A Galerkin Approach to Define Measured Terrain Surfaces with Analytic Basis Vectors to Produce a Compact Representation

Heather M. Chemistruck

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

Doctor of Philosophy

In

Mechanical Engineering

John B. Ferris

David J. Gorsich

Scott L. Hendricks

Daniel J. Inman

Saied Taheri

Robert L. West

November 1, 2010

Danville, VA

Keywords: Terrain surfaces, INS drift, Hilbert Space, Principle Component Analysis, Galerkin Method

Report Documentation Page

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE **01 NOV 2010**	2. REPORT TYPE	3. DATES COVERED **00-00-2010 to 00-00-2010**
4. TITLE AND SUBTITLE **A Galerkin Approach to Define Measured Terrain Surfaces with Analytic Basis Vectors to Produce a Compact Representation**		5a. CONTRACT NUMBER
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Virginia Polytechnic Institute and State University,Blacksburg,VA,24061-0002**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S)
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The concept of simulation-based engineering has been embraced by virtually every research and industry sector (Sinha, Liang et al. 2001; Mocko and Fenves 2003). Engineering and science communities have become increasingly aware that computer simulation is an indispensable tool for resolving a multitude of scientific and technological problems. It is clearly desirable to gain a reliable perspective on the behaviour of a system early in the design stage, long before building costly prototypes (Chul and Ro 2002; Letherwood, Gunter et al. 2004; Makarand Datar 2007; Ersal, Fathy et al. 2008; Mueller, Ferris et al. 2009). Simulation tools have become a critical part of the automotive industry due to their ability to reduce the time and money spent in the development process. Terrain is the principle source of vertical excitation to the vehicle and must be accurately represented in order to correctly predict the vehicle response in simulation. In this dissertation, non-deformable terrain surfaces are defined as a sequence of vectors where each vector comprises terrain heights at locations oriented perpendicular to the direction of travel. The evolution and implications of terrain surface measurement techniques and existing methods for correcting INS drift are reviewed as a framework for a new compensation method for INS drift in terrain surface measurements. Each measurement is considered a combination of the true surface and the error surface defined on a Hilbert vector space, in which the error is decomposed into drift (global error) and noise (local error). It is also desirable to develop a compact, path-specific terrain surface representation that exploits the inherent anisotropicity in terrain over which vehicles traverse. In order to obtain this, a set of analytic basis vectors is formed iii from Gegenbauer polynomials, parameterized to approximate the empirical basis vectors of the true terrain surface. It is also desirable to evaluate vehicle models and tire models over a wide range of terrain types, but it is computationally impractical to store long distances of every terrain surface variation. This dissertation examines the terrain surface, rather than the terrain profile, to maximize the information available to the tire model (i.e. wheel path data). A method to decompose the terrain surface as a combination of deterministic and stochastic components is also developed.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**227**	

A Galerkin Approach to Define Measured Terrain Surfaces with Analytic Basis Vectors
to Produce a Compact Representation

Heather M. Chemistruck

Abstract

The concept of simulation-based engineering has been embraced by virtually every research and industry sector (Sinha, Liang et al. 2001; Mocko and Fenves 2003). Engineering and science communities have become increasingly aware that computer simulation is an indispensable tool for resolving a multitude of scientific and technological problems. It is clearly desirable to gain a reliable perspective on the behaviour of a system early in the design stage, long before building costly prototypes (Chul and Ro 2002; Letherwood, Gunter et al. 2004; Makarand Datar 2007; Ersal, Fathy et al. 2008; Mueller, Ferris et al. 2009). Simulation tools have become a critical part of the automotive industry due to their ability to reduce the time and money spent in the development process.

Terrain is the principle source of vertical excitation to the vehicle and must be accurately represented in order to correctly predict the vehicle response in simulation. In this dissertation, non-deformable terrain surfaces are defined as a sequence of vectors, where each vector comprises terrain heights at locations oriented perpendicular to the direction of travel. The evolution and implications of terrain surface measurement techniques and existing methods for correcting INS drift are reviewed as a framework for a new compensation method for INS drift in terrain surface measurements. Each measurement is considered a combination of the true surface and the error surface, defined on a Hilbert vector space, in which the error is decomposed into drift (global error) and noise (local error). It is also desirable to develop a compact, path-specific, terrain surface representation that exploits the inherent anisotropicity in terrain over which vehicles traverse. In order to obtain this, a set of analytic basis vectors is formed

from Gegenbauer polynomials, parameterized to approximate the empirical basis vectors of the true terrain surface. It is also desirable to evaluate vehicle models and tire models over a wide range of terrain types, but it is computationally impractical to store long distances of every terrain surface variation. This dissertation examines the terrain surface, rather than the terrain profile, to maximize the information available to the tire model (i.e. wheel path data). A method to decompose the terrain surface as a combination of deterministic and stochastic components is also developed.

Dedication

This dissertation is dedicated to my parents, Ronald and Cathy Chemistruck, who have supported me and motivated me to succeed in all of my endeavors. They have taught me that anything is possible with enough dedication and hard work. And to the rest of my family and friends who have pushed me to keep going, especially my brothers Michael Chemistruck, Daniel Chemistruck and Andrew Chemistruck. And lastly to my dog Roxy who has been there to welcome me with a wagging tail at the end of each day.

Acknowledgements

I owe my deepest gratitude to my advisor, Dr. John Ferris for his unwavering support, encouragement and guidance throughout my graduate career. This dissertation would not have been possible without him or the support of my talented committee members and the current and former members of the Vehicle Terrain Performance Laboratory. I would like to give special thanks to my committee members, Dr. David Gorsich for his insight to the importance of terrain modeling and a wonderful experience working on site at TARDEC; Dr. Scott Hendricks for his excellence in education and mastery of mechanical systems; Dr. Daniel Inman for his excellence in education and world knowledge; Dr. Saied Taheri for understanding the impact of the terrain on tire and vehicles, and Dr. Robert West for his unwavering support of students and insight to mechanical systems.

A special thanks to the research sponsors who made this research possible- the Tank and Automotive Research Development and Engineering Command (TARDEC) of the Army and the Federal Highway Administration (FHWA).

Thank you to Virginia Tech for providing the facilities and resources to successfully conduct my research. Special thanks go out to the IT department at the Institute for Advanced Learning and Research (IALR) in Danville, VA and the Mechanical Engineering IT department at Virginia Tech for all of your help. More thanks go out to Cathy Hill, Sue Teel, Pam Patterson and Amanda Glenn for being exceptional individuals and going the extra mile to work with the "Danville-kids." I am indebted to the Graduate Studies Associate Department Head, Dr. Mary Kasarda for her continual support and guidance throughout my studies, especially for providing the "Danville-kids" with a strong theoretical math class to take. I would also like to thank Dr. Steve Southward for his help and support along my path towards earning my degree, specifically by taking the time out of his busy schedule to offer his perspective to overcome obstacles that I have met along the way.

Lastly, I would like to thank those closest to me for their unwavering support (you know who you are).

Table of Contents

List of Figures

List of Tables

1. Introduction

This dissertation is focused on formulating a method to define a compact yet accurate representation of terrain surfaces that can be universally applied to various terrain measurement systems and types of terrain. This work begins by identifying and removing a primary source of error inherent to all terrain measurement systems equipped with an Inertial Navigation system (INS), known as INS drift. By removing the effect of this drift, the precision of the terrain measurement is improved by an order of magnitude and the true terrain surface is estimated. Next, the terrain surface is decomposed into principle characteristics, such as elevation, bank angle and rutting, which are identified by basis vectors. The empirical basis vectors are approximated by analytic basis vectors through a Galerkin approach, which is demonstrated on several samples of typical US highway data. In this way, the set of analytic basis vectors are developed such that they can be used to characterize all terrain of a particular type. The analytic basis vectors are formed from the discretization of a set of generating analytic functions. The discretization of these functions is based on the capabilities of the particular terrain measurement system used to sample the terrain. This method eliminates the necessity to calculate empirical basis vectors for each set of data by implementing the analytical basis vectors corresponding to that type of terrain (e.g., US Highways). The projection of the terrain data onto the analytic basis vectors can be used to characterize the terrain data in terms of its principle components. This method allows this characterization to be independent of the means by which the terrain system discretely samples the terrain surface. The Vehicle Terrain Performance Laboratory (VTPL) Vehicle Terrain Measurement System (VTMS) is used to acquire the data on a sample of US Highways as a proof of concept for this research. US Highways were selected as a terrain type because of their extensive use by the driving community, offering approximately 75,000km (46,000 mi) of paved surfaces. In this way, the principle components of the terrain surface are developed and can be combined to form an increasingly accurate representation of the terrain surface. Lastly, using this method of decomposing the

1

terrain, each principle component of the terrain surface can be classified as deterministic or stochastic. If the terrain component can be classified as stochastic, then it can be modeled using the same approach by which a terrain profile can be modeled, resulting in a more compact statistical representation.

