside elevation should greater than 0.12m. This tolerance value is required since in many cases driveways are depressed over a ditch, either by design or due to settlement or erosion. The criterion for the presence of a driveway is that the ditch is continuously absent for more than 3.5m. These threshold values were determined in an iterative process that included many trials.

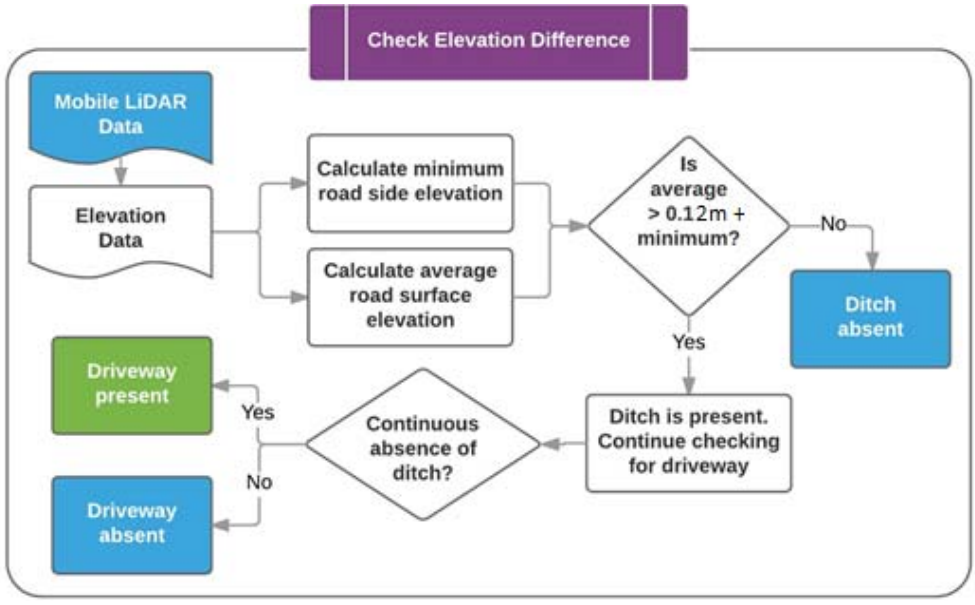


Figure 29: Identification of driveways based on cross section geometry.

4.1.2 Verification of Driveways Based on Material Type

A reference data set is defined for the road section under consideration. To improve the accuracy of material detection - as discussed earlier in Section 3.2 of this study, the dataset should contain known material reflectivity for road and roadside feature only. Reflectivity values for the suspected driveway area is extracted and tested against known values using methods developed in Section 3.2. If grass reflectivity values are used in the reference data to distinguish between driveway material and roadside material, grass driveways cannot be identified using this technique. Steps involved in this method are presented in Figure 30.

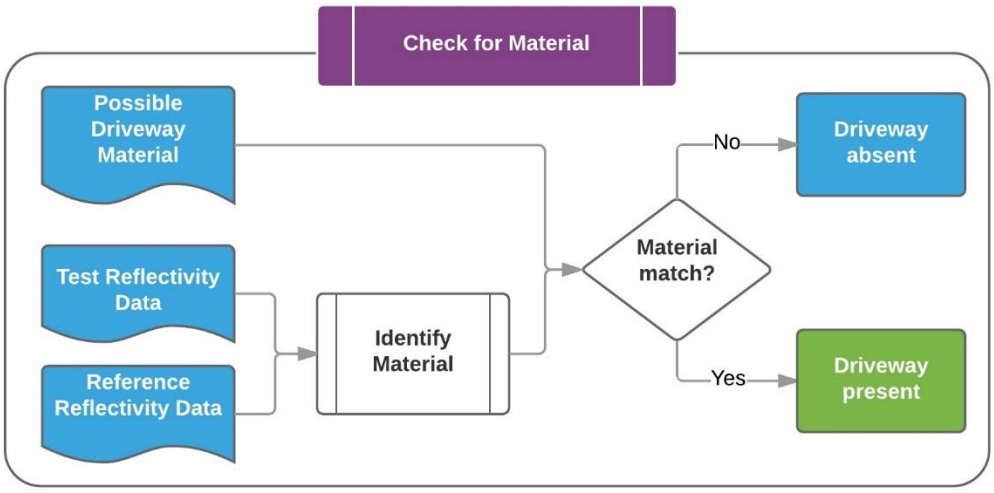


Figure 30: Verification of driveway detection based on material type.

4.1.3 Verification of Driveways Based on Smoothness

Final step in the developed method is to verify the suspected driveways based on measuring elevation change in the suspected driveway area. The basic assumption here is that the driveway has a smooth flat surface. This is based on the assumption that driveways are longitudinally smoother than dirt pile blocking the ditch, for example. Figure 31 depicts steps developed for this purpose. Elevation data for suspected driveway area is extracted. Maximum and minimum elevation in each longitudinal line (based on data grid) is calculated. To account for anomalies in the collected data, maximum elevation is defined as an average of ten high elevations, and minimum is defined as an average of ten low elevations. Driveway flatness (i.e., surface smoothness) is defined as the minimum of differences in elevation between maximum and minimum values of all longitudinal lines along the driveway (see Figure 32). This measure of smoothness ensures that the gradient in transverse direction does not influence smoothness measure. Based on this measure of smoothness, a value of zero corresponds to perfectly smooth surface. If smoothness is less than 0.02m, then the suspected area is verified as a driveway. Multiple trials are carried out to arrive at this threshold value.

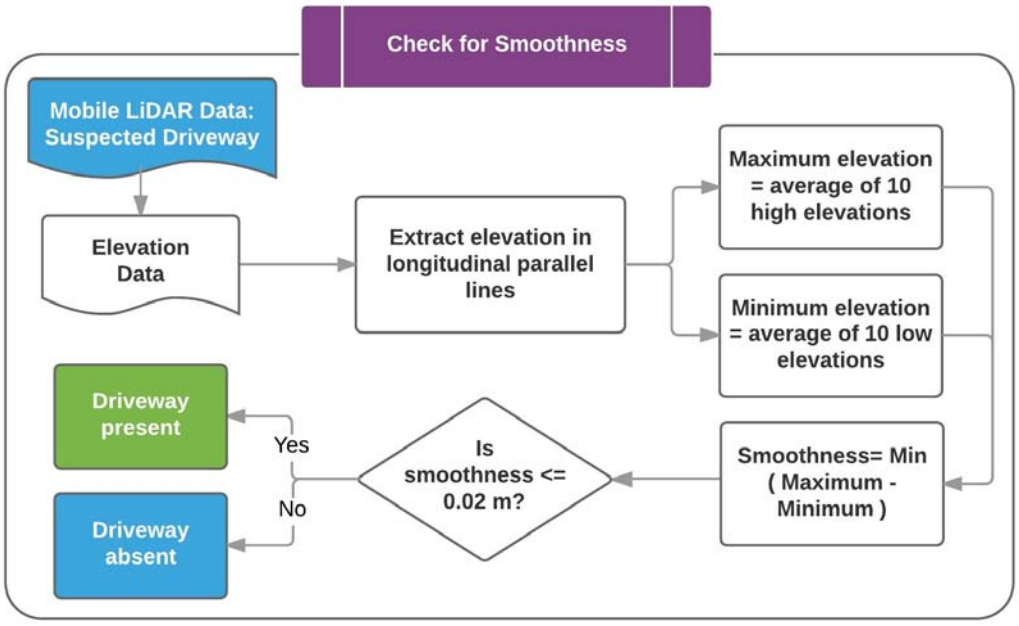


Figure 31: Verification of driveways based on surface smoothness

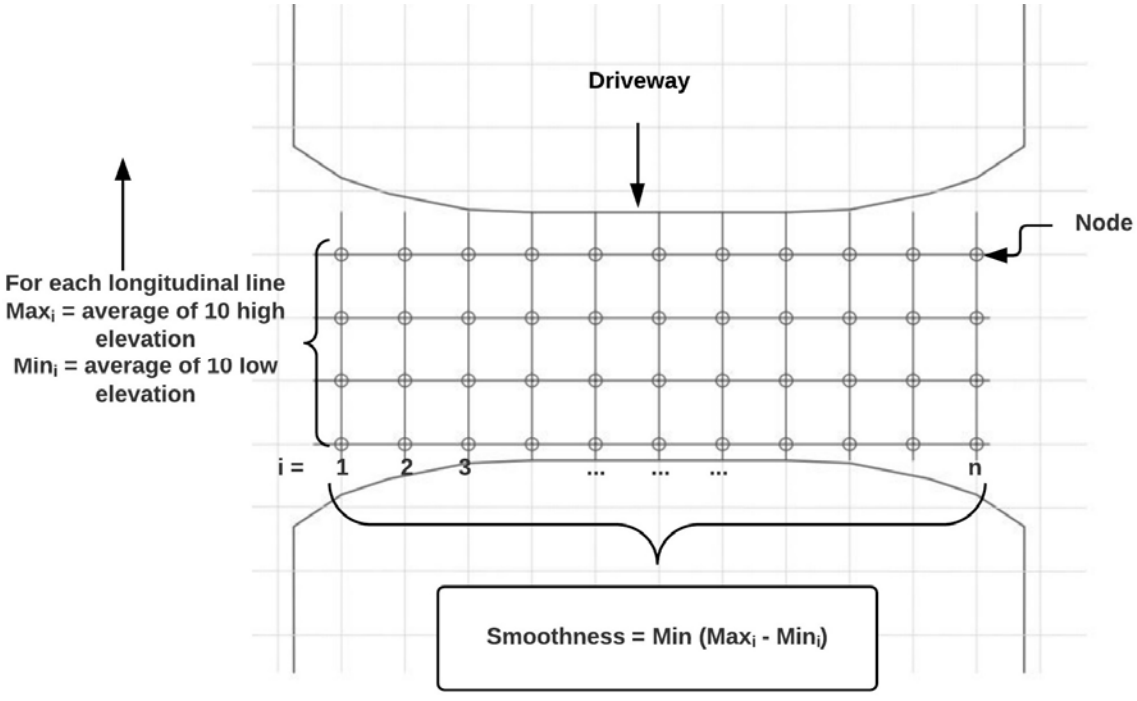


Figure 32: Calculation of driveway smoothness

4.2 Implementation

Implementation of the proposed method requires known (reference) reflectivity data set for grass and pavement materials and a priori information – whether grass driveways are present or absent. If grass driveways are presumed to be absent for a test section, then 'detection of driveways based on material' portion of developed identification method can be used. For the following tests being carried out, it is assumed that driveway material are either asphalt or concrete, but not grass. To test the proposed method a 5,300m section of FM2661 and another 5,000m section of FM320 are used. The FM2661 section has proper, undamaged driveways and intersections with clear separation from adjoining grass patches (see Table 7). On the other hand, the FM320 section has many imperfect dirt and dilapidated driveways unlike the ones used for development (see Table 8). The developed method was tested using these two scenarios. An example of driveway identification is presented in Figure 33. Blue dots in image on the right represent correctly identified driveway.

Table 7: Visual observation of driveways used for testing (FM2661 Section)

Driveway ID	Location Start (m)	Location End (m)	Material	Remark
1	240	254	Concrete	Driveway
2	402	412	Asphalt	Driveway / Adjoining Road
3	454	467	Asphalt	Driveway / Adjoining Road
4	763	794	Concrete	Driveway
5	854	865	Asphalt	Driveway / Adjoining Road
6	1229	1245	Dirt + Asphalt	Driveway
7	1275	1307	Asphalt	Wide Adjoining Road
8	1491	1502	Concrete	Driveway
9	1705	1707	Asphalt	Driveway / Adjoining Road
10	1823	1832	Asphalt	Driveway / Adjoining Road
11	2115	2128	Asphalt	Driveway / Adjoining Road
12	2451	2462	Concrete	Driveway
13	2701	2712	Asphalt	Driveway / Adjoining Road
14	2861	2882	Concrete	Driveway
15	3557	3568	Asphalt	Driveway / Adjoining Road
16	4081	4096	Concrete	Driveway
17	4344	4361	Asphalt	Adjoining Road
18	5201	5242	Asphalt	Wide Adjoining Road
19	5319	5329	Concrete	Driveway
20	5444	5458	Concrete	Driveway

Table 8: Visual observation of driveways used for testing (FM320 Section)

Driveway ID	Location Start (m)	Location End (m)	Material	Remark
1	1132	1137	Dirt (Gravel)	Driveway
2	1170	1176	Asphalt	Driveway
3	1203	1212	Concrete	Driveway
4	1248	1257	Dirt (Gravel)	Driveway
5	1287	1299	Dirt	Driveway
6	1361	1371	Dirt + Grass + Asphalt	Driveway
7	1442	1456	Dirt + Concrete	Driveway
8	1488	1497	Asphalt + Gravel + Dirt	Driveway
9	1599	1605	Dirt	Driveway
10	1724	1738	Dirt	Damaged
11	1871	1879	Gravel	Driveway
12	1993	2014	Asphalt	Adjoining Road
13	2499	2508	Concrete	Driveway
14	2601	2616	Concrete	Driveway
15	3025	3035	Dirt + Grass	Driveway
16	3040	3040	Concrete	Driveway
17	3160	3170	Concrete	Driveway
18	3204	3211	Dirt	Driveway
19	3250	3257	Grass + Dirt	Driveway
20	3262	3273	Concrete	Driveway
21	3332	3338	Concrete	Driveway
22	3464	3473	Grass + Gravel	Driveway
23	3635	3643	Asphalt	Driveway
24	3763	3773	Concrete	Driveway
25	3793	3802	Concrete	Driveway
26	3937	3944	Concrete	Driveway
27	3959	3970	Grass + Gravel	Driveway
28	4253	4262	Grass + Dirt	Driveway
29	4274	4282	Grass + Dirt	Driveway
30	4454	4484	Asphalt	Adjoining Road
31	4589	4598	Grass + Dirt	Driveway
32	4844	4852	Grass + Dirt	Driveway

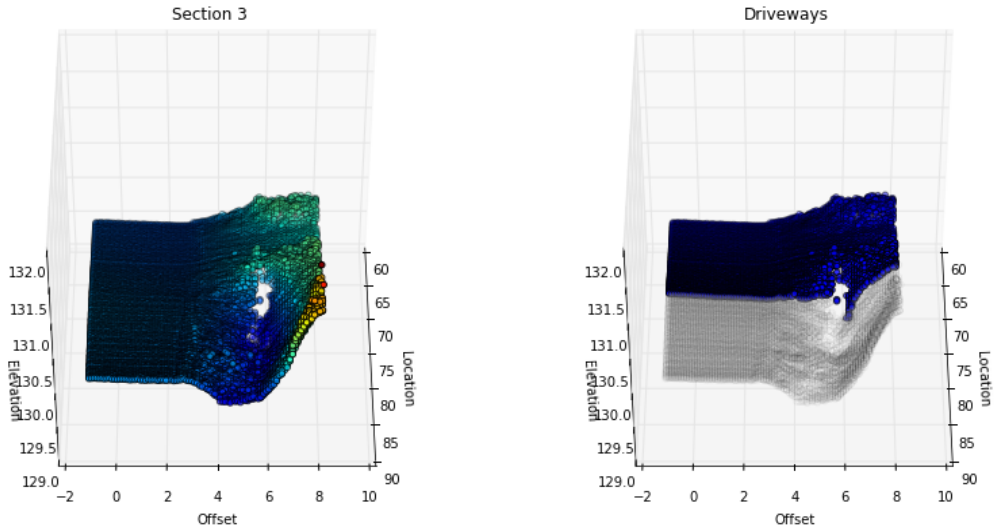


Figure 33: An example driveway identified by the developed method (FM320 Section). Left: Elevation, Right: Distance along road and offset from MLS vehicle.

4.3 Results

Table 9 summarizes the results of this identification method for driveways on FM2661 test sections. These results include estimated width of underlying pipes, start and end elevation difference, and driveway material type. All the 20 driveways (and adjoining roads) present in this test section were correctly identified. One asphalt driveway covered with dirt was misidentified as concrete. Larger elevation difference and small width might indicate scouring at each end. This could not be verified using MLS videos.

Table 10 presents the results of this identification method for driveways on FM320 test sections. Out of 32 driveways present, 22 were correctly detected by the developed method. Additionally, two false positives were found. Most of the driveways in this section have grass overgrowth, dirt and gravel cover on the driveways driving the reflectivity values to be between that of concrete and grass. Careful study of MLS video

revealed lack of proper ditch in many areas of this test section. The algorithm searches for ditch and continuous absence of ditch over a length is postulated as suspected driveways. Lack of proper ditch, therefore, aided in reduced accuracy in identification of driveways and underlying pipes.

Table 9: Results of driveways identification (FM2661)

Driveway ID	**Location Start (m)**	**Location End (m)**	**Width (m)**	**Elevation Difference (m) (ΔE)**	**Material**	**Driveway Presence Correctly Identified?**
1	245.91	251.51	5.6	0.24	Concrete	YES
2	401.81	414.01	12.2	-0.30	Asphalt	YES
3	454.81	464.61	9.8	-0.08	Asphalt	YES
4	776.71	790.11	13.4	-0.08	Concrete	YES
5	850.71	865.21	14.5	-0.17	Asphalt	YES
6	1234.71	1250.01	15.3	-0.07	Concrete	YES
7	1276.91	1306.41	29.5	0.33	Asphalt	YES
8	1489.31	1503.01	13.7	-0.06	Concrete	YES
9	1705.31	1716.81	11.5	-0.06	Asphalt	YES
10	1824.21	1832.01	7.8	0.20	Asphalt	YES
11	2116.61	2127.71	11.1	0.67	Asphalt	YES
12	2454.31	2459.21	4.9	-0.12	Concrete	YES
13	2699.01	2716.01	17.0	-0.68	Asphalt	YES
14	2862.31	2881.51	19.2	-0.27	Concrete	YES
15	3554.31	3569.31	15.0	-0.35	Asphalt	YES
16	4085.81	4097.91	12.1	0.72	Concrete	YES
17	4344.71	4362.01	17.3	0.26	Concrete	YES
18	5203.31	5239.41	36.1	-0.85	Asphalt	YES
19	5320.21	5330.71	10.5	-0.06	Concrete	YES
20	5443.91	5456.31	12.4	-0.02	Concrete	YES

Table 10: Driveways identification test results: FM320 Section

Driveway ID	Location Start (m)	Location End (m)	Width (m)	Elevation Difference (m)	Material	Driveway Presence Correctly Identified?
1	1133.91	1140.31	6.4	0.15	Concrete	YES
2	1172.01	1178.01	6.0	0.11	Asphalt	YES
3	1204.81	1211.01	6.2	-0.12	Concrete	YES
4	1231.81	1236.31	4.5	0.17	Concrete	NO
5	1250.01	1261.31	11.3	-0.23	Concrete	YES
6	1364.21	1370.51	6.3	0.23	Concrete	YES
7	1451.51	1455.21	3.7	0.56	Concrete	YES
8	1599.51	1604.31	4.8	0.01	Concrete	YES
9	1872.11	1878.71	6.6	0.30	Concrete	YES
10	1994.51	2014.31	19.8	0.25	Asphalt	YES
11	2015.81	2025.21	9.4	0.28	Concrete	NO
12	2499.91	2507.11	7.2	0.06	Concrete	YES
13	2602.21	2616.11	13.9	0.05	Asphalt	YES
14	3028.41	3031.61	3.2	-0.03	Concrete	YES
15	3160.21	3173.21	13.0	-0.24	Concrete	YES
16	3205.61	3210.91	5.3	-0.07	Concrete	YES
17	3334.11	3337.71	3.6	0.14	Concrete	YES
18	3469.21	3472.81	3.6	0.45	Concrete	YES
19	3637.51	3643.31	5.8	0.28	Asphalt	YES
20	3766.81	3772.81	6.0	0.31	Concrete	YES
21	3793.51	3802.41	8.9	0.13	Concrete	YES
22	4452.81	4486.21	5.4	0.15	Asphalt	YES
23	4586.91	4597.41	10.5	0.00	Concrete	YES
24	4848.01	4853.91	5.9	-0.36	Concrete	YES

5 DEVELOPMENT OF ROADSIDE FEATURES IDENTIFICATION METHOD

This section deals with development of methods for detecting grass areas and curbs from LiDAR data. Further relevant properties like grass condition, height and location of detected curbs are calculated. Sections of road from FM95, FM320 and FM2661 are used for developing the grass area extraction. Two sections from George Bush Dr. and FM320 are used for developing the curb detection algorithm.