The remainder of this chapter provides motivation for this research on developing a compact representation of terrain surfaces, especially for terrain profiler development and tire and vehicle simulations, as presented in Section 1.1. The problems that this research seeks to address are identified in Section 1.2. The thesis statement and scope of work are introduced in Section 1.3, followed by main contributions in Section 1.4. Lastly, a brief outline of the dissertation is presented in Section 1.5.

1.1 Motivation

Throughout a vehicle development program, it is necessary to possess chassis load data representing severe customer usage to ensure that the vehicle will perform as required. Unfortunately, actual loads are only available at the conclusion of the program. The design engineer is challenged with using predicted chassis loads early in the design process to design the 'best' components with the currently available knowledge. It is in these early stages that changes are relatively easy and inexpensive to make. However, measured chassis loads representing true loading conditions are not available until late in the program when changes to the design are extremely costly, if allowed to be implemented at all. More knowledge about a vehicle's true loading conditions early in the design process would reduce cost and timing for a new vehicle. Similar difficulties are faced by tire engineers since tractive efforts are largely dictated by the tire-terrain interaction (Shoop 2001). It is advantageous, therefore, to accurately predict these target chassis loads early in the program and to maintain a consistent process for predicting chassis loads as the design develops throughout the program.

In order to predict these target chassis loads, a set of terrain that will be consistently used to excite the vehicle must be determined at the beginning of a vehicle development program (Ferris and Larsen 2002). It is clear that the terrain is the main excitation to the vehicle (Aurell and Edlund 1989). Non-deformable terrain topology imposes a unilateral contact constraint on rolling tires to which the tires respond by

2

generating forces and moments. The terrain remains a consistent excitation to the vehicle, even as the vehicle design changes. Knowledge of this excitation, when applied in conjunction with high-fidelity tire and vehicle models, would allow chassis loads to be accurately predicted in vehicle simulations. Therefore, throughout the design process the system response to this consistent excitation can be calculated and compared for each vehicle design considered. Accurate terrain models would then provide the vehicle designer with a powerful tool to make informed design decisions early in the design process while changes are relatively inexpensive to implement. This will, in turn, shorten vehicle development time and reduce overall development costs.

The current practice for vehicle and tire simulations is to use terrain profile data as the input to these models. The terrain profile is an indexed set of terrain heights extending longitudinally along each wheel path. These profiles can be considered signals and in many cases can be modeled as a stochastic process, thus creating a compact representation of the terrain profile. The ability to capture these signals in a compact form is a major advantage of this simple representation of terrain. Although computationally efficient, this compact representation limits the available fidelity of data to excite tire models. It would be advantageous to the design engineer to implement terrain surface data as the input to tire and vehicle models since terrain surfaces are capable of capturing more detailed information about the tire contact patch.

The terrain surface can be represented as an indexed set of transverse vectors of terrain height, represented in matrix form. That is, each vector of the matrix comprises a set of terrain heights located perpendicular to the direction of travel of the vehicle. Although terrain surface data are more computationally expensive to use in simulation, they typically provide better estimates of the tire-road interaction. Toward this end, this work proposes a method to compactly represent terrain surfaces through representation by analytic functions.

1.2 Problem Statement

Terrain profiles, whether measured or modeled, are a useful tool for many groups. For example, civil engineers use terrain profiles of highways for pavement health monitoring to determine when a highway needs to be resurfaced. Vehicle engineers use

terrain profiles for physical simulations on shaker rigs to test the durability of a vehicle. An engineer working in modeling and simulation uses terrain profiles to create a virtual simulation to test a proposed vehicle design before it is built. While terrain profiles are the common practice, it would be advantageous to expand each group's capabilities to use terrain surfaces. Terrain is the consistent input into ground vehicle suspensions and is, therefore, vital to understand when designing or analyzing any ground vehicle.

The American Society of Testing and Materials (ASTM) defines a longitudinal profile as the perpendicular deviations of the pavement surface from an established parallel to the lane direction, usually the wheel tracks (ASTM 1989). However, this paper also includes surfaces that are not paved, therefore, throughout the rest of this work the term terrain will be used to represent on-road and off-road surfaces that a vehicle may traverse. Terrain profiles, as shown in Figure 1, consist of longitudinal distance traveled and corresponding vertical height. Note that the horizontal axis is distance traveled in meters and the vertical axis is the height in millimeters (this is to accentuate vertical profile characteristics).
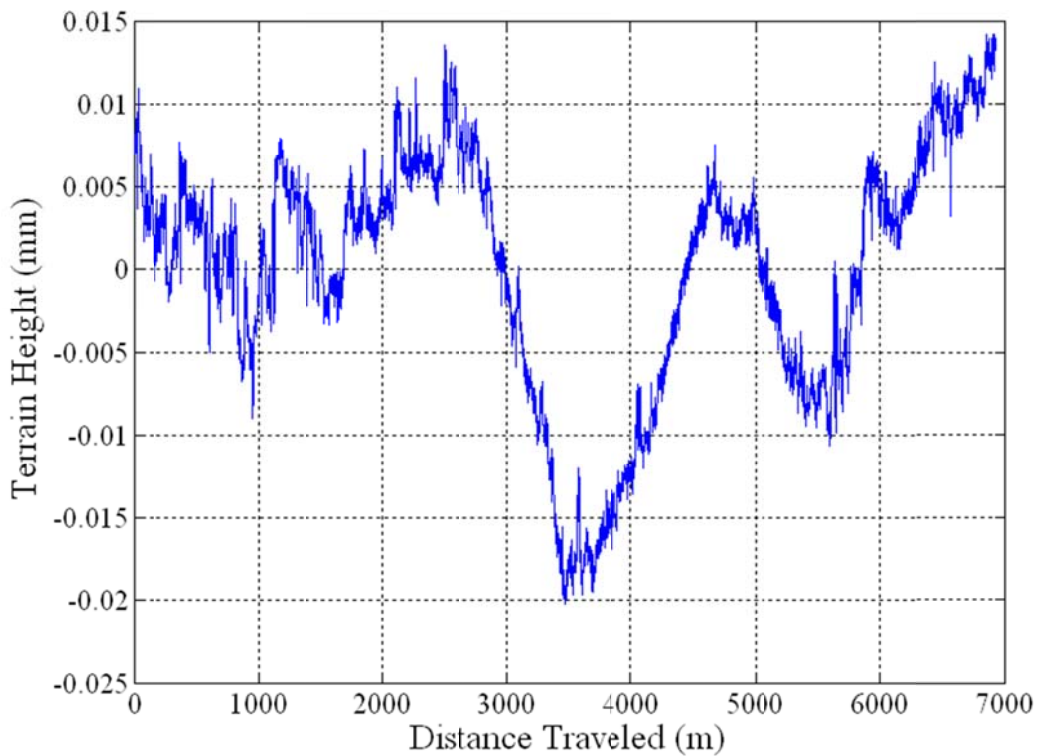
Figure 1. Sample longitudinal terrain profile.

Recent optical and computational advances have produced terrain measurement systems (Herr 1996) that can discretely sample the terrain at 941 points across a 4.2 meter wide transverse path, resulting in four-millimeter resolution of the terrain surface measurement. A sample of a measured cobblestone terrain surface, rendered as a contour plot, is shown in Figure 2. If desired, terrain profiles can be identified from the terrain surface measurements for analysis, but as it can be seen from Figure 2, there is more information captured in the surface measurement as compared with the profile measurement in Figure 1. These three-dimensional (3D) scanners are capable of differentiating between small localized disturbances and disturbances that will excite the vehicle chassis due to the fact that the scanner measures the entire width of the lane, instead of a single line down the road. This is particularly important for properly characterizing the envelopment property of the tire for accurate vehicle simulations, such that vehicle responses will not be affected by narrow dips and small stones in the terrain mapping (Karamihas 2005). Thus, the problem with using a single longitudinal profile is that it assumes that the profile represents the entire width of the tire contact patch that the tire is traversing, when the profile may actually contain a rock or a crack that produces a non-representative response. This is merely one problem that can be solved by providing more detailed knowledge of the terrain.
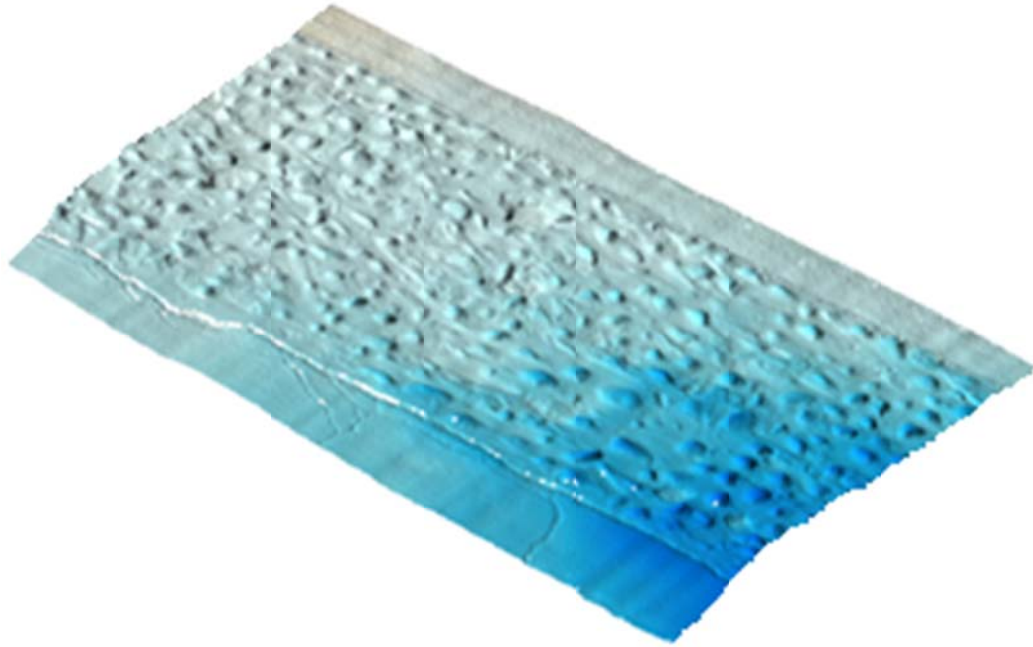
Figure 2. Sample terrain surface.

Despite the advantages of these 3D terrain measurement systems for terrain surface acquisition, several issues must be addressed. Current terrain measurement systems acquire approximately one-million data points per second. While this data acquisition rate improves the available signal bandwidth and allows sharp disturbances to be detected in both the transverse and longitudinal directions, for longer stretches of terrain, storing and manipulating these data is a daunting task. One benefit of the techniques developed in this work is a reduction in the steps required to post-process the measured terrain data for implementation in tire and vehicle simulations, as well as replacing measured terrain data with a stochastic terrain surface.

The potential of having a tool to accurately model and reproduce terrain surfaces has long been recognized, especially since previous studies have only been focused on characterizing and synthesizing terrain profiles. It is computationally impractical to simulate a vehicle traveling over long stretches of terrain, requiring choices to be made about the physical terrain features that are most appropriate for a given application. Toward this end, the methods developed in this work decompose the terrain surface into its principle characteristics, subjectively described as elevation, bank angle, crowning, and rutting to name a few. Each terrain component can then be statistically analyzed

individually, such as identifying the roughness contributed by each component. The overall contribution to the terrain surface could also be studied, as in the degree of elevation change or the degree of rutting is exhibited by this section of road. In this way, terrain surfaces can be grouped into meaningful sets with similar physical characteristics based on similarities of terrain components from which sets of surfaces can be selected to satisfy the needs of given applications.

Additionally, difficulties arise when two adjacent lanes have been measured and need to be concatenated to form one continuous terrain mapping. Accurate and representative concatenation of the two lanes is critical if a lane change maneuver is to be performed on the data set. A solution to this problem is developed in Chapter 3 in which the true surface is estimated from measured sets of terrain data.

1.3 Thesis Statement and Scope of Work

Thesis Statement: A high-fidelity yet compact representation of terrain surface types can be developed that are insensitive to the particular measurement system being used and allows for the study of principle terrain characteristics.

The primary goal of this research is to develop a technique to develop a compact representation of a terrain surface while improving the fidelity of the estimated terrain surface. This work focuses on non-deformable terrain, defined herein as terrain whose surface deformation due to a single vehicle traversing the surface is negligible, such as paved roads (both asphalt and concrete), gravel roads, and typical off-road trails; deformable terrain such as sand and snow are beyond the scope of this work. The focus of this work is further restricted to terrain surfaces that are anisotropic in nature, defined herein as having an inherent path defined over which vehicles travel and where that travel, over time, has defined obvious paths or lanes. Specifically this work pertains to nearly 6.3 million kilometers (DOT 2008) of non-deformable anisotropic terrain over which nearly 254.4 million vehicles travel each day (DOT 2006). The results of this study are intended for application in the accurate prediction of tire loads based on full-knowledge of the tire patch: lateral, longitudinal, and radial tire forces and the overturning, rolling resistance and self-aligning tire moments acting on the contact patch.

7

Better modeling of the contact patch will yield more representative tire loading, which will lead to better estimates of chassis loading scenarios and will inevitably result in better ground vehicle reliability predictions. Current application of this work is limited to the analysis of U.S. Highways, comprised of both asphalt and concrete, but the methods developed in this dissertation are applicable to any non-deformable anisotropic terrain.

1.4 Main Contributions

This research focuses on the development of a high-fidelity and compact representation of terrain surfaces that are insensitive to the terrain measurement system used to acquire the data. Nevertheless, the methodology developed in this work is not limited to the example cases. For example, any discretized polynomial or series can be used to approximate the empirical basis vectors, as long as it provides the smallest root mean square error solution. Additionally, these methods apply for any non-deformable terrain surfaces, but additional research will be needed to determine if they apply for deformable terrain. The main contributions of this research are:

1) The first representation of terrain surfaces as elements of a Hilbert Space, $\mathbf{H}$
    a. Terrain surfaces are explicitly defined as a sequence of vectors
    b. Each vector comprises terrain heights at locations oriented perpendicular to the direction of travel.
    c. A vector space is defined by the span of these vectors
    d. A weighted inner-product is defined and the norm is defined as the induced norm
2) A novel method to remove INS drift:
    a. The Hilbert space is decomposed into a Global subspace, $\mathbf{G}$, and the complementary local subspace, $\mathbf{G^c}$.
    b. A sequence of error vectors are defined such that each error vector is the difference between the measured terrain surface vector and the true surface vector.
    c. The three vectors comprising the measurement, error, and true surface at each longitudinal location have components in both the Global and Local subspaces.

d. The Global subspace is defined such that the component of the error vectors in the Global subspace can be modelled as a random-walk process and the elements of the error vectors in the Local subspace are elements of a zero-mean uncorrelated noise process.

e. The Global error, arising from Inertial Navigation System (INS) drift, is identified and removed from the measurements such that the standard deviation of the residual noise process (Local error) is within the resolution of the measurement system (+/-1mm)

f. The true surface is computed from the drift-free measurements using a non-linear expectation estimator.

3) The first rigorous definition of principle terrain characteristics:

a. Although principle terrain characteristics, such as elevation, bank angle, crowning, and rutting have been subjectively described and techniques for *ad hoc* measurement of these characteristics have been defined (AASHTO 2009), this research develops sets of empirical basis vectors for US Highways that are consistent between asphalt and concrete roads

b. The principle terrain characteristics are defined with analytic basis vectors that closely approximate these consistent empirical basis vectors by minimizing the root mean square error between the empirical and analytic basis vectors.

4) It is the first method that has been developed that can define terrain surfaces without sensitivity to the measurement system used to acquire the terrain data

a. Analytic basis vectors can be discretized based on terrain measurement system capabilities to sample the terrain

5) A novel method to stochastically represent a terrain surface

a. Components of terrain (known as $\sigma_{i,l}$) capture the magnitude of the principle terrain characteristics along the length of the terrain surface

b. Components of terrain are classified as deterministic or stochastic

c. Stochastic components of terrain can be modelled in the same manner as terrain profiles

      d. Truncated terrain surface is synthesized based on stochastically modelled components of terrain

1.5 Dissertation Outline

This work is organized as follows. Chapter 1 motivates the research and presents the scope of work, research objectives and main contributions. In the second chapter, recent literature pertaining to terrain measurement techniques and sources of error, such as drift contributed by the Inertial Navigation System (INS) are presented, followed by the definition of the coordinate system used throughout the developments in this work. The second chapter concludes with background on current methods used to study the roughness of the terrain and the latest terrain modeling techniques. A novel technique for removing the drift components of error from terrain measurements is developed in the third chapter. The fourth chapter develops a terrain surface representation that retains sufficiently high-fidelity information while possessing the simplicity of terrain profiles, as well as being insensitive to the choice of the terrain measurement system from which the data were acquired. A novel method to represent a stochastic terrain surface is developed in the fifth chapter. Applications of this research, conclusions that can be drawn for this dissertation and areas for future work are presented in chapter six.

2. Background

Terrain can typically be grouped into two major categories, the first is defined as empirical terrain, based on physical measurements of the terrain surface, and the second is based on mathematical models aimed at characterizing the physical measurements, defined here as analytical terrain. It is further important to distinguish between the concept of continuous and discrete, or analog and digital. There exists a continuous empirical terrain surface, known as the true surface. This surface has infinite bandwidth and maintains every detail of the terrain (Sayers and Karamihas 1998). Due to limitations of technology, it is only achievable to sample the true terrain surface at discrete locations, resulting in discrete empirical terrain surfaces, or digital terrain surfaces. Furthermore, there exist continuous analytical functions that are used to approximate continuous terrain surfaces. Additionally, synthetic terrain can be generated by the continuous analytical functions, and then discretized to match the same sampling rate of the empirical data for appropriate analysis of the two segments. With this differentiation, several areas of research that will be reviewed in this chapter under the empirical terrain classification are terrain measurement systems, including terrestrial terrain measurement systems and digital terrain measurement (DTM) systems, the coordinate system used for empirical terrain data, and roughness indices used to classify the components of terrain. Under the analytical terrain classification, state-of-the-art methods used to approximate empirical terrain data with mathematical models, including statistical analysis, terrain characterization, and approximation techniques will be reviewed. This chapter concludes with a concise summary of the state-of-the-art in terrain studies.