Developed algorithm was converted into multiple codes and subroutines for testing and verification. Table 11 summarizes features of roadway sections used for testing the developed algorithms.

Table 11: Description of actual sections used for testing roadside feature identification (grass area extraction, grass condition, presence of curb, curb height, and curb location)

Section	Pavement Type	Length (m)	Curbs	Ditches	Driveways (Intersection)
FM320 SB 1	Asphalt	1,027	Absent	Present	Present
FM320 SB 2	Asphalt	10,729	Absent	Present	Present
FM95 NB 1	Asphalt	1,026	Absent	Present	Present
FM95 NB 2	Asphalt	1,040	Absent	Present	Present
FM95 NB 3	Asphalt	1,060	Absent	Present	Present
FM2661 WB 1	Asphalt	1,985	Absent	Present	Present
GB WBOL 1	Concrete	815	Present	Absent	Present
GB WBOL 2	Concrete	815	Present	Absent	Present
GB EBOL 1	Concrete	932	Present	Absent	Present
GB EBOL 2	Concrete	933	Present	Absent	Present
PBY NB 1	Concrete	616	Present	Absent	Present

5.1 Grass Area Extraction

5.1.1 Formulation

Figure 34 depicts typical distribution of reflectivity values as we move from 1.5m from the MLS center toward the roadside grass area. Clear change in distribution is observed across the pavement edge. Therefore, it is possible to distinguish and extract grass areas from pavement material based on reflectivity. Subsequent detection methodology is based on this observation to a large extent.

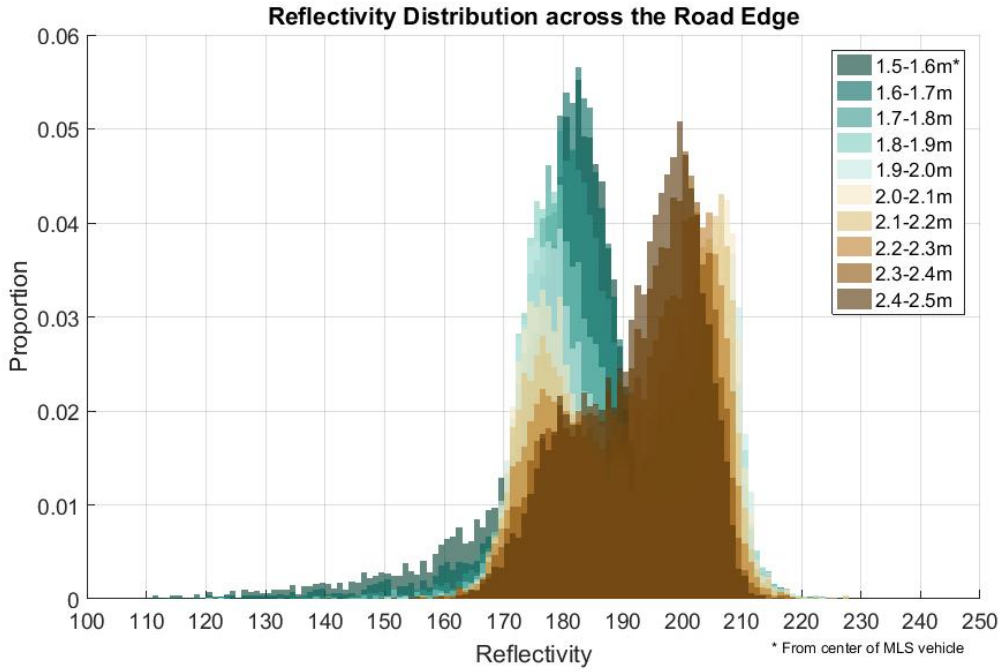


Figure 34: Change in reflectivity distribution across road edge.

5.1.1.1 Filtering of Reflectivity Values

Figure 35 depicts distribution of reflectivity values for a section of FM320. Ordinate axis corresponds to perpendicular distance from direction of travel, 0 being the center of the MLS vehicle. The blue dots in the reflectivity distribution represent grid points without reflectivity values. In general, density of collected data points decrease as we move away from the MLS vehicle toward the roadside. This is a problem when trying to extract grass area and classify grass condition. In addition, local variation and anomaly is observed in raw reflectivity values. In order to extract and identify aggregate features from reflectivity intensities, it is important to apply a filtering technique such that anomalies are accounted for. After carrying out multiple trials, a filtering technique was developed as presented in Figure 36.

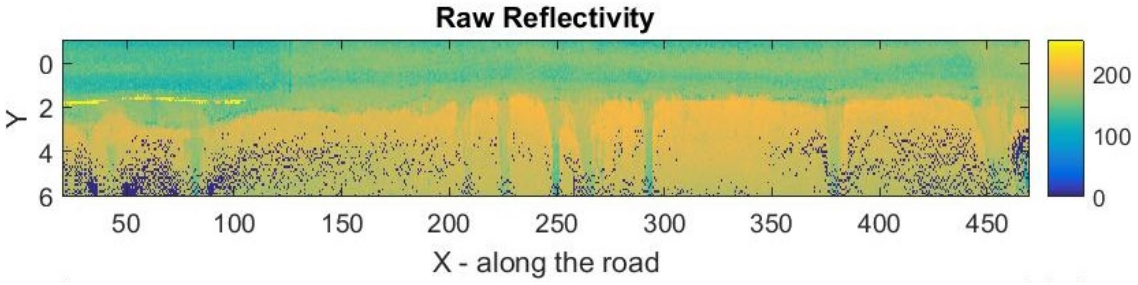


Figure 35: Distribution of reflectivity for a road section.

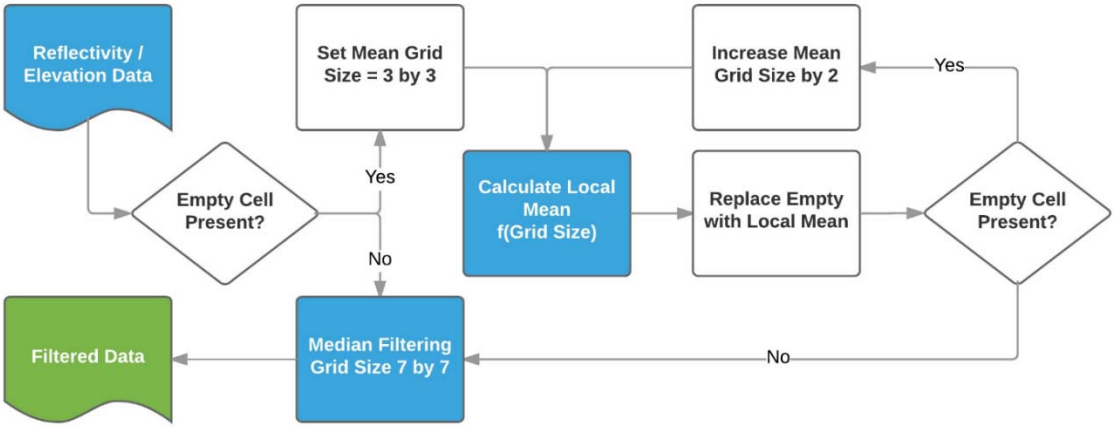


Figure 36: Developed filtering technique for reflectivity and elevation distribution

Empty cells, if present, are replaced with local mean. Initial grid size for local mean computation is 3 by 3. If there are any more empty cells, mean grid size is increased by 2 units (5 by 5, 7 by 7 and so on) until there are no empty cells. Median filtering is then applied to remove anomalies. A pictorial representation of the filtering process is given in Figure 37. Reflectivity distribution after application of attenuation correction on the filtered reflectivity is given in Figure 38.

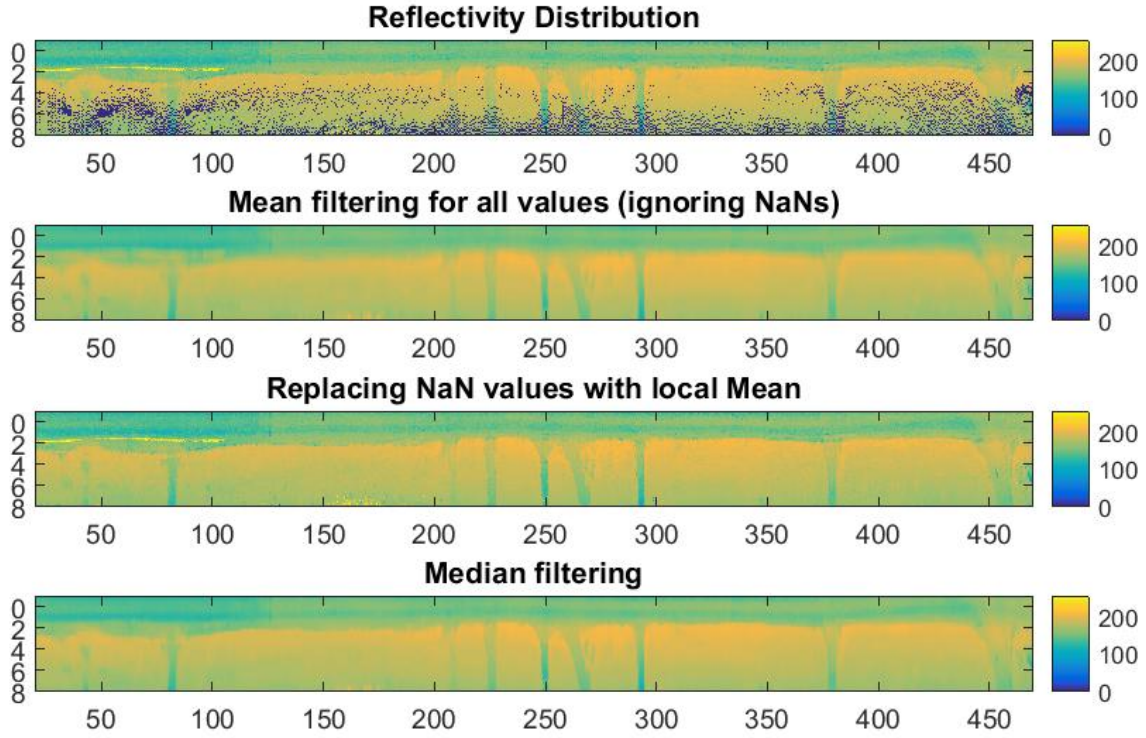


Figure 37: Application of filtering technique

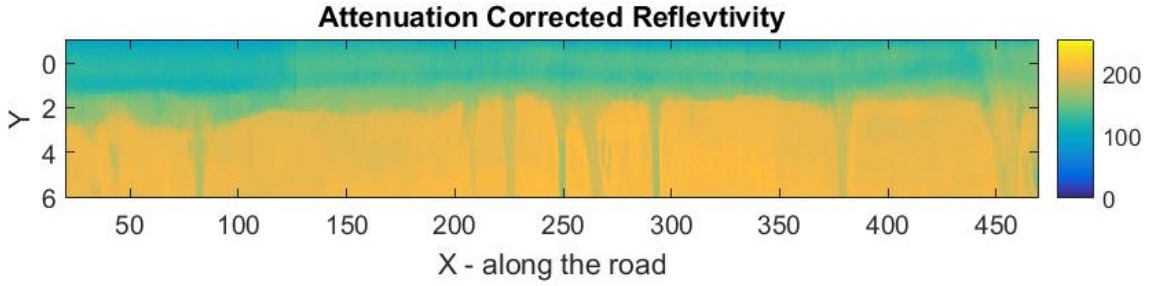


Figure 38: Reflectivity distribution after attenuation correction

5.1.1.2 Statistical Parameters for Extraction

Statistical approach is implemented for the extraction of grass area. As presented earlier in Figure 26 (Section 3.3.1), to be able to detect roadside grass for asphalt roads reliably (80% or higher confidence), the reflectivity value should be greater than 187.69.

For concert roads (Figure 27), the reflectivity has to be higher than 200 for the same confidence level. However, the upper limit of grass reflectivity is set to 213.99; which corresponds to nearly 100 confidence for both asphalt and concrete roads.

Steps involved in extraction of grass area for the method developed are presented in Figure 39. Figure 40 again highlights the importance of filtering and attenuation correction in extraction of grass area.

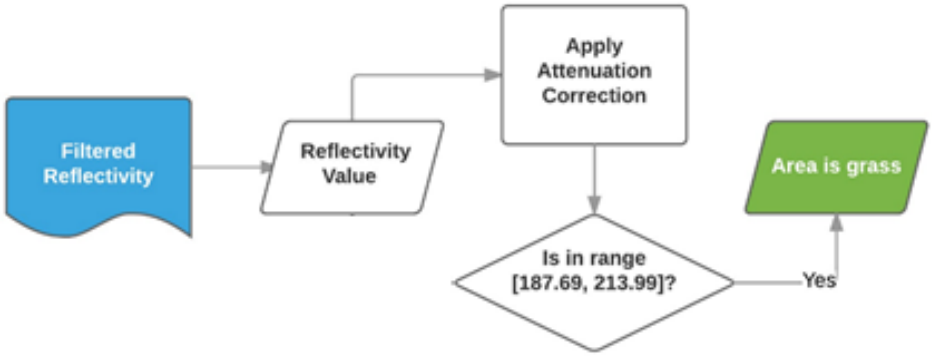


Figure 39: Steps involved in extraction of grass area

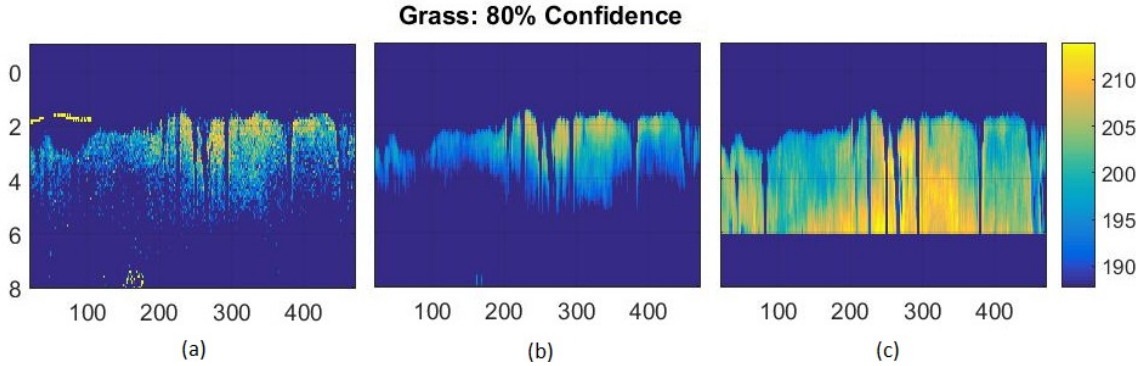


Figure 40: Road side grass extracted with 80% confidence for asphalt road using (a) raw reflectivity, (b) filtered reflectivity and (c) attenuation corrected reflectivity

5.1.1.3 Classification of Grass Condition

After grass area is extracted, areas with reflectivity closer to mean of grass distribution is considered as a 'good' grass and those further away in either direction are considered as a 'poor' grass section (Figure 26). Reflectivity values at the lower tail of the distribution tend to represent low-density grass (i.e., grass with bare spots). On the other hand, reflectivity values at the upper tail of the distribution tend to represent areas with water ponding. Based on a trial and error process for a section of FM320 (length = 450m), values within the central 68.2% of the normal distribution (i.e. 1 standard deviation on either direction of mean for a normal distribution) were considered to be representative of good grass as confirmed by the MLS video. A summary of these trials is presented in Table 12. Figure 41 shows the area of good grass as identified by this method. Steps involved in the method developed for grass type classification are presented in Figure 42 below.

Table 12: Selection of definition for 'good grass' based on central confidence interval.

Trial No.	**Central confidence interval**	**Reflectivity range for good grass**	**Area: Good grass (sq. m)**	**Area: Poor grass (sq. m)**
1	68.2%	[192.22, 208.68]	1357.21	339.6
2	50%	[194.91, 206.00]	1035.78	661.03
3	40%	[198.37, 202.54]	451.1	1245.71

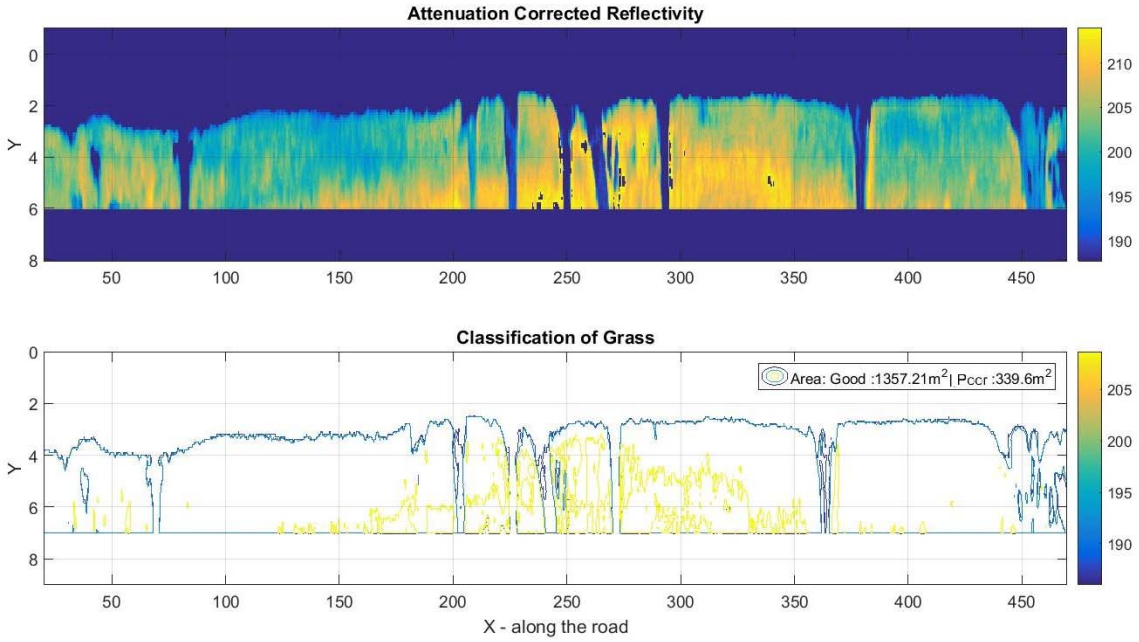


Figure 41: Test for selection of 'good grass' definition corresponding to 68.2% central confidence level.