2.1 Empirical Terrain

2.1.1 Terrestrial Terrain Measurement Systems

The technological developments in the area of terrain measurement has progressed from vehicle-response systems (Hveem 1960; Gillespie, Sayers et al. 1980) to the measurement of various types of terrain using vehicle-independent systems

(Dembski, Rizzoni et al. 2006; Dembski, Rizzoni et al. 2006; Kern and Ferris 2007; Wagner, Kern et al. 2007; Liu, Dembski et al. 2008). As a result of these technologies, terrain measurement systems have been developed with various capabilities and can be classified as one of the four types of systems (Perera and Kohn 2002; Wang 2006):

1) Manual Devices
2) Profilographs
3) Response type road roughness measuring systems (RTRRMs)
4) High-speed inertial profilers

Manual devices are typically used as reference devices for measuring terrain because the measurements are taken statically and do not contain the same error associated with dynamic systems. The standard test method for measuring road roughness with a manual device is described in ASTM Standard E1364 (ASTM 1996). The most common types of manual devices are the rod and level and the dipstick (Sayers and Karamihas 1998). The rod and level are most commonly used in surveying. The level provides the elevation reference, and the rod measures the height of the terrain relative to the level. The Dipstick is a type of walking profiler that is more automated than the rod and level system as it is equipped with an onboard computer for recording elevation height and an inclinometer to measure the angle of the surface. Despite the role of these devices as reference terrain measurement systems, they are time consuming and do not allow for a high sampling rate of the vertical elevation of the road.

A)

B)
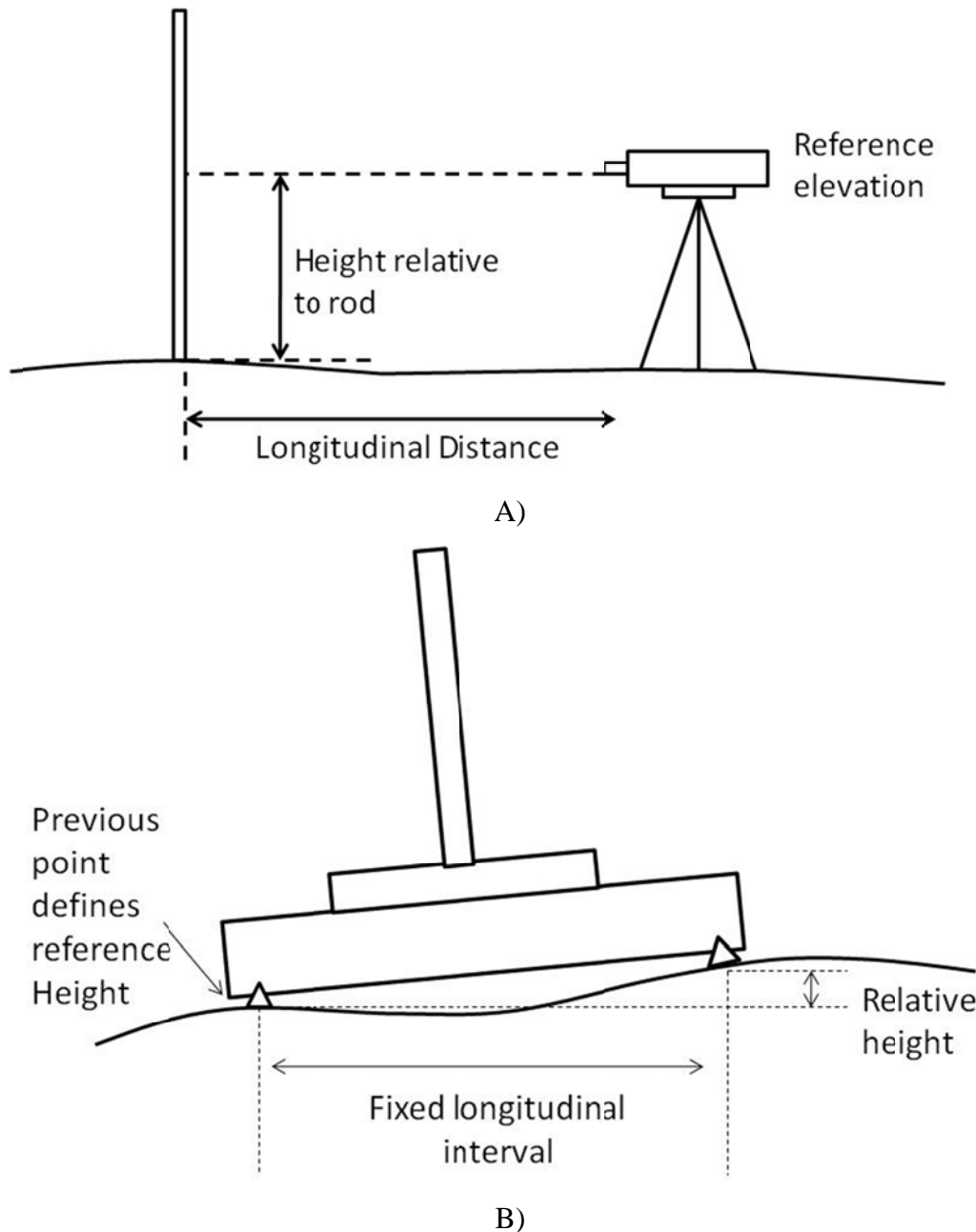
Figure 3. Samples of manual devices A) Road and Level B) Dipstick

Profilographs are most commonly used in evaluating the smoothness of paved surfaces. These systems are typically large, mobile, rigid beams with either a center wheel that is linked to a computer which records the movement of the wheel from the established datum, or modern systems can be equipped with a series of lasers to capture multiple profiles along the span of the device. Regardless of the setup, these measurement systems have been found to have inherent errors in their ability to

appropriately measure the wavelengths present on the road- some wavelengths have been amplified or attenuated as compared with the reference device. When comparing roughness index values between these systems and the reference device, it has been found that the measurements resulting from profilographs can totally miss the wavelength which affects ride quality and roughness indices. The misreading may result in a "smooth" roughness rating, when in actuality the pavement needs to be rehabilitated. More detail on roughness indices can be found later in this Chapter.

RTRRMs estimate the terrain surface by measuring and summing the vertical motion of the axle or road following wheel relative to the body of the host vehicle or trailer, with respect to horizontal distance traveled by the host vehicle or trailer. The measurement units are of average rectified slope (ARS) or more commonly known as mm/km or in/mi (Gillespie, Sayers et al. 1980; Sayers and Paterson 1986; Gillespie 1992; Wang 2006). The accuracy of the resulting effective terrain profile relies on the fidelity of the physical condition of the tire and vehicle, since they act as a mechanical filter to the terrain, as well as the operating speed of the host unit. Several examples of these RTRRMs include the Bureau of Public Roads (BPR) Roughometer, the CHLOE, Mays Meter and the PCA Meter to name a few. The BPR Roughometer, as seen in Figure 4A, consisted of a single-wheel trailer in which the mass, tire properties (i.e. temperature and pressure) and suspension system were standardized to minimize the effect of varying parameters on the roughness measurement (Hveem 1960; Gillespie 1992). The CHLOE, as seen in Figure 4B, featured a trailer equipped with two small wheels, located nine inches apart to measure the local road slope as the system was towed at low speeds along the road (Gillespie 1992). The Mays Meter, as seen in Figure 4C, was equipped with transducers on the host vehicle body and axle to measure the vehicle's suspension travel (1973; Perera and Kohn 2002). The PCA Meter was very similar to the Mays Meter in that it measured accumulated axle displacement, but this system gave a greater weight to larger displacements (Gillespie 1992). All of these systems are subject to being affected by the dynamics of the vehicle, particularly instability with time (a measurement completed in present day is not comparable to measurements made several years ago) and the roughness measurements are not transportable with measurements taken from different systems (Sayers and Karamihas 1998). In addition to these flaws, there are

numerous modeling difficulties to derive the original terrain profile in this process, as far as identifying and removing the mechanical filtering properties of the vehicle system (e.g., the enveloping property of tires precludes capturing the sharpness of terrain events). RTRRMs should only be used, therefore, to describe the terrain's influence on the host vehicle in terms of ride quality or making general statements about terrain roughness, not about generating the true terrain profile.