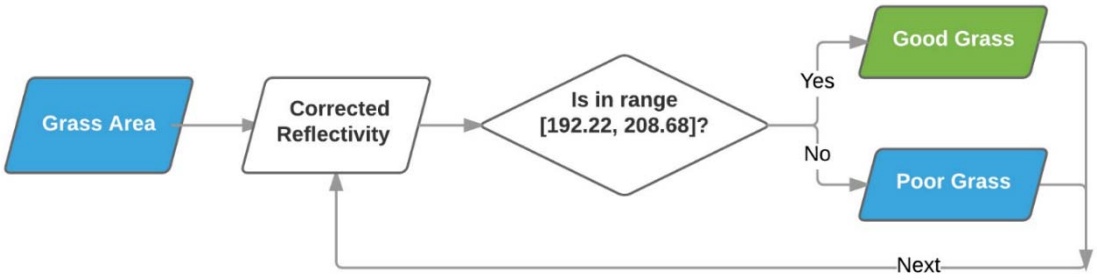


Figure 42: Steps involved in grass type classification method.

5.1.2 Results

Summary of tests carried out for extraction of grass area on road sections listed in Table 11 is presented in Table 13. Graphical representation of extracted area is presented in Appendix C.

For asphalt roads (FM320, FM95, and FM2661), 16.83% of the grass area is identified as poor grass in FM320 SB 1. This includes areas containing foreign objects, dirt, and standing water. For FM320 SB 2, 9.68% of the grass area is identified as poor grass area. Upon visual inspection of the MLS video, this part of FM320 is indeed found to have relatively small ratio of poor grass cover. Similarly, 6.38%, 4.76% and 6.75% of grass area in FM95 NB 1, 2 and 3 respectively were identified as poor grass areas. FM2661 was estimated to have 12.37% poor grass area. Inspection of MLS video corroborated this estimate.

Tests carried out on concrete road sections with side-walks and grass patches reveal larger estimate of poor grass area. This is attributed to difficulty in accurately separating grass areas from concrete side-walks based on reflectivity values. As such much of the side walk area was identified as poor grass, i.e. 42.86%, 50.81%, 35.82% and 42.2% for GB WBOL 1, GB WBOL 2, GB EBOL 1 and GB EBOL 2, respectively. Similarly, 52.21% of area is estimated to be poor grass for the final concrete road section with concrete sidewalk, namely, PBY NB 1. An example test section is presented in Figure 43.

Table 13: Grass detection test results

	Grass Parameters		
Test No.	**Section**	**Good Grass Area (sq. m)**	**Poor Grass Area (sq. m)**
1	FM320 SB 1	3,132.73	633.83
2	FM320 SB 2	38,417.89	4,119.25
3	FM95 NB 1	3,642.63	248.25
4	FM95 NB 2	3,733.99	186.53
5	FM95 NB 3	3,852.20	279.06
6	FM2661 WB 1	4,888.80	604.92
7	GB WBOL 1	1,860.48	1,395.69
8	GB WBOL 2	1,233.34	1,273.98
9	GB EBOL 1	2,372.76	1,324.39
10	GB EBOL 2	2,000.40	1,460.26
11	PBY NB 1	892.83	975.44

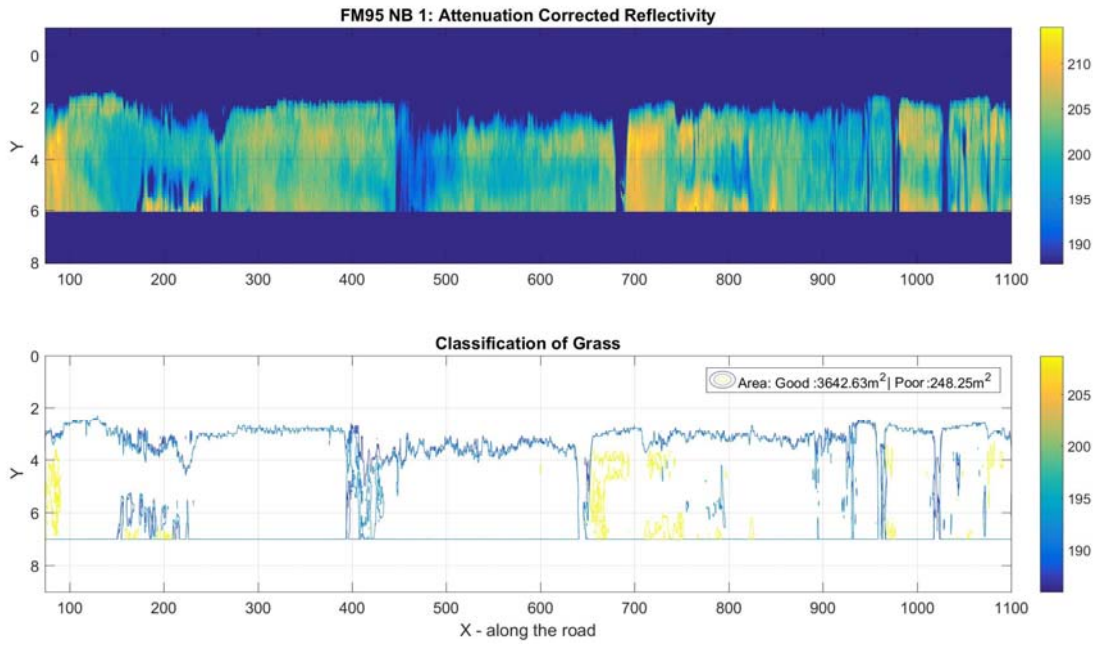


Figure 43: An example of grass detection test: FM95 North Bound Section 1

5.2 Curb Detection

5.2.1 Formulation

A curb is usually a stone or concrete edging bounding a road or a path. For a road section with curbs, a gutter is usually provided along the curb to drain out the pavement. Roads without side ditches are usually provided with a curb and gutter. This section describes the development of a method for detecting the presence of curb and calculating its height and location. A section of George Bush Dr. (with curbs) and a section of FM320 (without curbs) are used for the development of this algorithm. Both sections are 450m in length and both contain multiple driveways and intersections.

The developed method uses a filter template technique; which uses normalized cross correlation coefficients to identify features. Cross-correlation is a measure of similarity of two series as a function of displacement of one relative to other. Cross correlation is also known as sliding inner-product. A sample signal (shape of cross section at the curb) is called a filter. The moving correlation between the filter and a test signal (i.e., shapes of cross sections) is calculated. The location of highest correlation gives the location of best match along the cross section. The formula used for calculating normalized cross-correlation are as follows:

$$\rho_{XY} = \frac{1}{\sigma_X \sigma_Y} E[(X - \mu_X)(Y - \mu_Y)] = \frac{1}{\sigma_X \sigma_Y} \gamma_{XY}$$

where, X, Y = *two random variables with* μ_X, μ_Y *as means (expected value) and* σ_X, σ_Y *as respective standard deviations*

γ_{XY} = *covariance of random variables X and Y*

ρ_{XY} = *correlation between random variables X and Y*

Cross-correlation is calculated using this formula where *X* is a fixed filter signal and *Y* is a test signal of length equal to *X* at location τ. Cross-correlation at location τ is calculated as:

$$\rho_{XY}(\tau) = \frac{1}{\sigma_X \sigma_Y} E[(X - \mu_X)\,(Y_\tau - \mu_{Y\tau})] = \frac{1}{\sigma_X \sigma_Y}\,\gamma_{XY}\,(\tau)$$

Based on this definition, the cross-correlation coefficient is calculated by normalizing the features to unit lengths using the fast normalized cross-correlation method developed by J.P. Lewis (Lewis, 1995).

Location is defined in terms of x, y coordinates, where x is the distance from the beginning of the road section being analyzed and y is the distance from the MLS vehicle center toward the roadside (perpendicular to the direction of traffic). This technique is used in pattern recognition, feature detection, and signal searching (Tsai & Lin, 2003). Both one and two dimensional filters can be used. Although two dimensional filters was tested, single dimensional filter was used in the developed algorithm so that individual cross sections can be examined independently.

In order to create one dimensional filter for curb detection, elevation difference relative to center of MLS vehicle (ΔZ) near to the known curbs from multiple cross sections were averaged (Figure 44 and Figure 45). Filter length was varied from 0.3m to 2.1m. A filter length of 2.1 m was found to produce the most accurate location of curb (Table 14). A correlation threshold of 0.93 was found to be most suitable as it identifies all curbs present in George Bush Dr. with a small number of false positives for FM320.

The 0.93 correlation threshold along with a 2.1m filter were tested on George Bush Dr. and FM320 (Figure 46 and Figure 47).

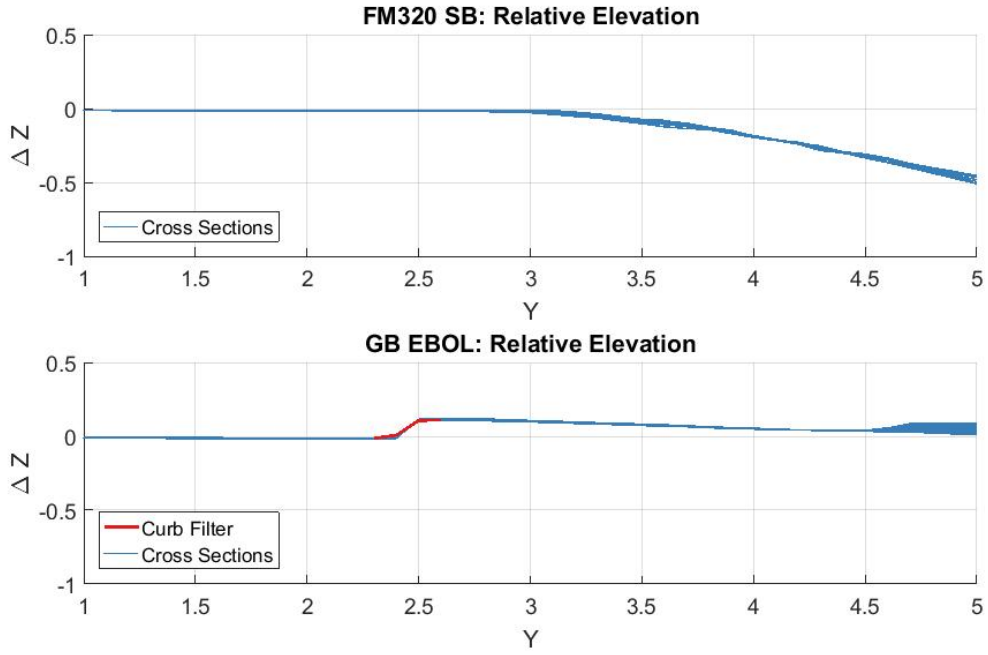


Figure 44: One dimensional filter 0.3m in length (Y=distance from the MLS vehicle toward the roadside, ΔZ= elevation difference relative to center of MLS vehicle)

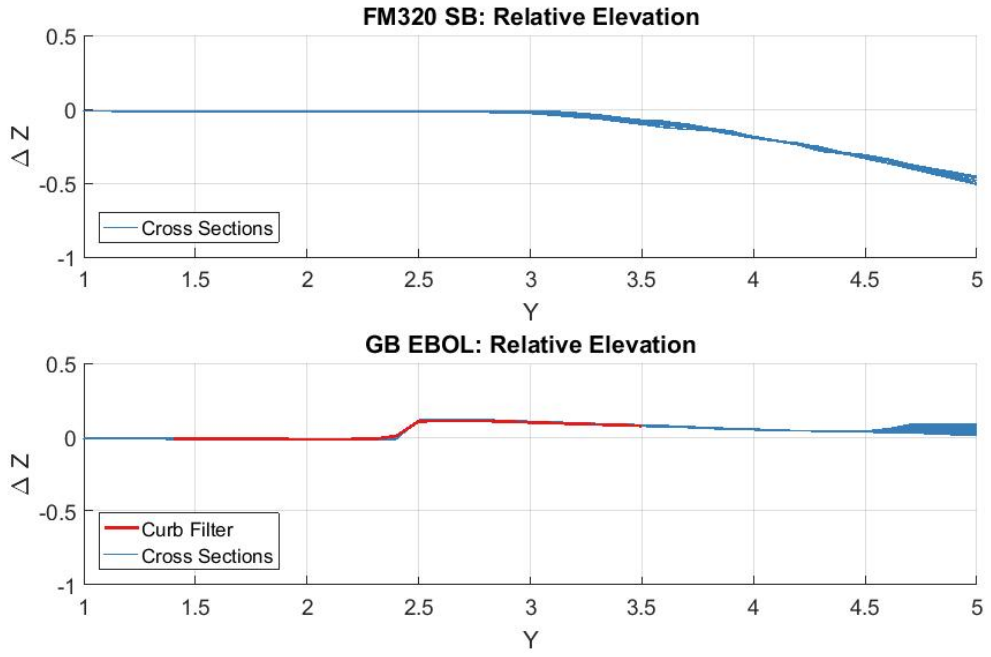


Figure 45: One dimensional filter 2.1m on length (Y=distance from the MLS vehicle toward the roadside, ΔZ= elevation difference relative to center of MLS vehicle)

Table 14: Trials with varying correlation threshold for identification of curbs

Test Section	**Curbs Present in Realty?**	**Correlation Threshold**	**Number of Cross Sections Detected to have Curbs**	**Percentage of Detected Cross Sections with Curbs**
George Bush. Dr	Yes	0.9	2931	65.13%
FM320	No	0.9	203	4.51%
George Bush. Dr	Yes	0.92	2923	64.96%
FM320	No	0.92	135	3.00%
George Bush. Dr	Yes	0.93	2917	64.82%
FM320	No	0.93	109	2.42%
George Bush. Dr	Yes	0.94	2849	63.31%
FM320	No	0.94	76	1.69%
George Bush. Dr	Yes	0.95	2563	56.96%
FM320	No	0.95	43	0.96%

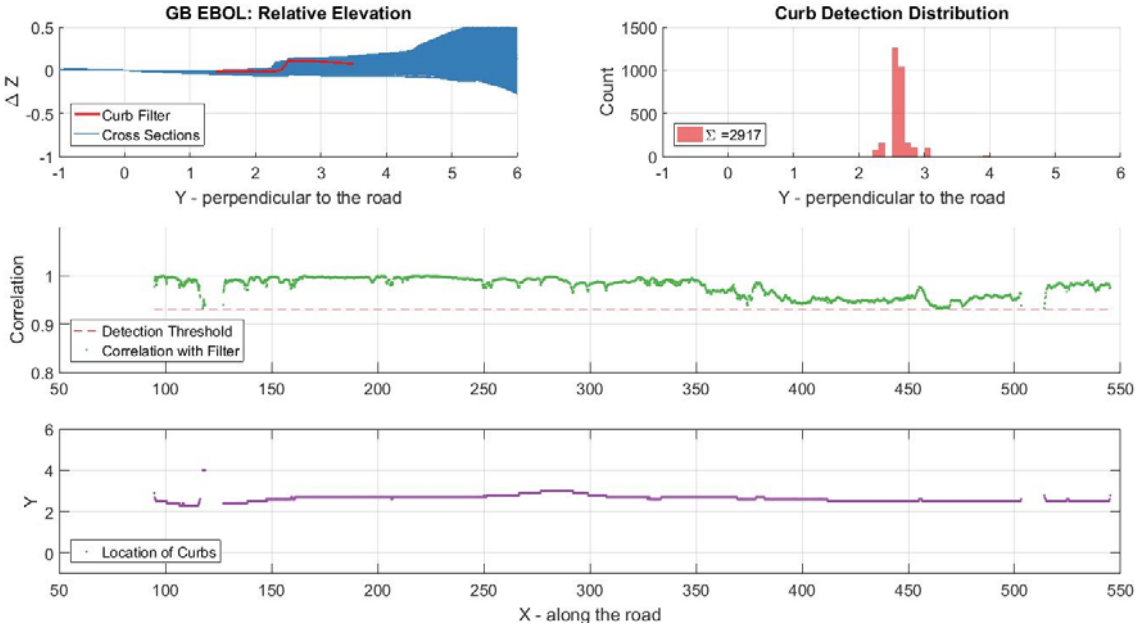


Figure 46: Validity of 0.93 correlation threshold and filter length of 2.1m for George Bush Dr. (curb present)

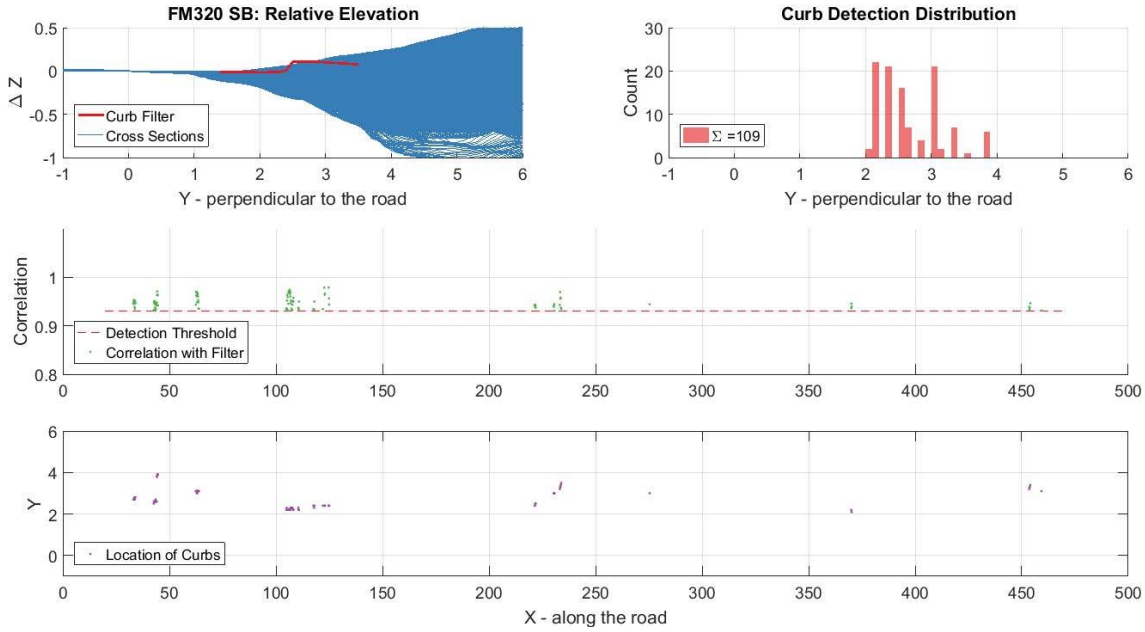


Figure 47: Validity of 0.93 correlation threshold and filter length of 2.1m for FM320 (road without curb)

Figure 48 depicts steps involved in curb detection method developed. Normalized cross correlation calculation is carried out between curb detection filter and rescaled elevation. A search limit is set so that any foreign features far away are not misidentified as curbs. Maximum correlation value for each cross section is compared with correlation threshold defined (0.93). When correlation exceeds this threshold, a curb is detected and the location of maximum correlation gives the transverse location of curb (y coordinate). Longitudinal location (x coordinate) is calculated based on location of the cross section as distance from the beginning of the road section. Maximum elevation difference within two nodes in either direction of the y coordinate (of the detected curb) gives the height of the curb.

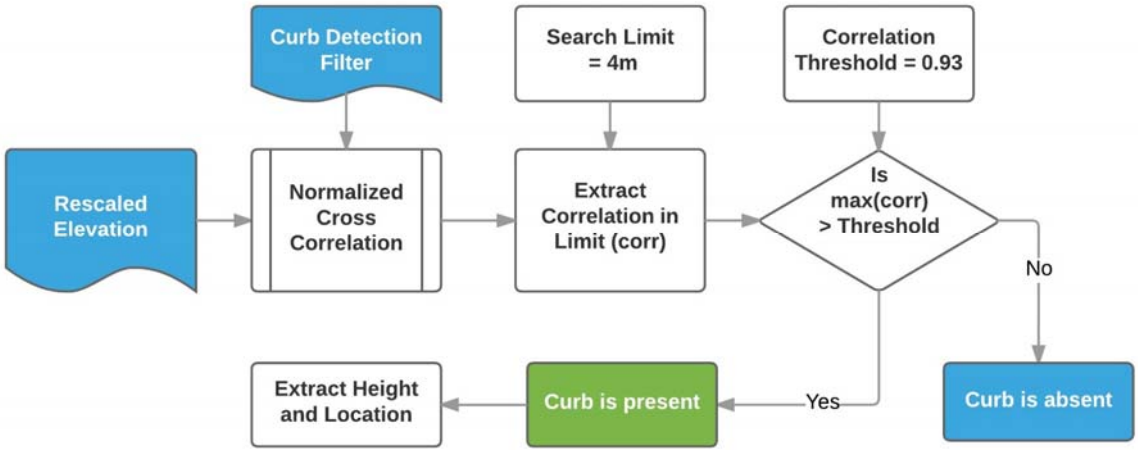


Figure 48: Steps involved in curb detection method.

5.2.2 Results

Results from carrying out the curb detection algorithm on 11 actual road sections (listed earlier in Table 11) are presented in Table 15. For example, the algorithm detected that 93.47% of cross sections on GB WBOL 1 and 78.50% of cross sections GB WBOL 2 have curbs. These estimates have been corroborated by observing MLS videos. The absence of curbs for approximately 7% - 20% of these roads is explained by the presence of multiple intersections. The average height of curb for these locations was found to be 13.5cm. A smaller detection rate was obtained for PBY NB 1 (75.32% of road length). This rates can be explained by the presence of more frequent intersections along this road. Average curb height for PBY NB 1 was found to be higher (18cm). Location of detected curbs (distance from MLS vehicle center in the transverse direction of the road) ranges from 2.5m to 2.7m for George Bush Dr. (GB) sections and 3.3m for Penberthy Road section (PBY NB 1). These values were also verified by visual comparison of the MLS video.

For the sections without curbs the developed method detected false positives for few of the cross sections (e.g., FM320 SB 1 with 2.29% of cross sections detected to have curbs). However, these false positives are random because the algorithm evaluates each cross section independently. Examination of the cross sections around these false positives (Appendix D) indicates that these are in fact random false positives (i.e. surrounding cross sections are true negatives. In all of these cases, average height of falsely detected curbs ranges from 4cm to 8cm (clearly unrealistic values). Also, the location with respective to MLS vehicle center varies widely.

Table 15: Curbs detection test results

Section	Detection Percentage	Average Curb Height (m)	Curb Location* (Modal) (m)	Average Filter Match (Correlation)
FM320 SB 1	2.29%	0.05	2.20	0.949
FM320 SB 2	4.92%	0.08	2.00	0.953
FM95 NB 1	0.94%	0.07	3.20	0.943
FM95 NB 2	2.39%	0.04	2.00	0.947
FM95 NB 3	0.00%	NA	NA	NA
FM2661 WB 1	0.04%	0.06	2.90	0.939
GB WBOL 1	93.47%	0.13	2.60	0.956
GB WBOL 2	78.50%	0.14	2.70	0.964
GB EBOL 1	90.25%	0.13	2.50	0.970
GB EBOL 2	82.04%	0.14	2.50	0.967
PBY NB 1	75.32%	0.18	3.30	0.967

* From center of MLS vehicle

Graphical representation of these results for each road section are presented in Appendix D, depicting distribution of all cross sections along with location and height of curbs when detected. Location in transverse direction is presented with respect to MLS vehicle center. Histogram of transverse location of curb is also presented. Results from two of these test road sections (one with curbs and one without curbs) are presented in Figure 49 and Figure 50, respectively, as examples.

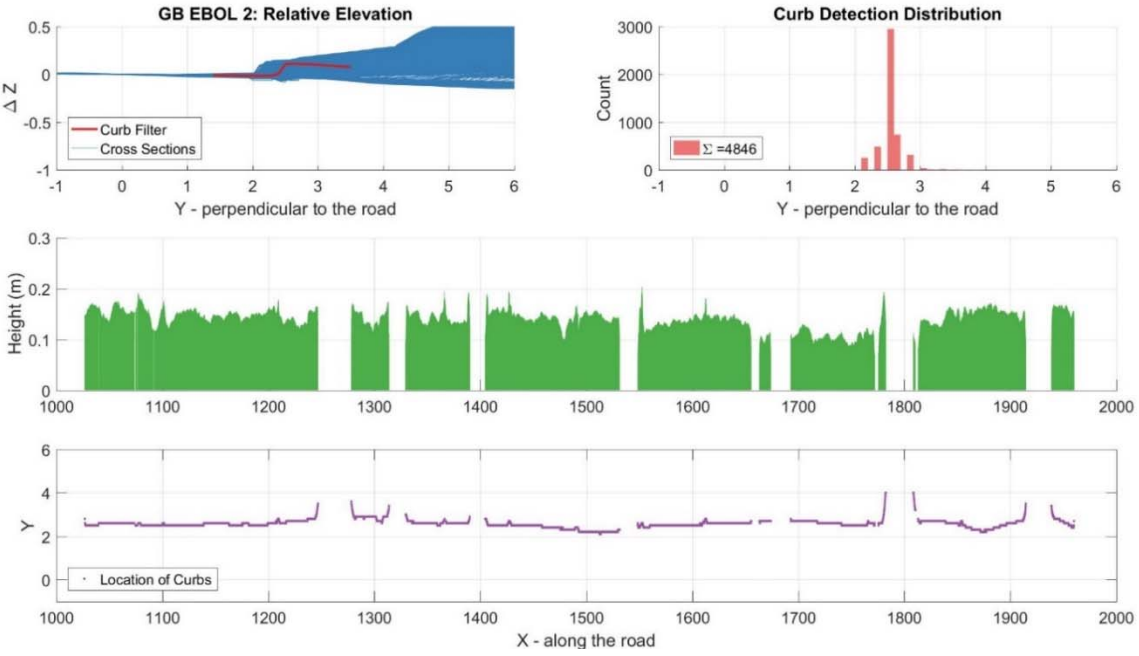


Figure 49: Curb detection results for GB EBOL 2 (road with curb)

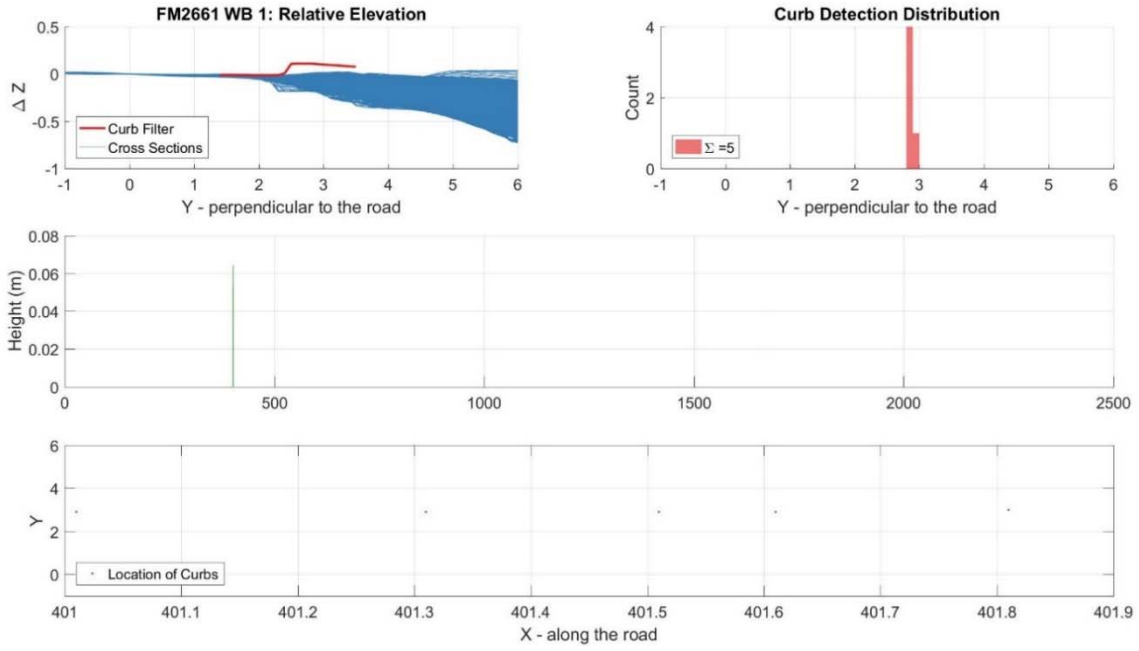


Figure 50: Curb detection results for FM2661 WB 1 (road without curb)

6 SUMMARY OF RESEARCH EFFORTS, CONCLUSIONS, AND RECOMMENDATIONS

The aim of the study was to aid in application of MLSs for roadway inventory and drainage condition assessment. The study attains its objective by developing computational methods for pavement surface type identification, grass detection, driveways and underlying pipe identification, grass area extraction, and curb detection. Each method was tested and validated using data from actual road sections in Texas. The ability to detect aforementioned features reliably using automated means is an initial step to further the cause of MLS acceptance and implementation. Next, the study conclusions and suggested future works are discussed for each method.

6.1 Development of Surface Type Identification Method

6.1.1 Conclusions Related to Pavement Surface Type Detection

- For section-by-section surface type identification, a method based on skewness and closeness to mean of reference reflectivity distributions for known surface types produced the most accurate results. The accuracies of these detections are:
    - Greater than 95% when identifying between asphalt and concrete surfaces.
    - Between 83% and 90% when identifying between concrete, seal coated, dense graded and open graded section.
- Increasing reference data size or test road section length has no significant effect on accuracy. This can be attributed to the homogeneous distribution of reflectivity

for each surface type and the fact that reflectivity intensity is measured using 8-bit discrete values.

- Accuracy increases when adjacent sections are considered. This is most beneficial for network-level applications.

6.1.2 Conclusions Related to Grass Detection

- The effect of reflectivity attenuation is pronounced for roadside. Therefore, in order to develop a reliable grass detection method, attenuation correction is required.
- The developed linear attenuation equation works well for a distance up to 6m from center of MLS vehicle toward the roadside.
- The accuracy of detection is higher for asphalt roads than for concrete roads due to closeness in reflectivity distribution between grass and concrete.
    - When the pavement is asphalt (i.e., the identification is asphalt vs. grass), grass is identified accurately for 98.7% of the samples.
    - When the pavement is concrete (i.e., the identification is concrete vs. grass), 90.3% of grass samples were correctly identified and 16.3% of concrete points were misidentified as grass
- Generalizable graphs were developed (Figure 26 and Figure 27) to obtain the confidence level for any given attenuation-corrected reflectivity value to be grass.

6.1.3 Future Works

- Calibrate the developed methods for reduction in reflectivity due to pavement surface and grass wetness. Also, account for the effect for season on grass. In this research, data were collected in December / January.
- Pavements in good condition were observed to have lesser spread (less kurtosis) in reflectivity distribution compared to cracked pavements (sealed or unsealed cracks) (refer Appendix A). Future work could consider distinguishing between pavement in good condition and pavement in poor condition.
- Reflectivity values can be used in association with imagery data to potentially improve the accuracy of surface type and grass detection.

6.2 Development of Driveways and Underlying Pipes Identification Method

6.2.1 Conclusion

- It is possible to detect driveways and distinguish it from topographical features using a combination of elevation cross sections, material detection, and surface smoothness.
- Accuracy of identification is dependent on the condition of ditches and driveways, as follows:
    - For road sections with well-maintained ditches, paved and well-maintained driveways, the identification accuracy of the developed method is 100% for the test case.

- For road sections with numerous dilapidated and grass/dirt covered driveways, the identification accuracy of the developed method is about 69% for the test case.

6.2.2 Future Works

- The accuracy of driveway width estimation and start - end elevation difference can be further verified using direct field measurements.
- The developed method assumes that pipes are present beneath the driveways and adjoining roads. The method can be improved if a reliable method for detecting pipe inlets and protrusion can replace this assumption.
- Use of filtering technique, as presented in section 5.2 can be explored further to detect driveways and underlying pipes for possible improvement in accuracy for roads with poorly-maintained or grass driveways or roads without ditches.

6.3 Development of Roadside Feature Identification Method

6.3.1 Conclusions Related to Grass Area Extraction

- Filtering to remove local anomaly and variation was necessary to extract any aggregate information from reflectivity distributions. The developed filtering technique - replacing absent data in grid with local mean followed by median filtering - was found to be most suitable.
- Grass area is extracted using reflectivity limits corresponding to 80% confidence for lower end (for asphalt road). Inspection of MLS video showed that the extracted area closely matches the actual grass area for asphalt roads. For concrete roads, the extracted area was less accurate.

- Inspection of MLS video showed that classification of grass type into good and poor based on reflectivity distribution produced reasonable results for roads without sidewalks. For roads with side-walks, much of the side-walks was incorrectly classified as poor grass.

6.3.2 Conclusions Related to Curb Detection

- Presence of curb, along with height and location, can be detected using filter template technique (commonly used in image processing). Specifically, a long filter and a correlation threshold of 0.93 gives best results.
- Using this method, location of intersection and driveways can be identified for road sections with curbs.
- False positives of curb detection for roads without curbs are mostly below 2.5% when each cross section is examined independently.

6.3.3 Future Works

- Independent attenuation correction for range, angle, instrument and material can be explored to further improve accuracy.
- To improve accuracy of extraction for roads with sidewalks, surface smoothness (elevation variation) can be explored and incorporated in the algorithm to differentiate grass from smooth sidewalk surfaces.
- Definition of good and poor grass can be explored further to improve classification precisions.
- Inclusion of imagery data might help to improve the accuracy of grass area extraction.

- The developed methods do not account for the meandering driving path of the MLS. Future work could address this limitation.
- The developed methods could be extended to detect other features such as inlets, pipes, guardrails, and barriers.

REFERENCES

1. Ahmed, M., C. Haas, and R. Haas. Toward low-cost 3D automatic pavement distress surveying: the close range photogrammetry approach. *Canadian Journal of Civil Engineering,* Vol. 38, No. 12, 2011, pp. 1301-1313.

2. Bray, J., B. Verma, X. Li, and W. He. A Neural Network based technique for automatic classification of road cracks. *International Joint Conference on Neural Networks*, Vancouver, Canada, 2006, DOI 10.1109/IJCNN.2006.246782.

3. Gavilan, M., D. Balcones, O. Marcos, D. F. Llorca, M. A. Sotelo, I. Parra, M. Ocana, P. Aliseda, P. Yarza, and A. Amirola. Adaptive Road Crack Detection System by Pavement Classification. *Sensors*, Vol. 11, No. 10, 2011, pp. 9628-9657.

4. Guan, H., J. Li, Y. Yu, and R. Zhai. Iterative Tensor Voting for Pavement Crack Extraction using mobile laser scanning data.. *IEEE Transactions on Geoscience and Remote Sensing,* Vol. 53, No. 3, 2015, pp. 1527-1537.

5. Hans, Z., R. Tenges, S. Hallmark, R. Souleyrette, and S. Pattnaik. Use of LiDAR-Based Elevation Data for Highway Drainage Analysis: A Qualitative Assessment. *Mid-Continet Transportation Research Symposium,* Ames, Iowa, USA, 2003.

6. Herold, M., and D. Roberts. Spectral Characteristics of Asphalt Road Aging and Deterioration: Implications for Remote-Sensing Applications. *Applied Optics*, Vol. 44, No. 20, 2005, pp. 4327-4334.

7. Horkaew, P., S. Puttinaovarat, and K. Khaimook. River Boundary Delineation From Remotely Sensed Imagery Based on SVM and Relaxation Labeling of Water Index and DSM. *Journal of Theoretical and Applied Information* Technology, Vol. 71, No. 3, 2015, pp. 376-386.

8. Jonsson, P., J. Casselgren, and B. Thornberg. Road Surface Status Classification using Spectral Analysis of NIR Camera Images. *IEEE Sensors Journal*, Vol. 15, No. 3, 2015, pp. 1641-1656.

9. Kaseko, M. S., and S. G. Ritchie. A Neural Network-based methodology for pavement crack detection and classification.. *Transpn. Res.-C,* Vol. 1, No. 4, 1993, pp. 275-291.

10. Kashani, A. G., M. J. Olsen, C. E. Parrish, and N. Wilson. A Review of LIDAR Radiometric Processing: From Ad Hoc Intensity Correction to Rigorous Radiometric Calibration. *Sensors (Basel, Switzerland)*, Vol. 15, No. 11, 2015, pp. 28099-28128.

11. Kaufhold, J., and A. Hoogs. Learning to Segment Images Using Region-Based Perceptual Features. *IEEE Computer Vision and Pattern Recognition (CVPR)*, Vol. 2, 2004, pp. 954-961, DOI: 10.1109/CVPR.2004.1315268.

12. Large, A. R. G., and G. L. Heritage. *Laser Scanning for the Environmental Sciences.* Wiley-Blackwell, Chichester, UK, 2009.

13. Lewis, J. P. Fast Template Matching. *Vision Interface 95, Canadian Image Processing and Pattern Recognition Society,* Qubec City, Canada, May 15-19, 1995, pp. 120-123.

14. Mohammadi, M. Road Classification and Condition Determination using Hyperspectral Imagery. *International Archives of Photogrammetry and Remote Sensing*, Vol. 39, 2012, pp. 141-146.

15. Moussa, G., and K. Hussain. A new technique for automatic detection and parameters estimation of pavement crack. *4th International Multi-Conference on Engineering and Technological Innovation*, 2011, DOI: 10.13140/2.1.3191.2001.

16. Mustaffara, M., T. C. Lingb, and O. C. Puanb. Automated Pavement Imaging Program (APIP) for pavement cracks classification and quantification - a photogrammetric approach. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences,* Vol. XXXVII, Part B4, 2008.