A)

B)
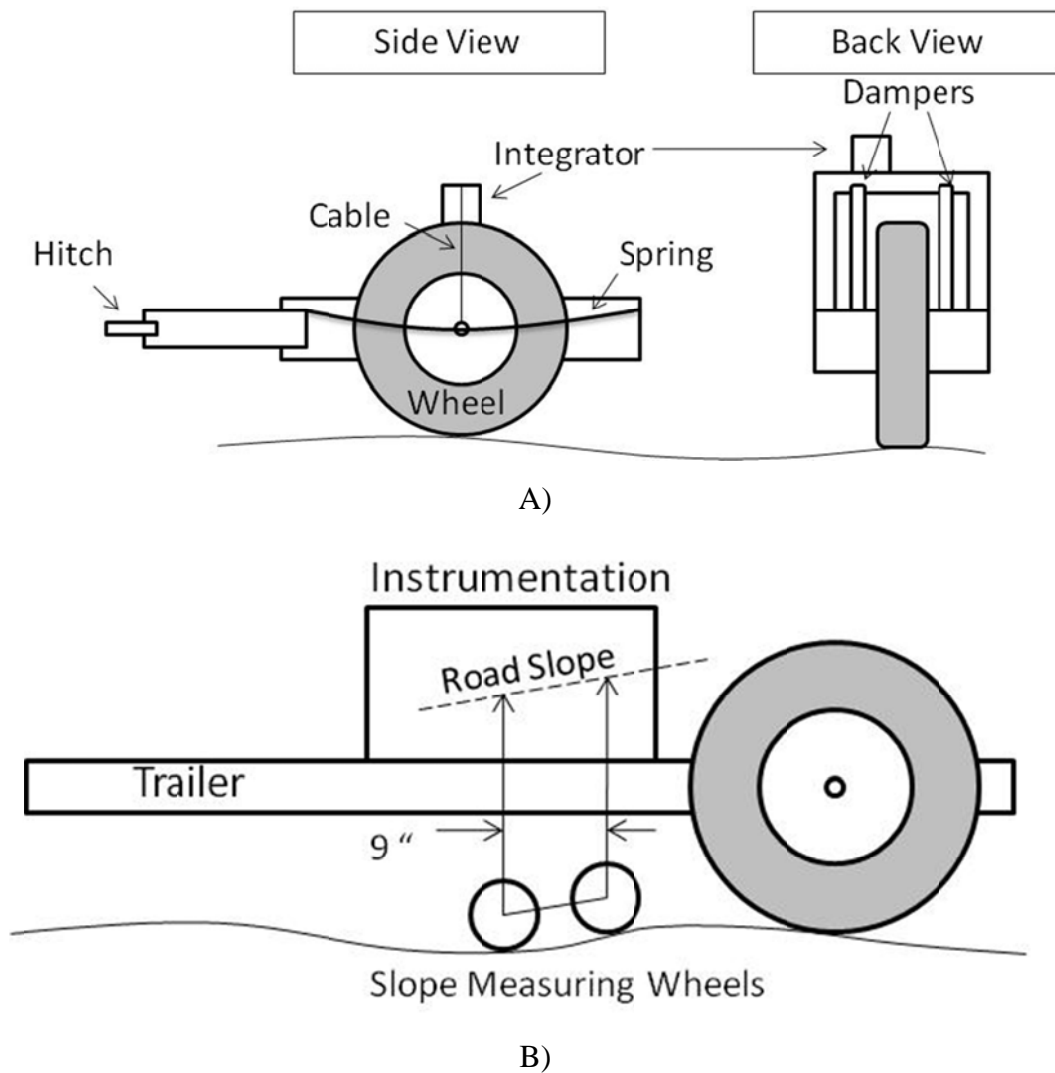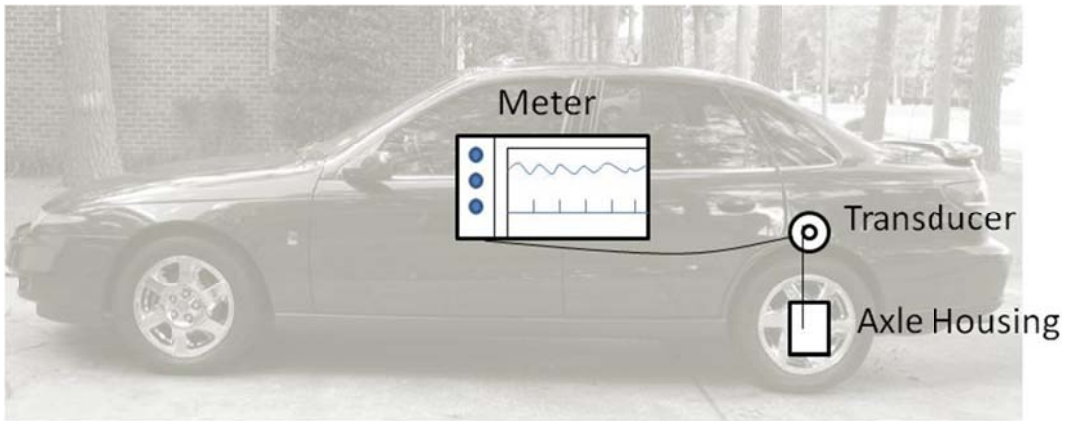
C)

Figure 4. Samples of RTRRMs A) BPA Roughometer B) CHLOE C)Mays Meter (Gillespie 1992).

A more precise method is direct measurement of terrain with high-speed inertial profilers, as shown in Figure 5. These systems are typically equipped with some combination of elevation sensors, accelerometers and distance measuring systems (Perera and Kohn 2002). The elevation sensors measure the distance from the terrain to the sensor, whether it is a laser, ultrasonic, optical, or infrared sensor. The accelerometers are used to measure the vertical acceleration of the host vehicle with respect to the elevation sensor, and used to establish an inertial reference. Lastly, the distance measuring system such as a global positioning system (GPS) or speedometer is used to provide the horizontal distance traveled. General Motors Research (GMR) developed the first high-speed inertial profiler in the 1960s (Spangler and Kelly 1966). This profiler implemented a road-following wheel that extended below the body of the host vehicle to measure the relative distance between the body and the road surface. An accelerometer was attached to the body and, having integrated the signal twice, measured the absolute instantaneous vertical position of the body. The resulting road profile is then the difference between the integrated signal and the measurement from the elevation sensor. A major fault in this system is that it relies on accelerometer data to remove unwanted body motion from post-processed data files. The reliability of these traditional accelerometer-based systems suffers when vehicle speed falls below 5 m/s (Walker and Becker 2006) and in other low frequency environments. Despite this, and the limitation to sample a single longitudinal profile per measurement, the design of the GMR road

16

profiler lead the way in technological advances for better, more consistent sampling of the terrain surface.
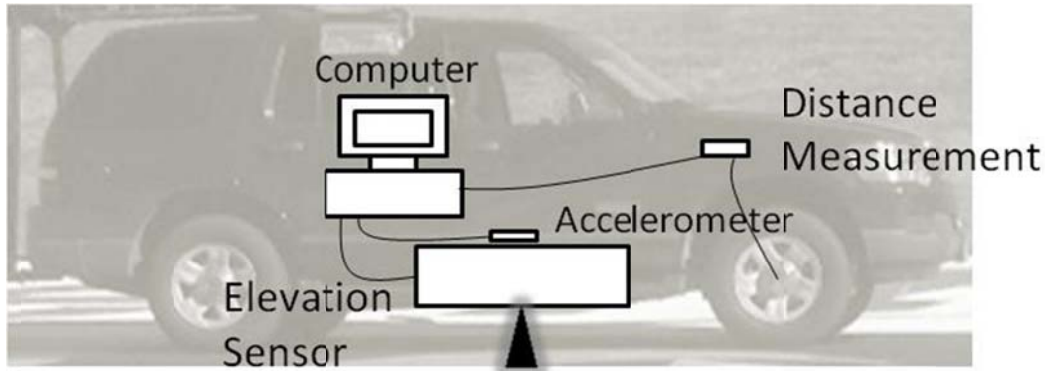
Figure 5. Basic diagram of a high-speed inertial profiler.

In 1987, The Australian Road Research Board (ARRB) substituted a laser for the road-following wheel (Prem 1987). In this system, the terrain is sampled optically, rather than mechanically as had been the standard practice. Presently, research is being conducted that examines the implications of this fundamental change. For example, the laser has a much finer resolution than the tire contact patch, capable of detecting small localized disturbances that would be enveloped or bridged by the tire of a typical passenger vehicle (Smith 2009).