17. Nguyen, T. S., M. Aliva, and S. Begot. Automatic detection and classification of defect on road pavement using anisotropy measure.. *European Signal Processing Conference,* Volume Glasgow, United Kingdom, 2009, pp. 617-621.

18. Noronha, V., M. Herold, D. Roberts, and M. Gardner. Spectrometry and Hyperspectral Remote Sensing for Road Centerline Extraction and Evaluation of Pavement Condition. *National Consortium on Remote Sensing in Transportation,* Denver, CO, 2002.

19. Omer, R., and L. Fu. An Automatic Image Recognition System for Winter Road Surface Condition Classification. *International IEEE Conference on Intelligent Transportation Systems*, Funchal, Portugal, 2010, DOI: 10.1109/ITSC.2010.5625290.

20. Omidalizarandi, M., and M. Saadatseresht. Segmentation and Classification of Point CLouds from Dense Aerial Image Matching. *International Journal of Multimedia and Its Applications,* Vol. 5, No. 4, 2013, pp. 33-51.

21. Ong, G. P. R., S. Noureldin, and K. C. Sinha. Automated Pavement Condition Data Collection Quality Control, Quality Assurance, and Reliability. *SPR-3111, Federal Highway Administration*, Washington, DC, 2010.

22. Rajab, I., M. H. Alawi, and M. A. Saif. Application of Image Processing to Measure Road Distresses. *WSEAS Transactions on Information Science & Applications,* Vol. 5, No. 1, 2008, pp. 1-7.

23. Shen, X., and W. Bao. The Remote Sensing Image Matching Algorithm Based on the Normalized Cross-Correlation and SIFT. *Journal of the Indian Society of Remote Sensing,* Vol. 42, No. 2, 2014, pp. 417-442.

24. Sun, L., and Z. Qjan. Multi-scale wavelet transform filtering of non-uniform pavement surface image background for automated pavement distress identification.. *Measurement,* Volume 86, 2016, pp. 26-40.

25. Teomete, E., V. R. Amin, H. Ceylan, and O. Smadi. Digital image processing for pavement distress analyses. *Mid-continent transportation symposium,* Ames, Iowa, USA, 2005.

26. Timm, D. H., and J. M. McQueen. A Study of Manual Vs. Automated Pavement Condition Surveys. *IR-04-01, Alabama Department of Transportation, Highway Research Center*, Auburn University, Auburn, Alabama, 2004.

27. Tsai, D. M., and C. T. Lin. Fast normalized cross correlation for defect detection. *Pattern Recognition Letters,* Vol. 24, No. 15, 2003, pp. 2625-2631.

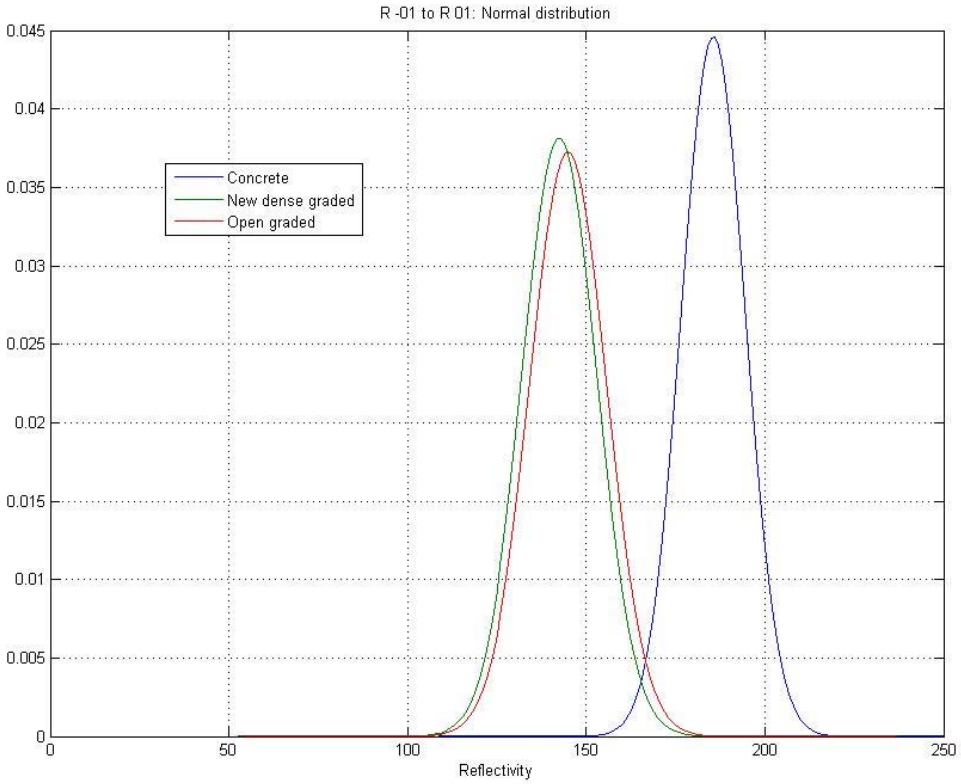
28. Williams, K., M. J. Olsen, G. V. Roe, and C. Glennie. Synthesis of Transportation Applications of Mobile LIDAR. *Remote Sensing*, Vol. 5, No. 9, 2013, pp. 4652-4692.

29. Yoo, H. J., F. Goulette, J. Senpauroca, and G. Lepere. Analysis and Improvement of Laser Terrestrial Mobile Mapping Systems Configurations. *International Archives of Photogrammetry and Remote Sensing*, Vol. 38, No. PART 5, 2010, pp. 633-638.

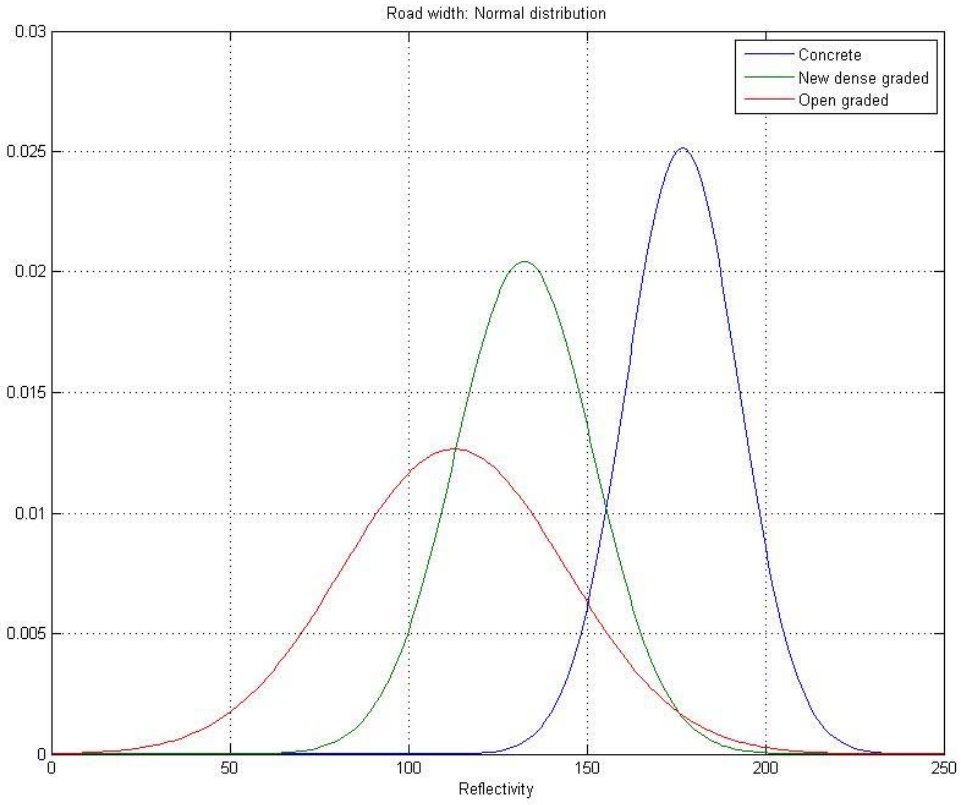
30. Yu, Y., J. Li, H. Guan, and C. Wang. 3D Crack skeleton extraction from Mobile LiDAR point clouds. *International Geoscience and Remote Sensing Symposium (IGARSS),* Quebec, Canada, 2014, DOI: 10.1109/IGRASS.2014.6946574.

APPENDIX A

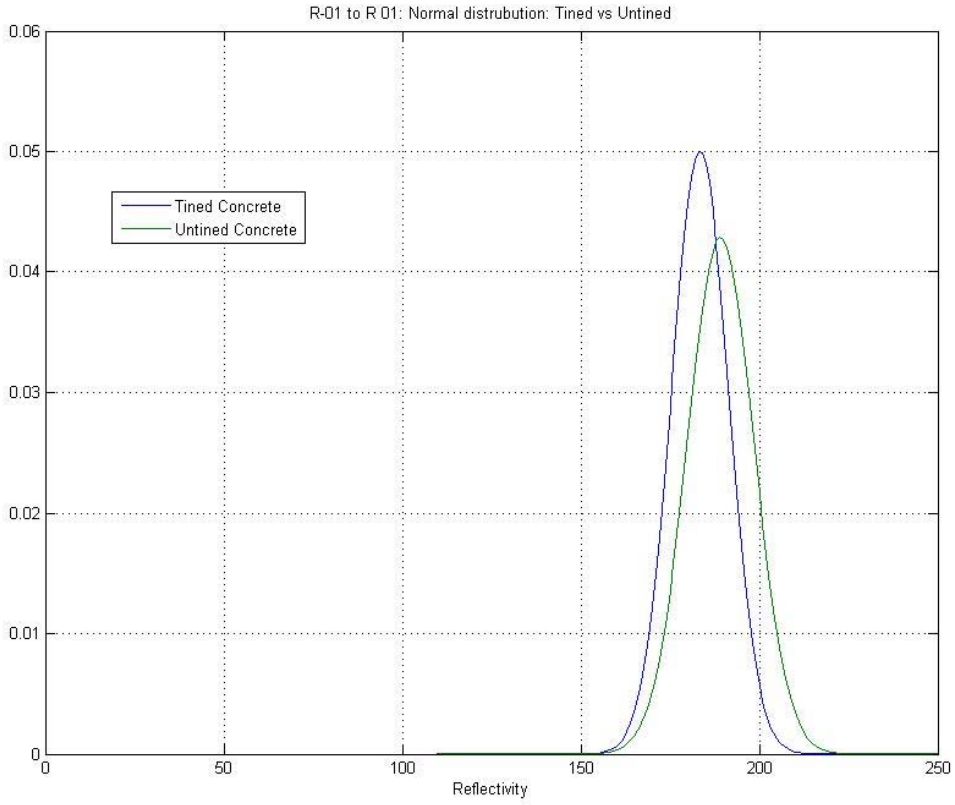
REFLECTIVITY DISTRIBUTIONS



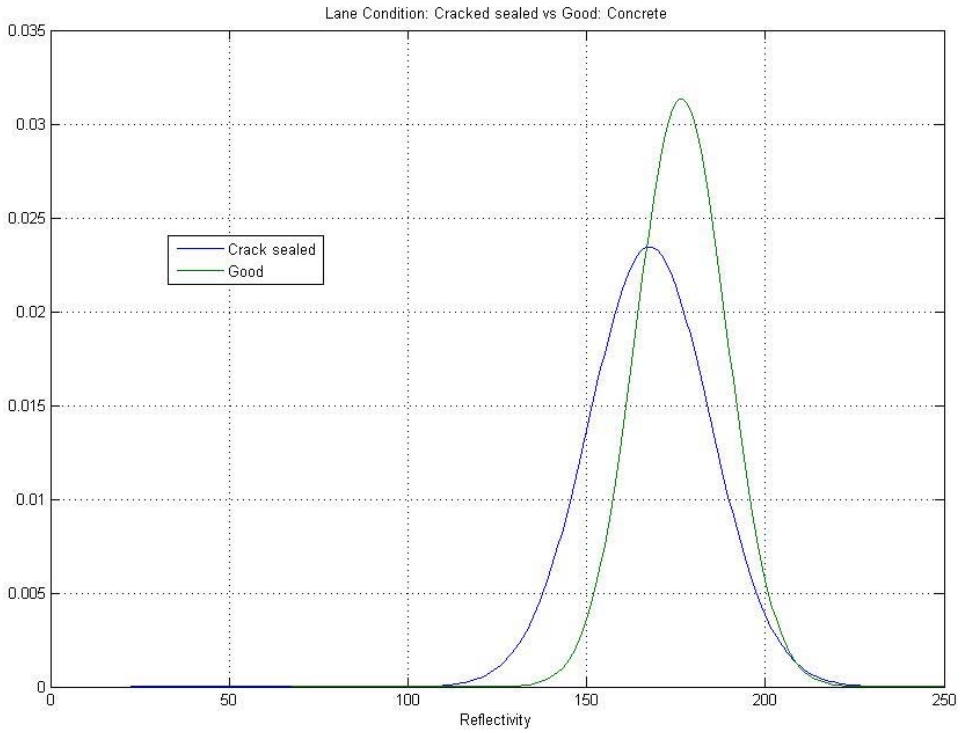
Normalized reflectivity distribution in driven lane of road section



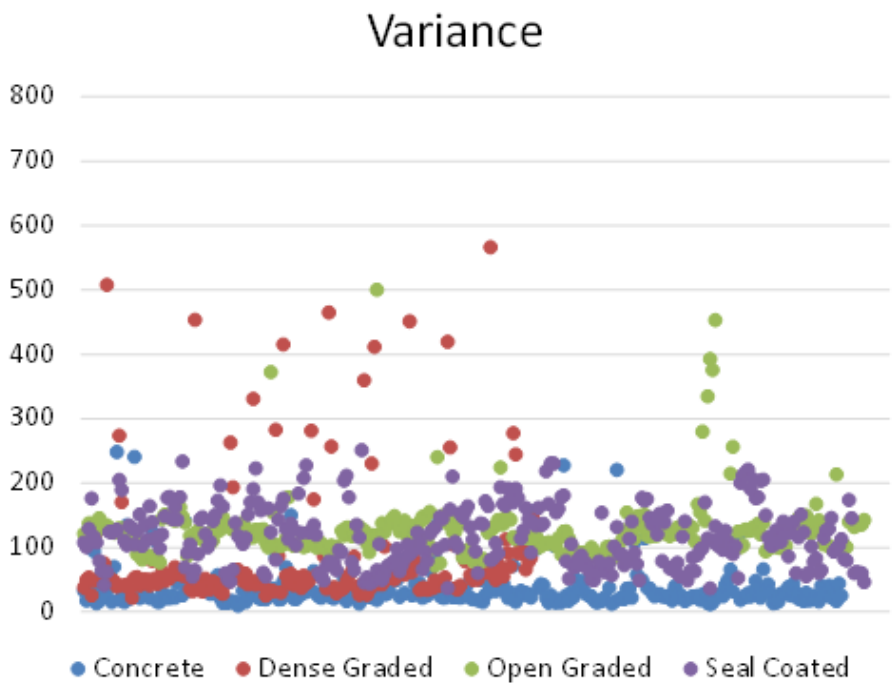
Normalized reflectivity distribution considering whole width of road section:



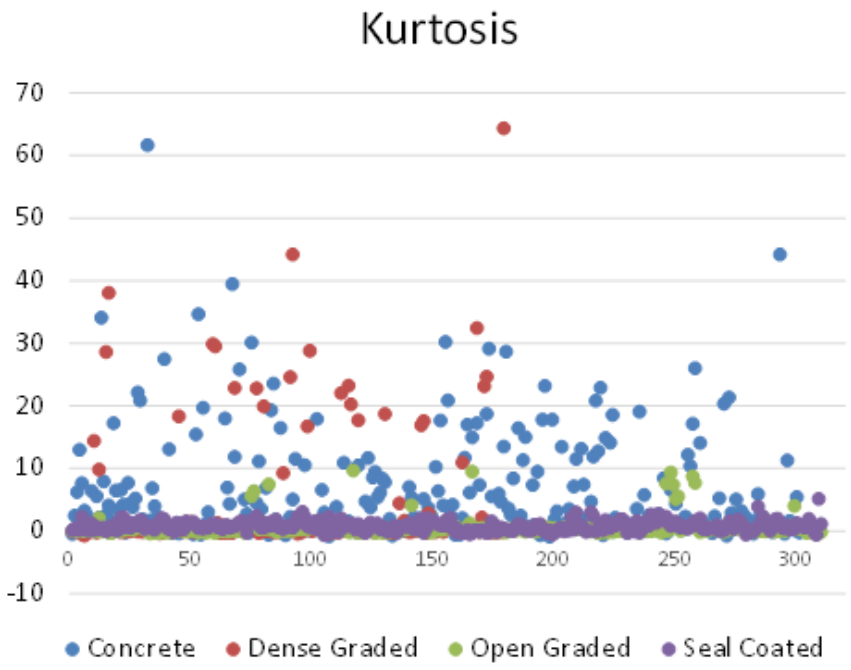
Normalized reflectivity distribution of tined and un-tined concrete



Normalized reflectivity distribution of defect free and crack sealed pavement



Distribution of variance of reflectivity values (unit length: 0.05 miles, 264', 80.47m)



Distribution of kurtosis of reflectivity values (unit length: 0.05 miles, 264', 80.47m)

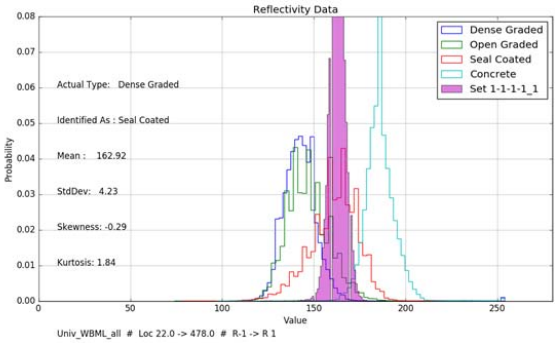
APPENDIX B

SOME TESTS OF DIFFERENT SECTIONS

Some of the road sections not represented by the reference data is tested. The sections considered here have some distinguishing characteristics that is likely to make pavement surface identification difficult. The results of such tests is presented below along with a picture and a brief description of the section in consideration. Identification between the four categories followed by asphalt vs concrete identification is presented on the right. Reflectivity histogram for the test section is also presented.

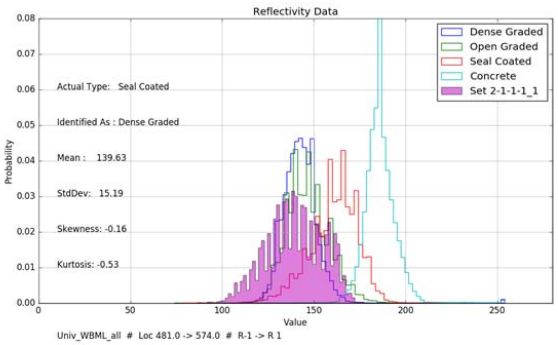
1: Dense Graded

Identification: Seal Coated, Asphalt



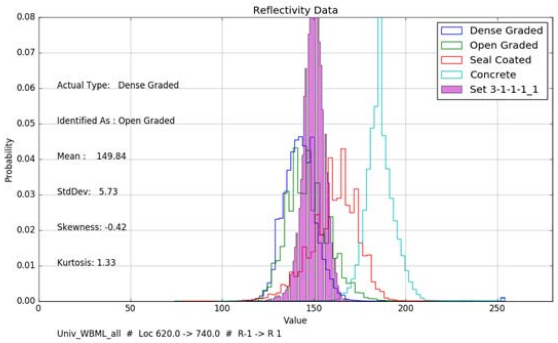
Description: Dense graded asphaltic pavement placed within last 3 years.

2: Seal Coated **Identification: Dense Graded, Asphalt**



Description: Seal coat over dense graded over a bridge deck. Joint for the bridge running transversely in the pavement can be observed. The outside lane shows signs of bleeding in the wheel-paths. This bleeding is somewhat expected due to the breaking condition and acceleration action of stopping and moving vehicles at the signal.

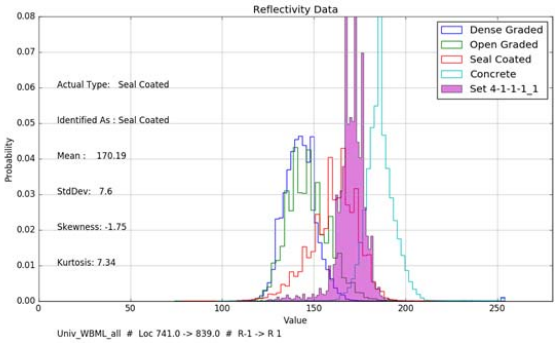
3: Dense Graded **Identification: Open Graded, Asphalt**



Description: Dense graded mix placed within the last 3 years. The small oily spots are a little difficult to explain. They could come from the oil from the seal coat on the other side of the sign. It could also be over-compaction that brought some oil to the surface.

4: Seal Coated

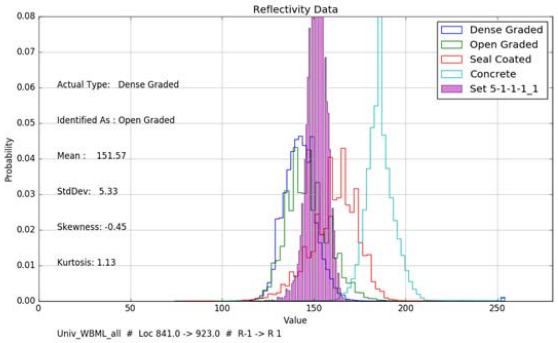
Identification: Seal Coated, Concrete



Description: Seal coated section with a fairly small cover stone. There is also some crack seal present.

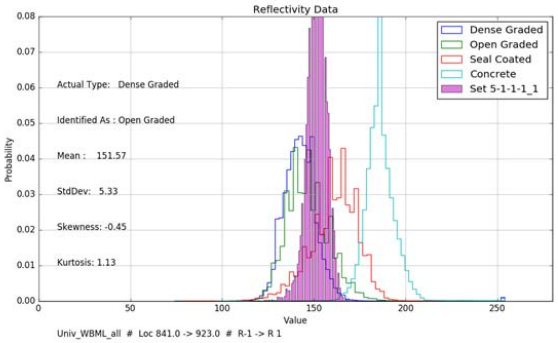
5: Dense Graded

Identification: Open Graded, Asphalt



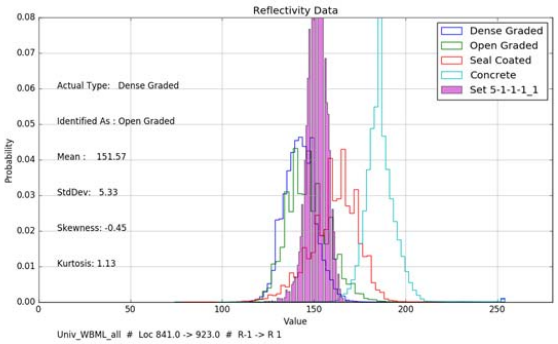
Description: Dense graded section, placed within the last 6 months.

6: Dense Graded **Identification: Open Graded, Asphalt**



Description: The dense graded surface in this section is likely to be more than 10 years old. Some crack sealing is also observed.

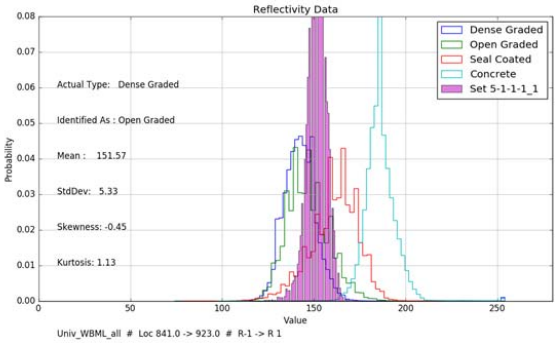
7: Dense Graded **Identification: Open Graded, Asphalt**



Description: Old dense graded section with significant crack seal. The dense graded surface is most likely to be more than 10 years old. With the age of the mix and the compaction in the wheelpaths, more oil could be near the top, creating a reflection difference.

8: Dense Graded

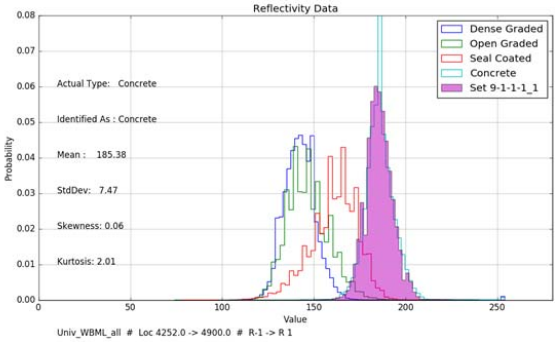
Identification: Open Graded, Asphalt



Description: This section is similar to the previous one.

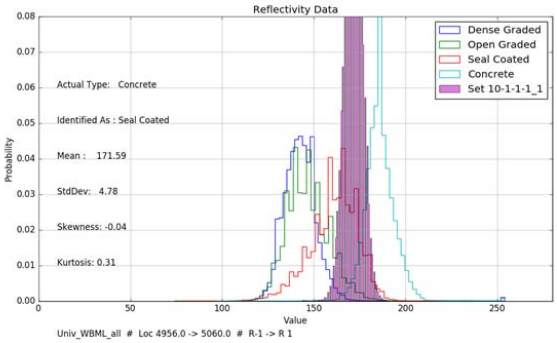
9: Concrete

Identification: Concrete



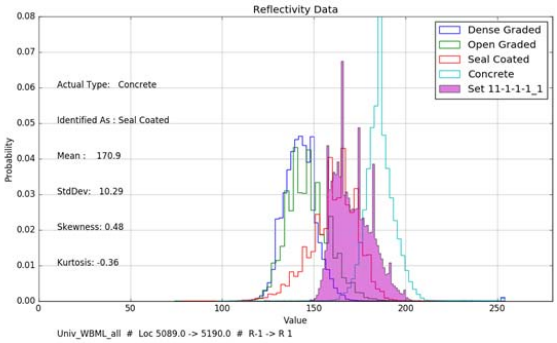
Description: Old concrete pavement (about 15 years) with closely spaced transverse joints.

10: Concrete **Identification: Concrete**



Description: About 15 year old concrete section with very little heavy truck traffic.

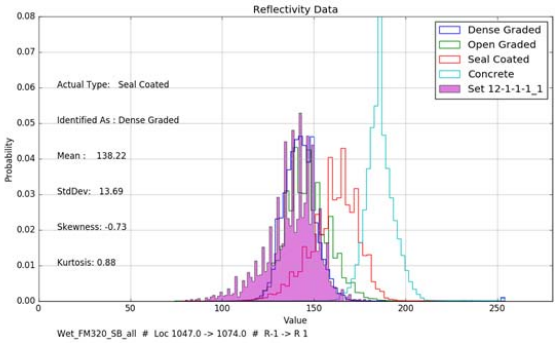
11: Concrete **Identification: Concrete**



Description: This section is similar to the one before.

12: Seal Coated, Wet

Identification: Dense Graded, Asphalt



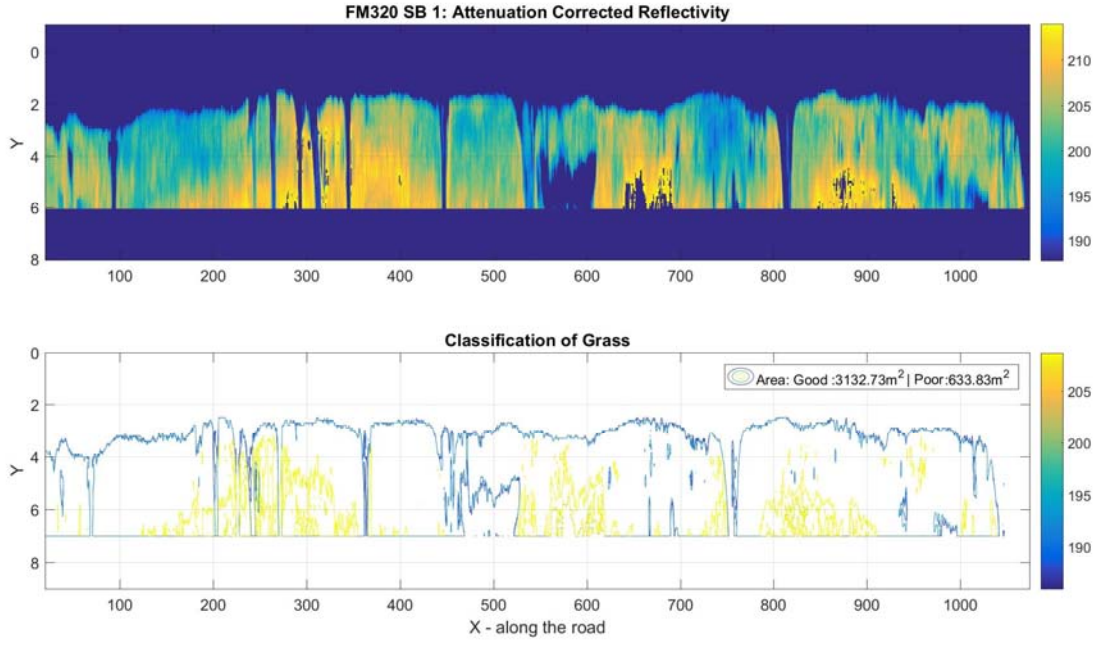
Description: Wet seal coated section. Rock appears to have bean sheared off creating oil spot that is clearly visible.

APPENDIX C

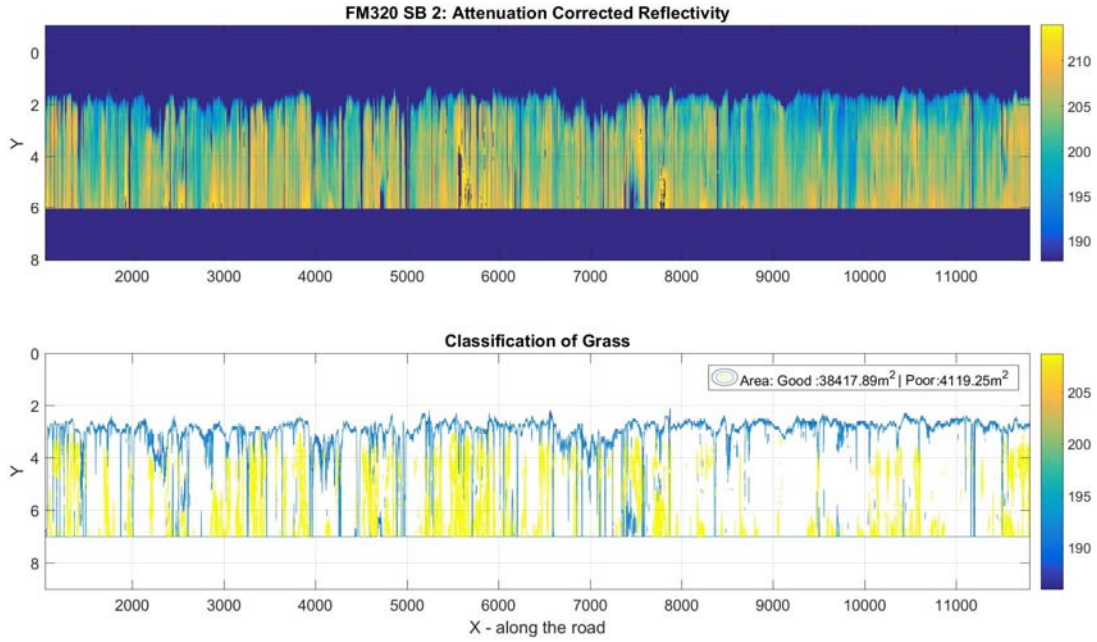
GRASS AREA EXTRACTION TESTS CARRIED OUT

Results of tests carried out for the algorithm developed for grass area extraction is presented in this section. A total of eleven sections were tested out of which six had roadside grass ditch and five had a combination of side walk and grass strips. All the sections had grass area broken by frequent driveways and intersecting roads.

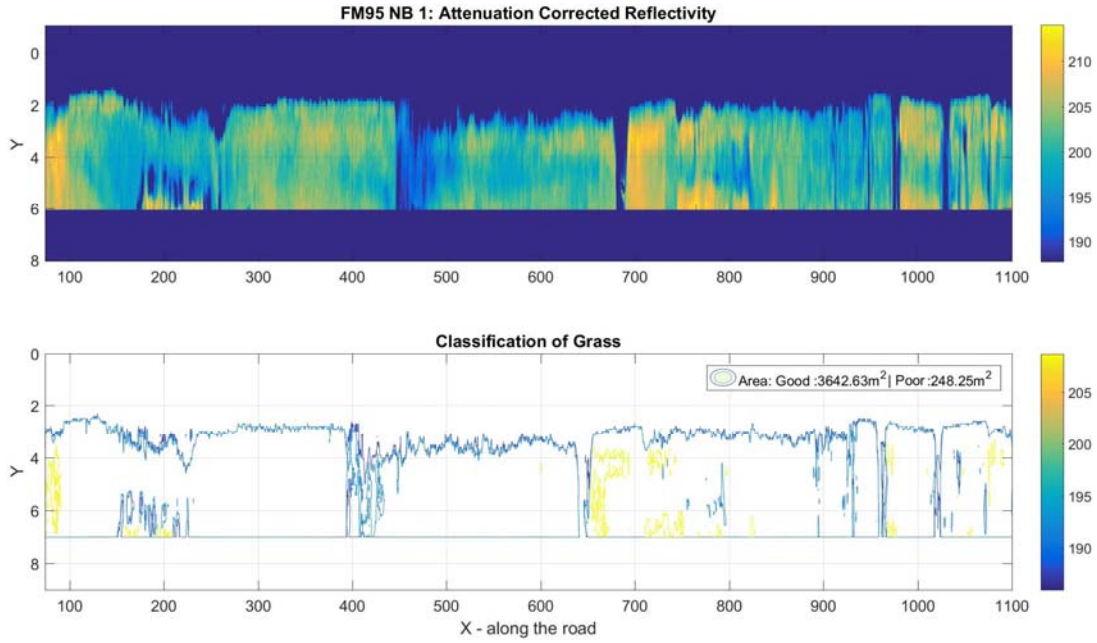
1. FM 320 South Bound 1 (FM320 SB 1)



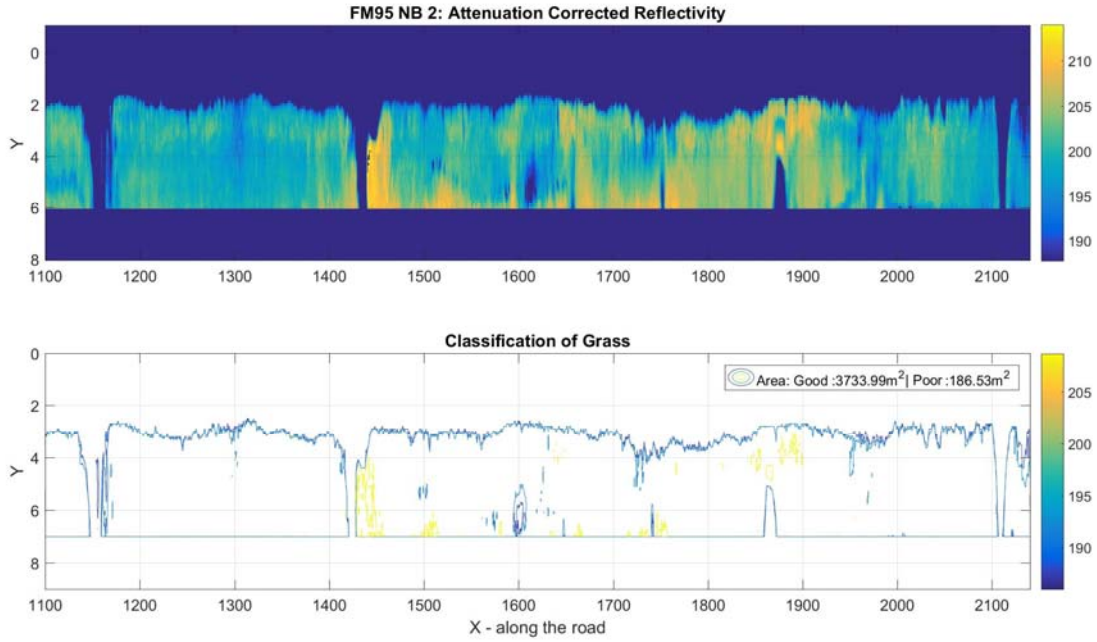
2. FM 320 South Bound 2 (FM320 SB 2)



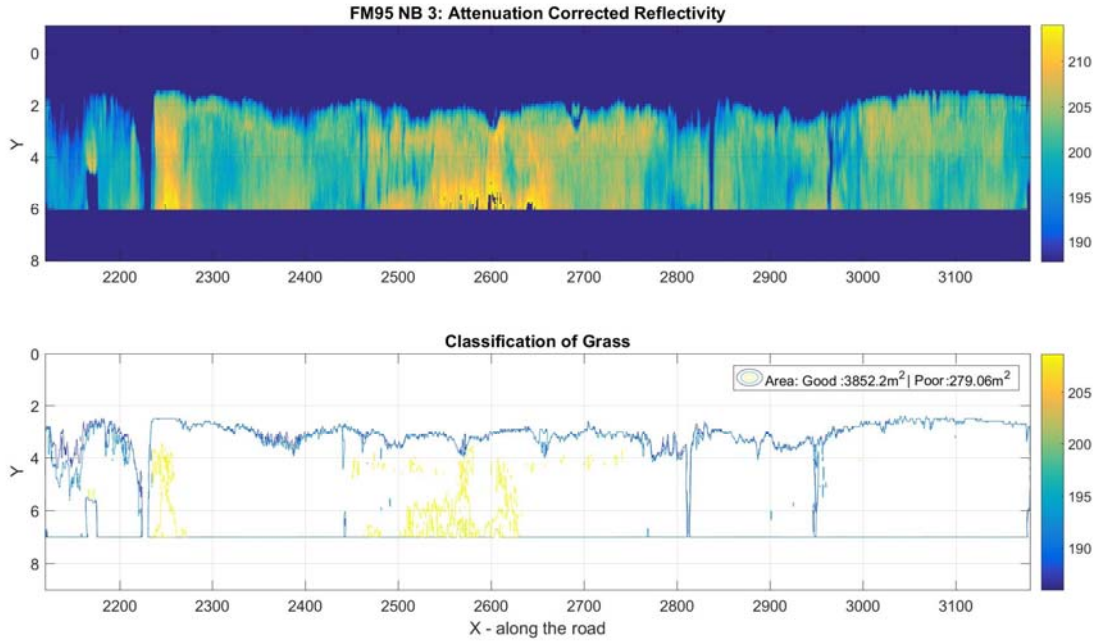
3. FM 95 North Bound 1 (FM95 NB 1):