As computer power increased and signal processing techniques evolved, high-speed inertial profilers became more capable (Kern and Ferris 2007; Wagner, Kern et al. 2007). Recent advances have produced 3D terrain measurement systems which can scan a wide transverse path with vertical resolution of less than a millimeter. Current terrain surface measurement systems incorporate a scanning laser (Herr 1996) that is rigidly mounted to the body of a host vehicle (Kern and Ferris 2007; Wagner, Kern et al. 2007; Liu, Dembski et al. 2008). The Vehicle Terrain Measurement System (VTMS) (Kern and Ferris 2007; Wagner, Kern et al. 2007), shown connected to the host vehicle in Figure 6, was used to acquire the data for this work. The scanning laser, affixed at the rear of the vehicle, acquires 941 data samples transversely across a 4.2 meter wide path each millisecond. In this particular system, the laser is emitted continuously and reflected against a rotation prism so that the system acquires data at equal samples in time. The laser system has a 90 degree field of view and the resulting parallax conversion produces

17

results in a Cartesian coordinate system. This vehicle traverses the terrain while simultaneously acquiring terrain measurements. When the vehicle encounters a disturbance, the laser translates and rotates with the body of the host vehicle. To obtain accurate terrain measurements, the motion of the vehicle must be accurately measured so that it can be removed from the laser measurement. Modern systems, including the VTMS, use an Inertial Navigation System (INS) to measure the vehicle movement (Kennedy, Hamilton et al. 2006). The accuracy of the INS depends on the alignment of the Inertial Measurement Unit (IMU) to the laser and satellite coverage of the Global Positioning System (GPS). Many factors affect the INS error. For example, small misalignment between the scanning laser and the Inertial Measurement Unit (IMU) will compound the error that can be anticipated from any Differential GPS (DGPS) (Smith and Ferris 2008). Once the data from the scanning laser and the INS are translated to a common coordinate system and synchronized in time, the laser data are nominally corrected for host-vehicle body motion and transformed to a global coordinate system.

Figure 6: Host Vehicle Terrain Measurement System (VTMS)

2.1.2 Digital Terrain Measurement Systems

Digital Terrain Modeling originated in the 1950's with the original purpose to aide in highway routing and design, with terrain data acquired photogrammetrically (Miller and Laflamme 1958). Photogrammetry is the practice of determining the geometric properties of an object from photos. A Digital Terrain Model (DTM) is defined as "an ordered array of numbers that represents the spatial distribution of terrain characteristics. The spatial distribution is represented by an XY horizontal coordinate system and the terrain characteristic which is recorded is the terrain elevation, Z (Doyle 1978)." Similar to DTMs, Digital Elevation Models DEM) are defined differently from DTMs because they capture different terrain characteristics. These models are defined by latitude φ and longitude coordinates λ, with the terrain elevation, h. The data from both models are organized either in matrix form by XY coordinates and the corresponding Z elevation or by equations of the surface defined by polynomials or Fourier series (Doyle 1978).

The coordinate system used in this early work was based on the State Plane Coordinate System. In the 1930s, the United States was divided into specific regions and a specific Cartesian coordinate system was assigned to each region. In the 1950's the coordinate system transitioned to the Universal Transverse Mercator coordinate system, which is commonly known as the WGS84 ellipsoid used in GPS presently (Doyle 1978). For the purpose of highway design, an XY horizontal coordinate system was established with the X-axis in the general direction of the proposed highway alignment. As is described in section 2.1.4 of this work, the global coordinate for terrestrial terrain measurement systems is oriented with Y in the direction of vehicle travel, and X perpendicular to vehicle travel.

DTM data can be acquired from existing topographical maps, photogrammetric stereomodels, or more recently, airborne laser scanners. Before computers were common place, review of topographical maps was popular, but time consuming, and did not offer the same resolution that is available from airborne laser scanners of today. A better method used in the 1950's was photogrammetric stereomodels, where elevation data was formatted in numerous ways, such as contour lines, profiles with elevations recorded at regular intervals, or at geographic points along major terrain features (Doyle 1978).

19

Photogrammetric systems are limited by good lighting, where elevation points can be affected by shadows from trees. A notable photogrammetric system, the Gestalt Photomapping System, was capable of outputting an array of digital elevation data at a spacing of 182mm at photographic scale, corresponding to approximately 14.3m on the ground scale when using photographs at 1:80,000 scale (Kelly, McConnell et al. 1977; Rose 2001). Similar to terrestrial terrain measurement systems, as technology advanced, the area of digital terrain measurement systems moved from manual methods and photogrammetrics to airborne laser scanners. Airborne laser scanner data offers detailed 3D digital terrain mappings with approximately one elevation point per square meter over a 1000x1000 km area.

Airborne laser scanners, also known as light detection and ranging (Lidar), are used for the derivation of topographic data and the generation of DTMs. Most airborne digital terrain measurement systems are mounted on an aircraft or helicopter, with data taken in a strip-wise manner. The laser, operating near infra-red (IR) frequency ranges, acquires data at frequencies ranging from 1-83kHz resulting in elevation accuracy of a few centimeters (Briese and Pfeifer 2001). The scanning lasers used on these systems are insensitive to shadows, unlike the photogrammetric systems, however the aperture size on the laser is about half as small as that as the photogrammetric systems and requires more laser scans to be taken to cover the same area of terrain (Kraus and Pfeifer 1998). One version of an airborne laser scanner is the TopEye which is suspended from a helicopter. It is equipped with a vertical scanning direct detection laser radar operating at a wavelength of 1.06μm and pulse rate between 2 and 7kHz (Elmqvist, Jungert et al. 2001). The operational altitude is between 60-900m and is able to produce point position, intensity of reflection as well as multiple return or double echo data. When flying between 10-25m/s at altitudes of 120-375m, the point density of the acquired data varies from 2-16 points per square meter (Elmqvist, Jungert et al. 2001). These systems along with terrestrial systems that are equipped with GPS and INS are inherently plagued by systematic errors. These systematic errors are discussed in the next Section.

2.1.3 GPS and INS Error Sources

Typically, an INS is capable of establishing a geodetic position with two centimeter accuracy (Smith 2009). It is clear that large localized disturbances in the

measured terrain surface will cause erroneous vehicle simulation results. In addition, slowly varying height changes that are artifacts of the INS drift will skew the absolute height of the terrain surfaces in the global reference frame (Fritsch and Kilian 1994; Kraus and Pfeifer 1998). For independent lane scans, this problem may not affect the application of the terrain data, since the absolute height of the terrain may not be of interest. If a measurement, for instance, is taken over perfectly level terrain for one mile, a slowly varying height change of several centimeters will not hinder simulation effectiveness. However, if data were recorded along the first lane of a two lane highway early in the day and a few hours later the second lane was measured, then any vertical misalignment of the two lanes resulting from the INS drift would produce an artificial vertical shift between lanes. The misalignment between the two lanes would cause undesired excitations into a model undergoing a lane change maneuver. Similar to a method developed by Smith, Kraus suggested identifying terrain data that exist in two sets of data and compare those points to known survey data, and adjust the surface accordingly (Kraus 1997; Smith 2009).

Airborne laser scanners are plagued with the same systematic errors as the terrestrial systems, as both systems are equipped with an INS and a GPS. The difference is the magnitude of the INS drift. Experimentation has shown the root mean square (R.M.S.) error of airborne systems to be 10.7cm or more in vertical elevation accuracy as compared with survey data (Pfeifer, Stadler et al. 2001). This is a full order of magnitude greater than that seen by terrestrial applications, where terrestrial measurement systems may see a +/-10mm deviation due to INS drift (Chemistruck, Binns et al. 2010). Many different errors and uncertainties arise when working with airborne laser scanner data due to the amount of filtering that needs to be completed to reduce the acquired data to terrain data only (Li 1994; Lopez 1997; Kraus and Pfeifer 1998; Briese and Pfeifer 2001; Pfeifer, Stadler et al. 2001; Kraus, Briese et al. 2004). Focusing on the contributions of error from the GPS and INS alone, studies have shown three major results. The first is that the mean square error in determining the height of an elevation point as referenced by the GPS is worse than the accuracy of the laser scanner measurement. Second, the GPS phase ambiguities based on on-the-fly initialization for fast moving objects results in +/-10cm error in identifying the origin of the polar coordinates. All airborne laser

scanners use polar coordinates for data acquisition. Lastly, the INS orientation has an error of up to 0.01 degrees, corresponding to a terrain point accuracy of +/-1cm for a flying height of 1000m (Kraus and Pfeifer 1998). Data acquired from airborne laser scanners have also shown discrepancies in height when trying to stitch two sets of data together.

The accuracy of the INS is dependent on instrumental biases and environmental biases, such as the number of satellites used by the receiver, the satellite-receiver distance, atmospheric effects, satellite and receiver clock offsets, phase ambiguities, and others (Sardon, Rius et al. 1994; Sardon and Zarraoa 1997; Sarma 2008). To decipher the systematic errors based on the INS and GPS as compared with the measurement sensor, Smith developed an excitation event and calibration surface to identify and isolate the error of the measurement sensor, in this case a scanning laser (Smith and Ferris 2009; Smith 2009). He also developed an ad hoc method to compensate for error contributed by INS drift, but this method fails to retain the absolute height (Smith and Ferris 2008). According to Fritsch, "the INS drift can be estimated from the long term stability of the GPS observations and on the other hand the larger observation noise of the GPS observations could be compensated by the short term stability of the INS (Fritsch and Kilian 1994)."

A strong effort has been made to characterize the GPS and INS errors by categorizing the errors based on the error origin. Specifically, several techniques have been used to estimate instrumental biases. One of the first proposed techniques, from preoperational GPS systems in the 1990's, applies the least squares method to estimate the coefficients of a two-dimensional quadratic model for vertical Total Electron Count (TEC) from a single GPS receiver. TEC is defined as the number of electrons in a column 1 m2 square extending from the base station to the satellite with unites of e/m2 (Coco, Coker et al. 1991). Coco et al. applied Lanyi and Roth's technique to study the variability of the GPS instrumental biases (Coco, Coker et al. 1991). As technology advanced, ionospheric delay was investigated through the development of ionospheric maps. The ionosphere acts as a dispersive medium to GPS signals, thus ionospheric propagation delays can be removed by the use of two frequencies- L1 and L2 (Coco, Coker et al. 1991). The bias for the two GPS frequencies (L1=1575.42 MHz and

L2=1227.60 MHz) and their difference, is referred to as differential instrumental bias and will produce systematic errors in the estimates of the ionospheric delays (Sardon and Zarraoa 1997). Ionospheric delays are vertical delay estimates at specified Ionospheric Grid Points (IGPs). Estimation of the IGP delay is limited by instrumental biases. Instrumental biases are the difference of the propagation paths of L1 and L2 signals and is directly due to the circuitry in the GPS satellite and receiver hardware (Sarma 2008).

Wilson and Mannucci applied two techniques based on surface harmonics and triangular interpolation for the development of global/regional ionospheric maps (Mannucci, Hajj et al. 2004). When GPS systems became completely operational in the mid-1990's, regional and global TEC maps were developed to improve the overall accuracy of the GPS system. These operational systems implement the dual frequencies, L1 and L2, which were previously introduced. A better technique for estimating instrumental biases and TEC was proposed by Sardon through modeling stochastic parameters of the GPS errors with a Kalman filter (Sardon, Rius et al. 1994). Sardon does not explicitly identify noise, multipath, differential delay between L1 and L2 antenna phase centers or selective availability in his stochastic model, rather he lumps the errors into one term. More recently, Sarma et al. applies singular value decomposition (SVD) algorithm to estimate the instrumental biases from data for several dual frequency GPS receivers (Sarma 2008). A complete investigation of the accuracy of these systems has been conducted by Binns, Smith and Wang (Wang 2006; Smith and Ferris 2008; Smith and Ferris 2009; Smith 2009; Smith 2009; Binns 2010). A method for removing the error contributions from the GPS and IMU is developed in Chapter 3, without requiring the identification of each error source, such that the effects of INS drift in the measured surface are removed. The next section defines the terrain coordinate system used for theoretical development in this dissertation.

2.1.4   Defining the Coordinate System

The coordinate system for non-deformable, anisotropic terrain topology is defined in this Section. Terrain surfaces in this work will be defined using a curved regular grid (CRG) format; however it is important to note that other gridding methods exist. DTMs are typically organized in three types of data structures, square-regular grid, triangulated irregular network (TIN) and contour-based networks. Square-regular grids are most

computationally efficient, but fail to handle abrupt changes in the terrain surface. TINs are typically constructed with Delaunay triangulation and are most suitable for identifying rough terrain features (Petrie and Kennie 1987; Yoon 2003). Contour-based networks are best for identifying general trends in the terrain surface, and are particularly good for hydrological applications. Other gridding techniques exist, but the techniques briefly discussed here are most common for terrain applications (Wilson and Gallant 2000). The CRG format is discussed in detail next as it is the basis for the theoretical discussion in this work.

The result of data acquisition and signal processing from the VTMS is a point cloud of terrain height data, located in non-uniform locations in the horizontal plane, as shown in Figure 7a (Kern and Ferris 2007; Wagner and Ferris 2007). Airborne terrain measurement systems also acquire data in a non-uniform grid-spacing, but these systems are faced with additional post-processing issues to remove vegetation and buildings from the data set (Kraus and Pfeifer 1998; Kraus, Briese et al. 2004). Once the data from the scanning laser(s) and the INS are translated to a common, body-fixed coordinate system, synchronized in time, and nominally corrected for host-vehicle motion, the data are transformed to a global coordinate system (x, y, z). The result is a point cloud of terrain height data, located in the horizontal plane as shown in Figure 7a (Kern and Ferris 2007; Wagner and Ferris 2007). A path coordinate, u, and a perpendicular coordinate, v, are introduced to form a curved regular grid (CRG) in the horizontal plane, as shown in Figure 7b (Gimmler, Ammon et al. 2005; Rauh and Mossner-Beigel 2008; VIRES October 2008). Discrete longitudinal locations along coordinate u are defined as vector u and are indexed by i; discrete transverse locations are defined as vector v and are indexed by j. The terrain height corresponding to each discrete horizontal grid point (u, v) is determined by examining the vertical heights in the localized area around that horizontal location. The spatial interpolation method (e.g., mean, median, inverse-distance-to-a-power, and kriging) must be carefully chosen based on the application (Stein 1999; Stein, Meer et al. 2002; Detweiler and Ferris 2008; Detweiler and Ferris 2009; Wackernagel 2010), the accuracy of the measurement system (Pfeifer, Kostli et al. 1998; Pfeifer, Stadler et al. 2001; Detweiler and Ferris 2008; Smith and Ferris 2008; Smith 2009), and the local homogeneity and isotropic nature of the terrain.. In this way,

each terrain surface comprises a sequence of transverse scans, perpendicular to the direction of vehicle travel, that have been nominally corrected for body motion. This body motion correction relies on the ability to accurately measure the position and orientation of the vehicle, typically using an INS that is prone to small amounts of drift (Fritsch and Kilian 1994; Pfeifer, Kostli et al. 1998; Wagner, Kern et al. 2007). A method to correct for INS drift is presented in Chapter 3 of this dissertation.
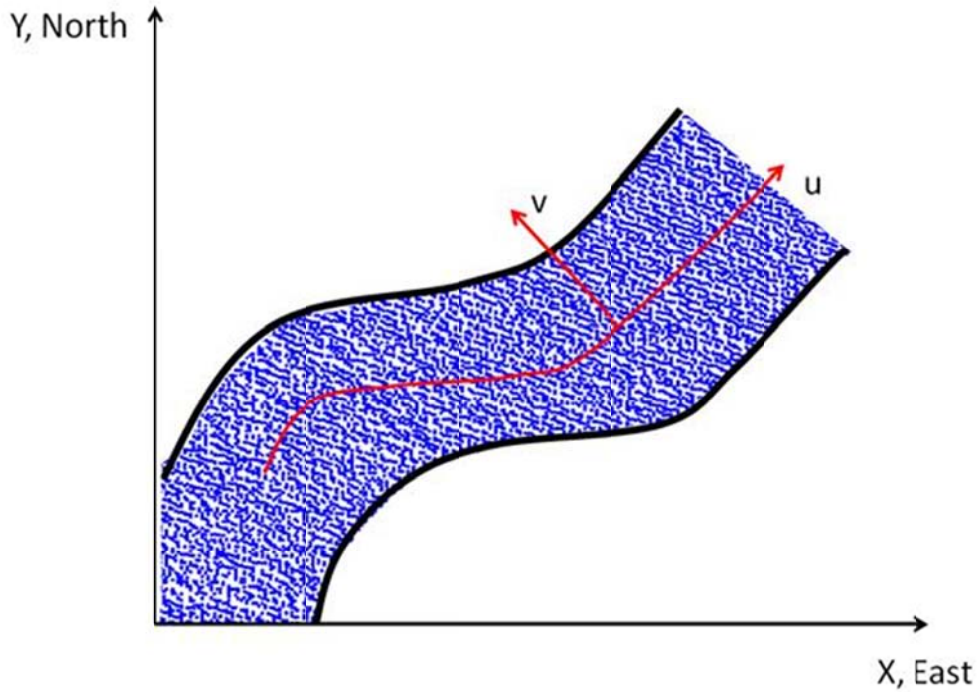
(A)
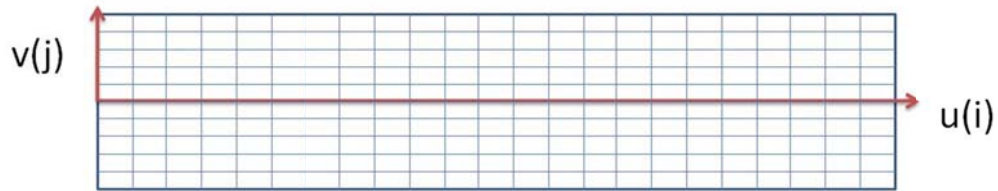
(B)

Figure 7: Creating a uniform grid (Curved Regular Grid) in the horizontal plane: A) cloud data measured in vehicle centered coordinate system B) CRG example.

The discretized terrain surface is defined as a sequence of vectors of terrain elevation as shown in Figure 8. Each vector is oriented perpendicular to the direction of travel, nominally corrected for body motion, and indexed by i (VIRES October 2008).