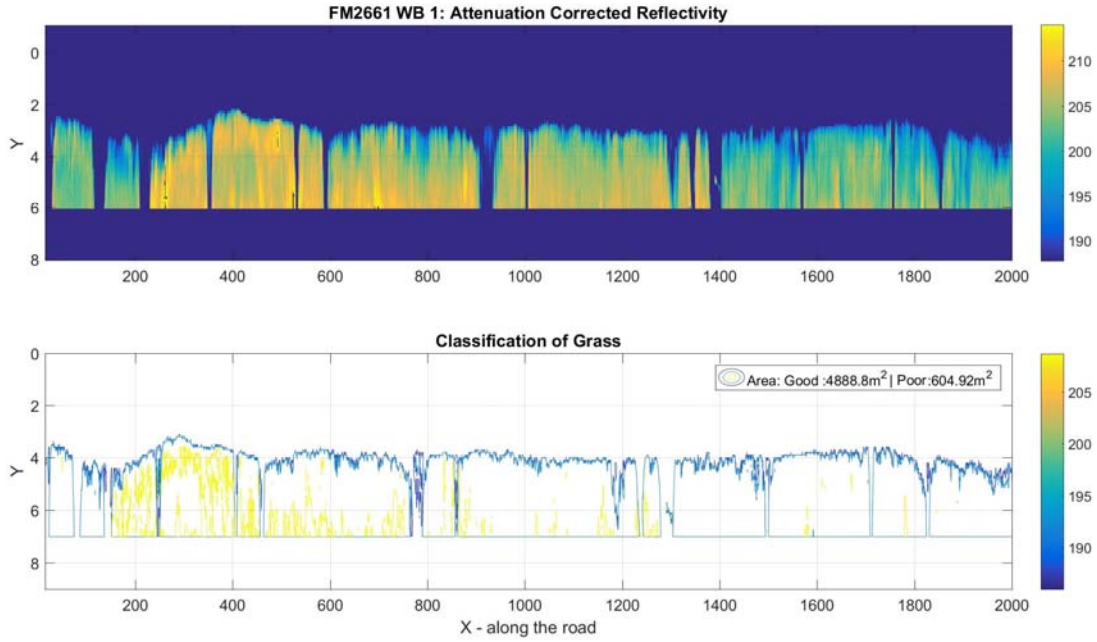
4. FM 95 North Bound 2 (FM95 NB 2):



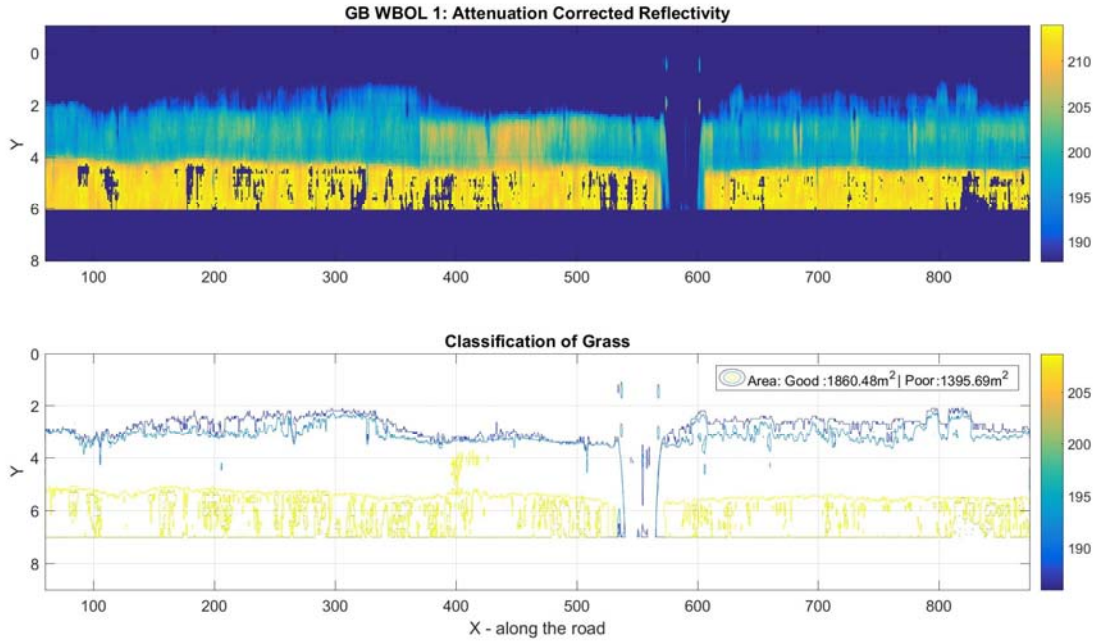
5. FM 95 North Bound 3 (FM95 NB 3):



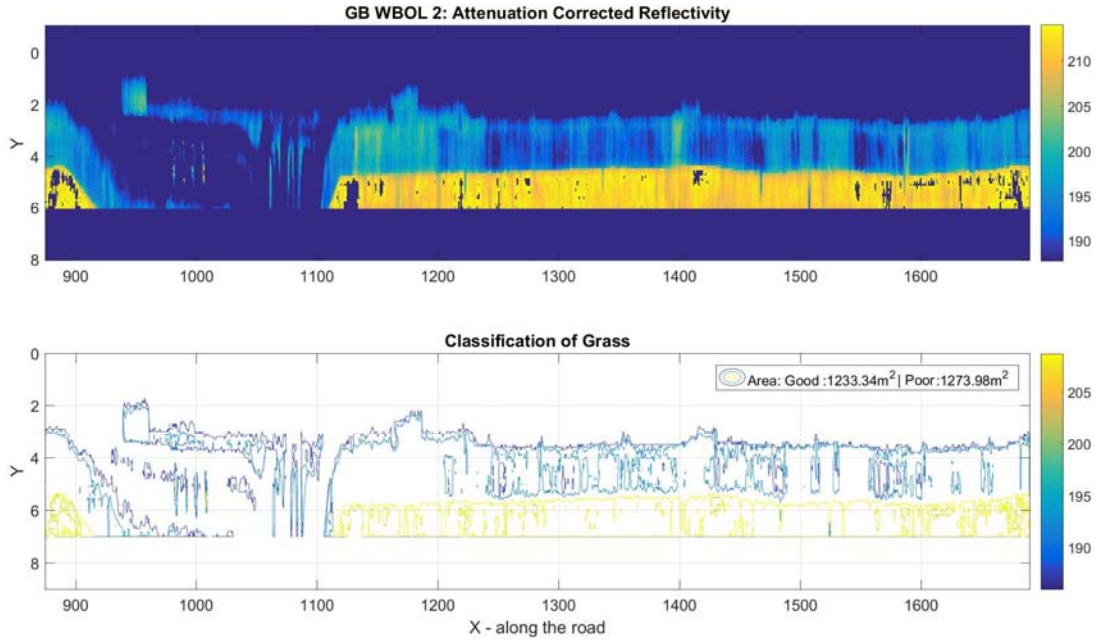
6. FM 2661 West Bound 1 (FM2661 WB 1):



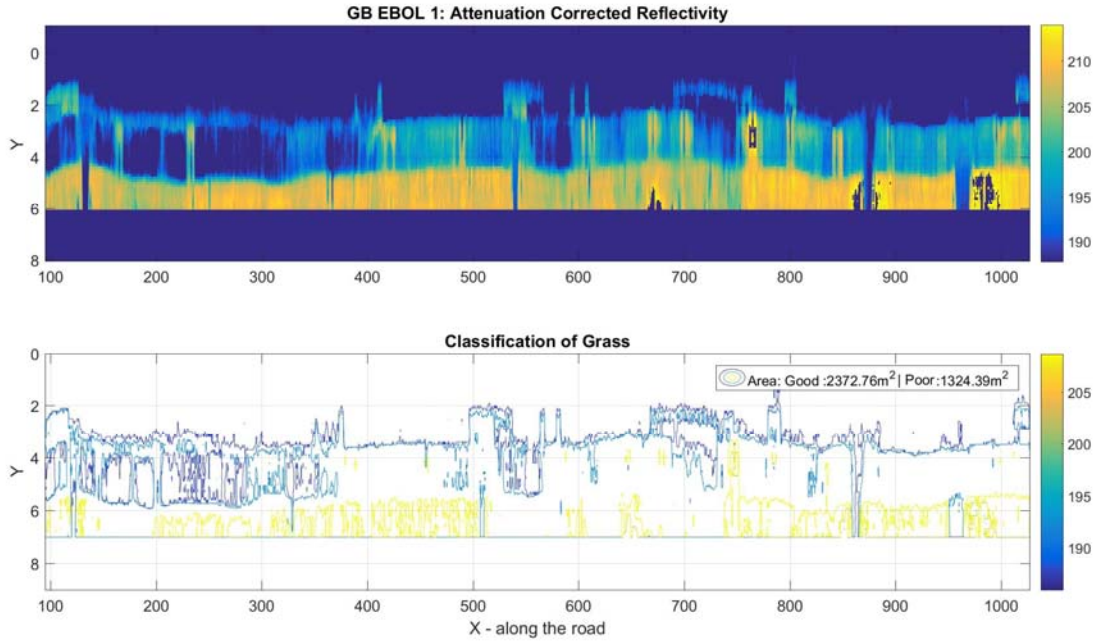
7. George Bush Dr. West Bound 1 (GB WBOL 1):



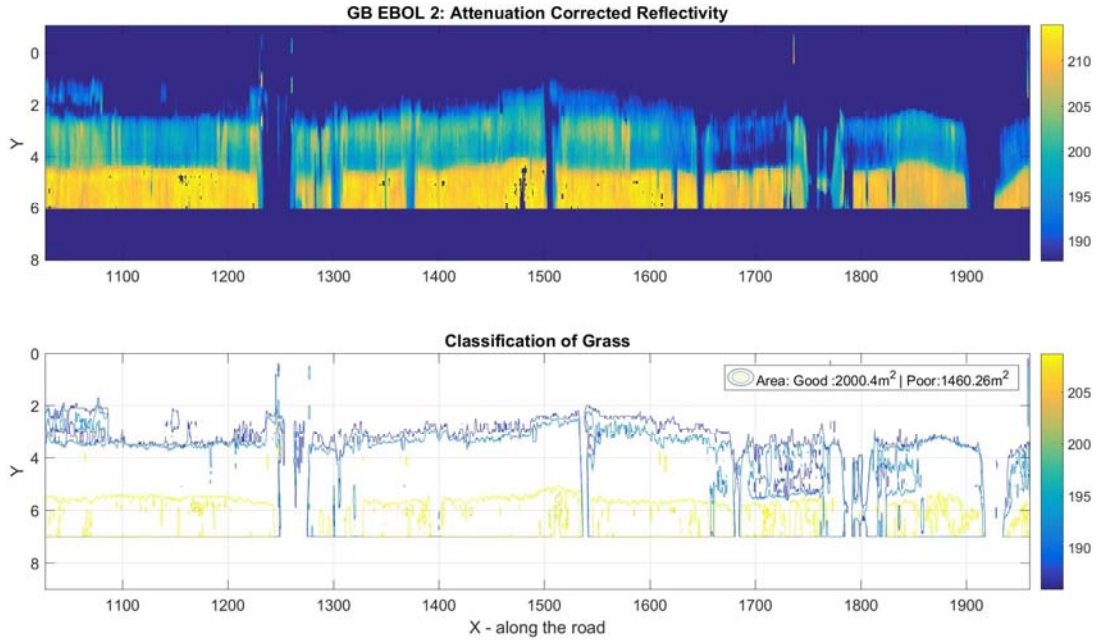
8. George Bush Dr. West Bound 2 (GB WBOL 2);



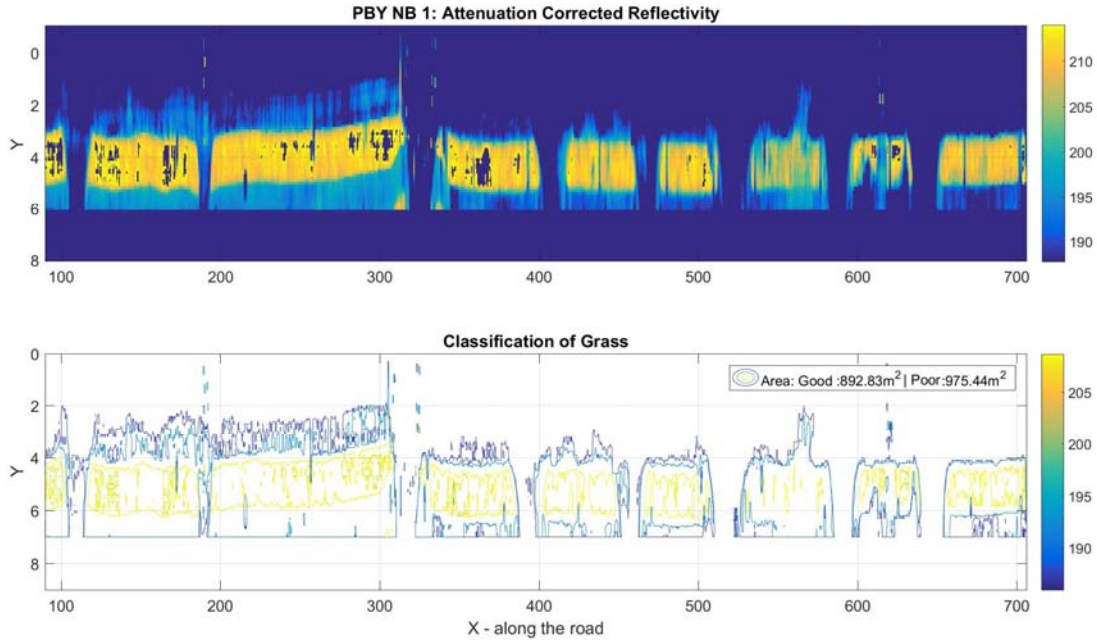
9. George Bush Dr. East Bound 1 (GB EBOL 1):



10. George Bush Dr. East Bound 2 (GB EBOL 2):



11. Penberthy Road North Bound 1 (PBY NB 1):

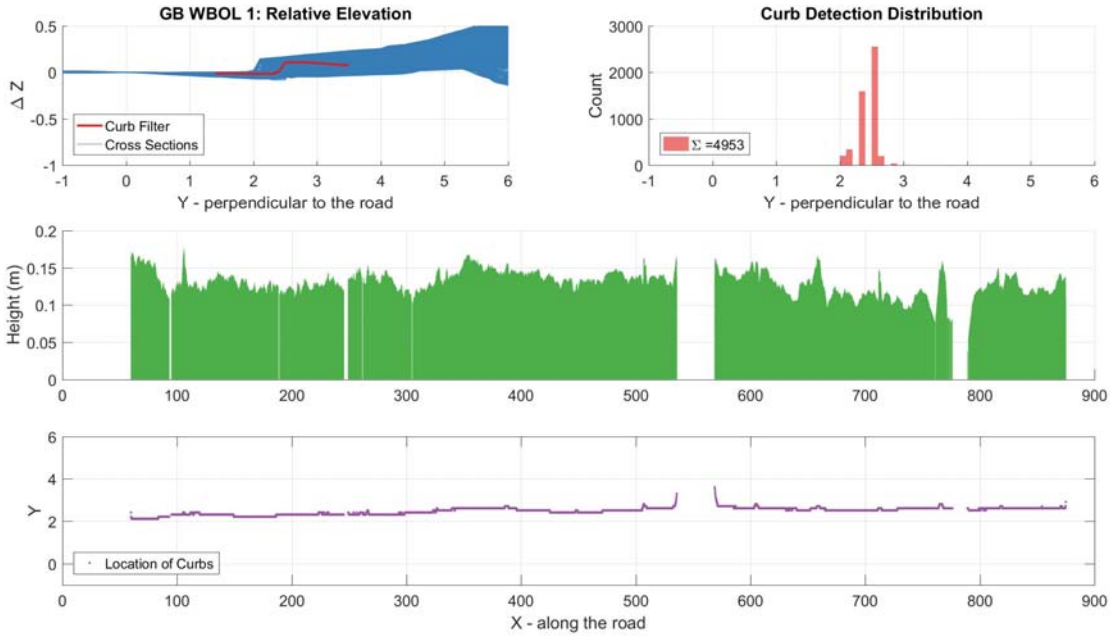


APPENDIX D

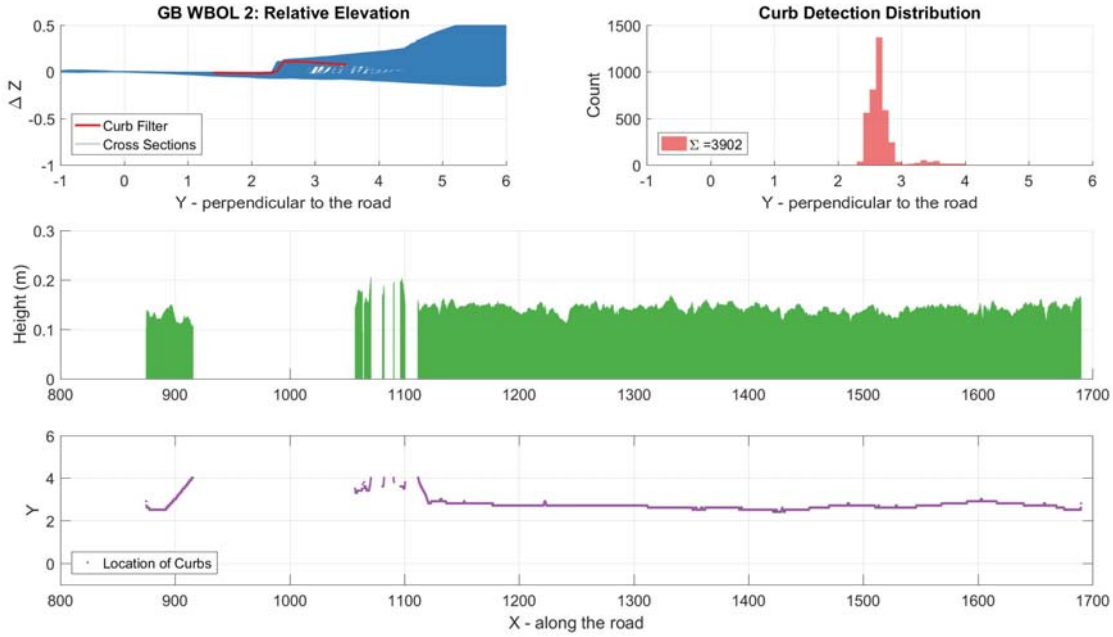
CURB DETECTION TESTS CARRIED OUT

Results of tests carried out for the algorithm developed for curb detection is presented in this section. A total of eleven sections were tested out of which five concrete road sections had road side pavement and six seal-coated farm-to-market road sections did not. All the sections had frequent driveways and intersecting roads -- intermittently breaking continuous curbs where present.

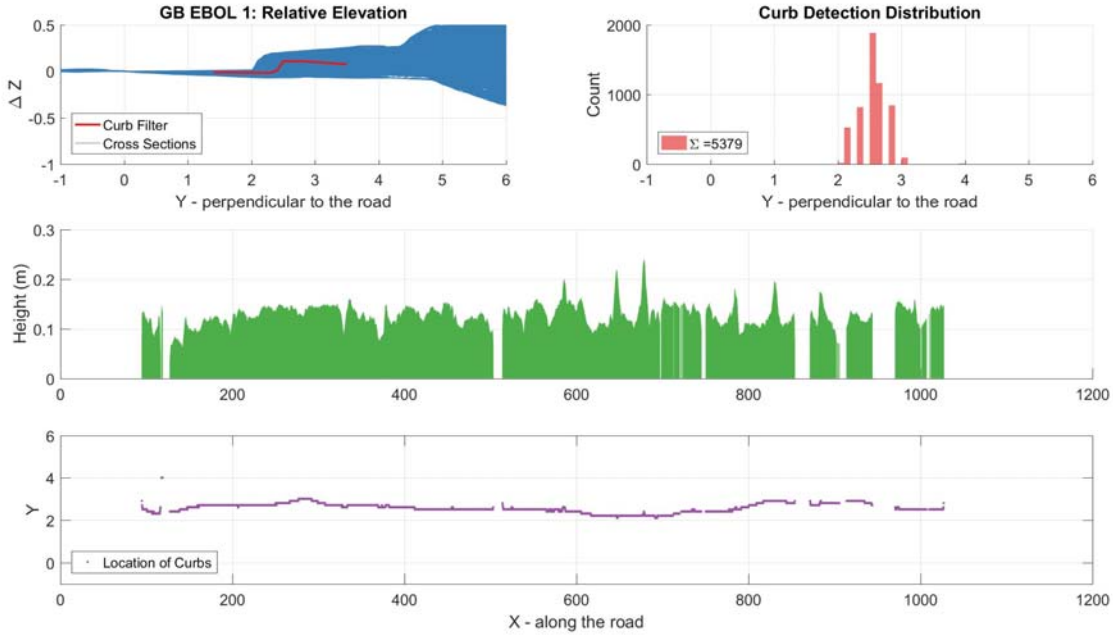
1. George Bush Dr. West Bound 1 (GB WBOL 1):



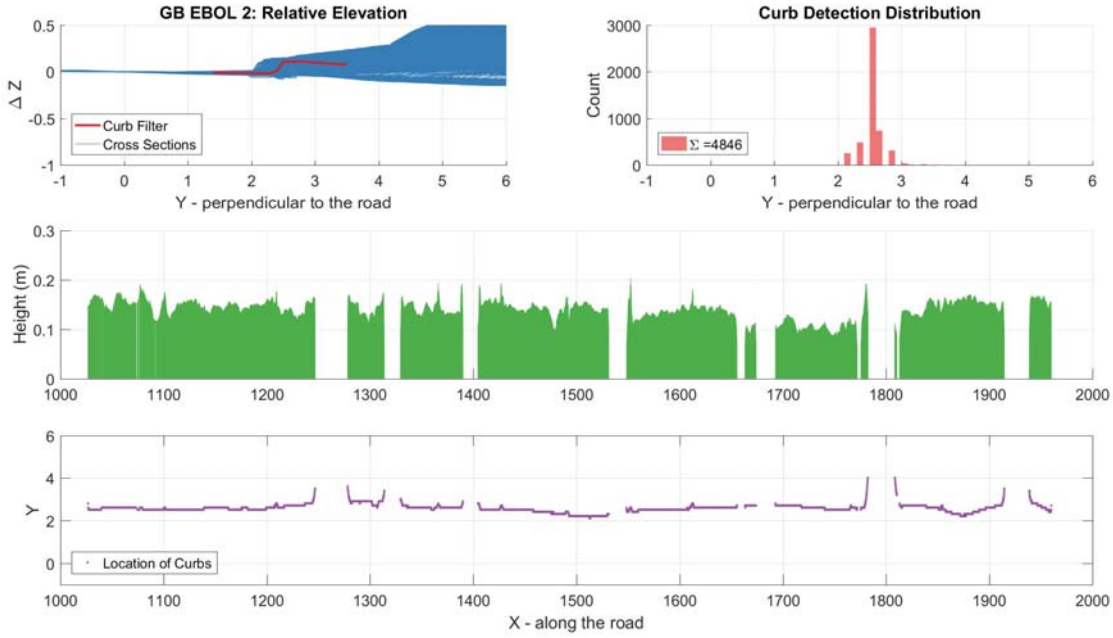
2. George Bush Dr. West Bound 2 (GB WBOL 2);



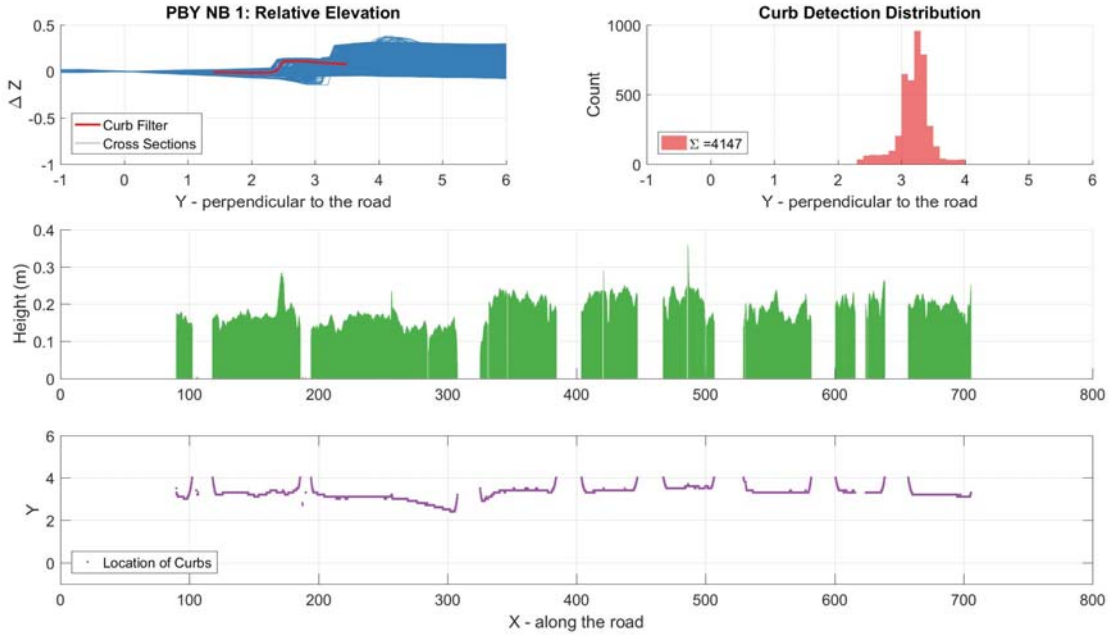
3. George Bush Dr. East Bound 1 (GB EBOL 1):



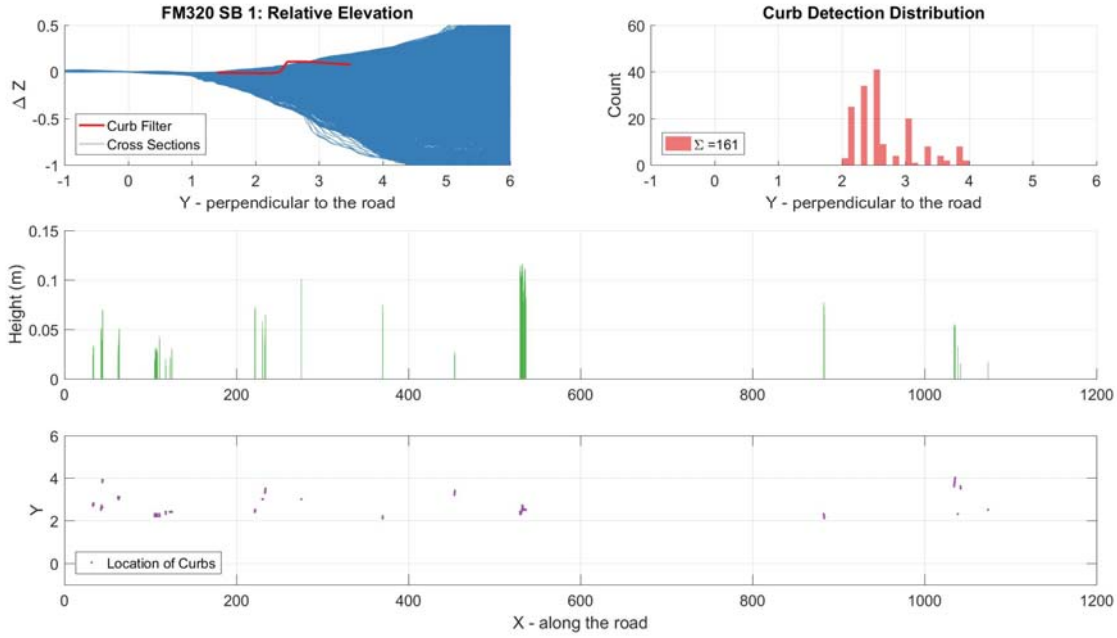
4. George Bush Dr. East Bound 2 (GB EBOL 2):



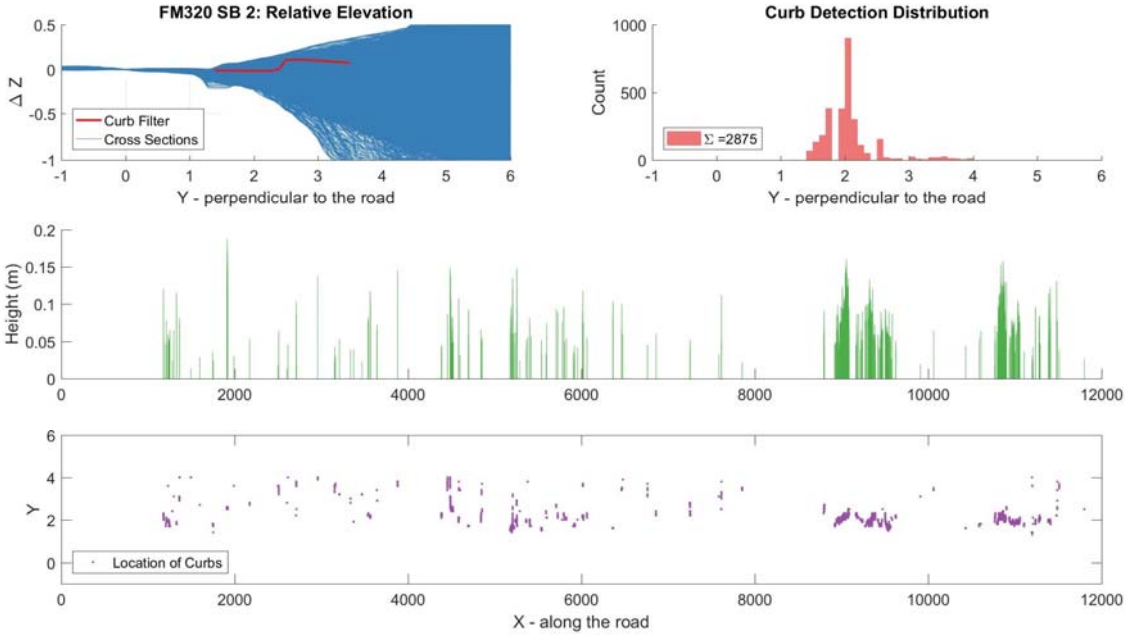
5. Penberthy Road North Bound 1 (PBY NB 1):



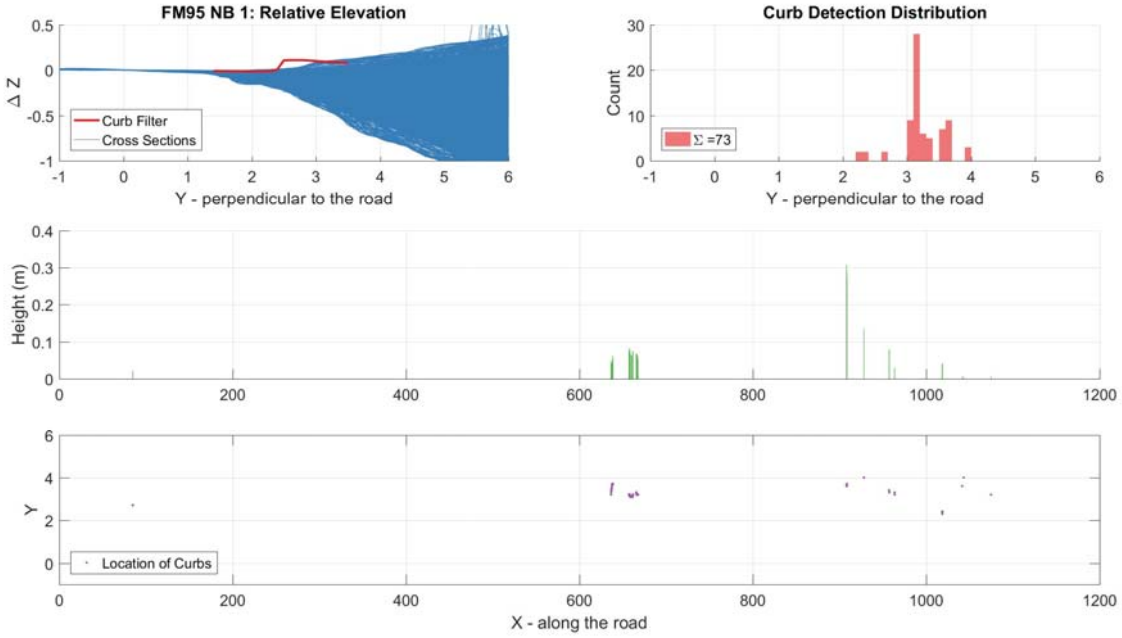
6. FM 320 South Bound 1 (FM320 SB 1)



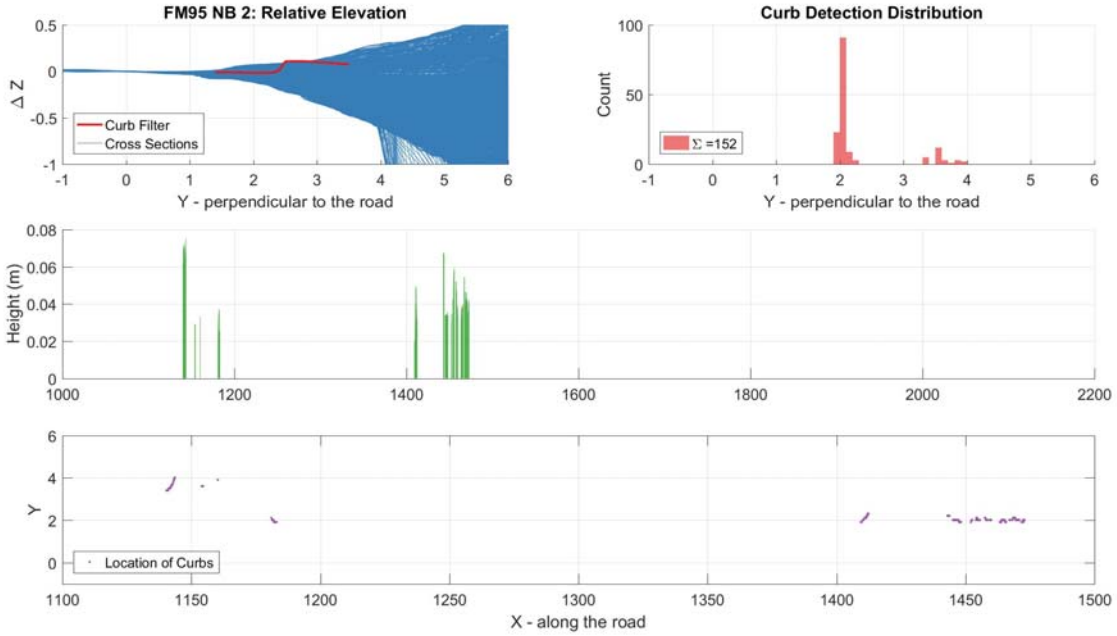
7. FM 320 South Bound 2 (FM320 SB 2)



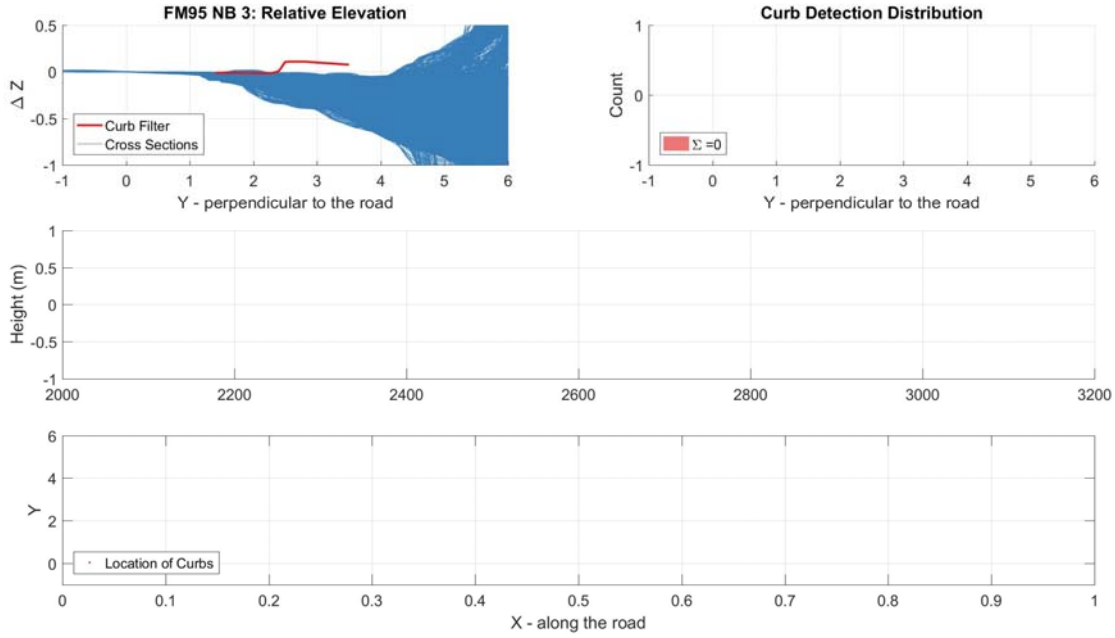
8. FM 95 North Bound 1 (FM95 NB 1):



9. FM 95 North Bound 2 (FM95 NB 2):



10. FM 95 North Bound 3 (FM95 NB 3):



11. FM 2661 West Bound 1 (FM2661 WB 1):