The number of discrete transverse locations and the corresponding length of vector v will vary according to the measurement system. If single point lasers are used, then the length of v will equal the number of lasers. Alternately, if a scanning laser is used that measures a full lane width at irregularly spaced intervals, then the length of v depends on the sampling interval for the spatial interpolation. Now consider several measurements of the terrain surface. Each measurement is considered a realization of an underlying stochastic process. An additional index, k, is introduced to indicate a particular realization of the process; the total number of realizations being considered is r, so that k takes on values 1, 2, …, r. The motivation for defining the terrain in this manner is for use in removing drift from the terrain measurements and for analyzing 3D terrain surfaces for developing new surface modeling approaches. The next section reviews terrain roughness indices as they apply to analyzing longitudinal profiles, however, in this work they will be used to analyze the components of terrain.
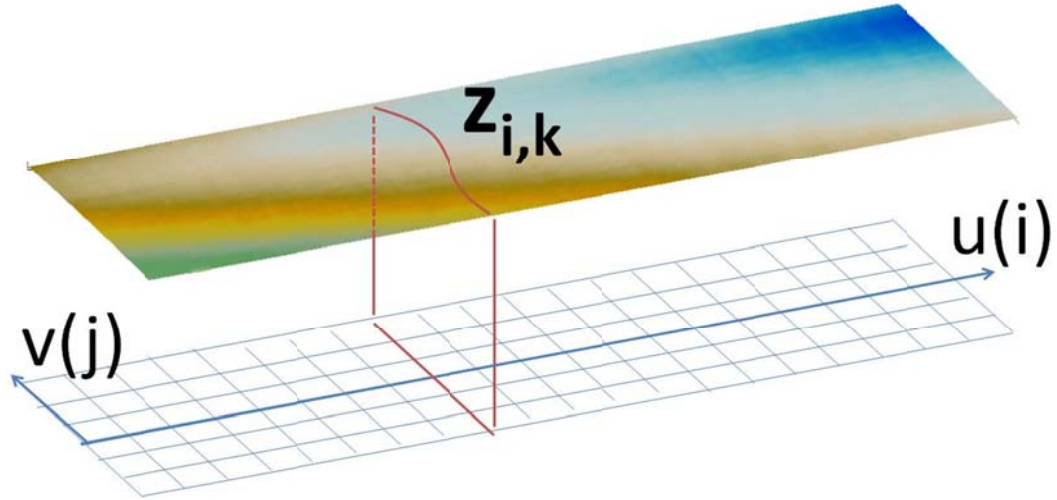
Figure 8. Defining the surface coordinate system (u, v, z) on a CRG format.

2.1.5 Terrain Roughness Indices

Terrain roughness is the most important indicator of pavement performance because it directly affects the way in which pavements serve the traveling public. Terrain roughness impacts vehicle dynamics; road maintenance; and road users in terms of health, safety, travel times, vehicle operating costs, and many other facets. It is defined by ASTM E867-82A as "The deviations of the surface from a true planar surface with characteristic dimensions that affect vehicle dynamics, ride quality, dynamic loads and

drainage" (ASTM 2006). It is vital to consider the road and travelling vehicle as inseparable parts of a unique complex system affecting the dynamics and loading of both. Roughness indices are measures of unevenness of the pavement in relation to vehicle response and, as such, are used as metrics to judge the ability of stochastic models to characterize the terrain.

Clearly, two roads that share a single roughness index need not have the same characteristics, but they share some general measure of roughness. The use of a single number to describe the overall condition of a road has been questioned, and other approaches for classifying terrain roughness using newly proposed indices or a combination of indices have been proposed (Sayers and Karamihas 1996; Papagiannakis and Raveendran 1998; Liu and Herman 1999; Ahlin, Granlund et al. 2004; Rouillard 2004; Capuruço, Hegazy et al. 2005; Kropáč and Múčka 2005; Kropáč 2008). Since roughness indices statistically describe the unevenness of the terrain, a major application is to judge the quality of the synthetic terrain produced from the terrain models. The index values for the synthetic terrain and the measured terrain can be computed and the distributions of the index for the synthetic and measured terrain then compared. Similarities in the distributions indicate how well the model captures the terrain unevenness in the synthesized data set, insofar as a single roughness index is able to characterize terrain unevenness. Discrepancies in the distributions would suggest that the terrain model is incapable of sufficiently modeling a particular terrain and a different terrain model should be used. In this work, the roughness indices will be used to analyze the principle components of terrain and compare the results with the roughness index of a longitudinal profile taken from the terrain surface and draw conclusions.

2.1.3.1   Quarter-Car Roughness Indices

The indices presented in this Section calculate the terrain roughness with some mathematical operation on the accumulated suspension stroke over the total distance traveled as determined by a quarter car model, as shown in Figure 9. The sprung mass, M_s, is the portion of the vehicle body mass supported by one wheel. The unsprung mass, M_u, is the mass of the wheel, tire and half of the axle/suspension at one corner of the vehicle. The tire spring rate is depicted by K_t, and the suspension spring rate is depicted

by K_s. The damping coefficient of the suspension system is C_s. The measure of suspension travel over time is show as *z(t)*, and the distance traveled by the model is L. The parameters chosen for this model are Golden Car parameters, please refer to Appendix B for details. The parameters do not describe an average American passenger as commonly thought; rather the parameters were selected to match the sprung mass response of a typical passenger vehicle. Specifically, the spring rates were selected to match the two major resonant frequencies, body and axle bounce, with the damping coefficient much higher than in most cars and trucks to improve correlation with a wide variety of response-type measurement systems (Sayers 1995).
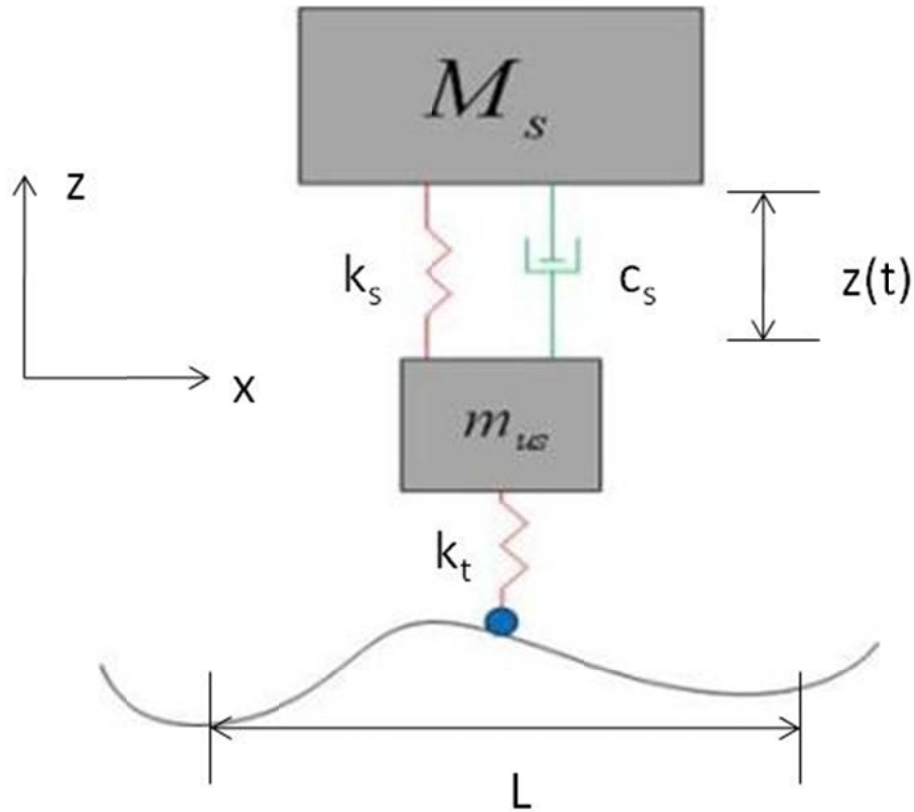
Figure 9. Quarter-car model used to calculate roughness indices based on inertial profiling devices.

International Roughness Index (IRI)

The standard index used for determining road roughness is the International Roughness Index (IRI) (Sayers 1989; Sayers 1995). The IRI has been adopted by road agencies worldwide for use in pavement health monitoring and determining the best

construction practices to develop the smoothest roads possible. It is considered an indicator of pavement serviceability, independent of the particular equipment used to measure it, internationally and geographically transferable, and time invariant (Hajek, Kazmierowski et al. 1998). The IRI is computed from a single longitudinal profile with a measurement resolution of 0.5mm to sufficiently capture the characteristics of the terrain of interest. The terrain profile is assumed to have a constant slope between sampled elevation points and is smoothed with a moving average having a base length of 250mm. The smoothed profile is then filtered using a quarter car model at a simulated speed of 80 km/hr. The accumulated suspension motion is linearly accumulated and divided by the profile length to yield the IRI, having units of elevation per distance (mm/km or in/mi), as shown in Equation 1, with detailed calculations provided in Sayers (Sayers 1989; Sayers 1995). In Equation 1, $\dot{z}_s$ is the motion of the sprung mass [m/s], $\dot{z}_u$ is the motion of the unsprung mass [m/s], L is the length of the profile [m], and V is the velocity at which the simulation was executed [m/s]. This index will be used for analyzing the principle components of terrain in Chapter 4 of this dissertation.

$$IRI = \frac{1}{L} \int_0^{L/V} |\dot{z}_s - \dot{z}_u| dt \qquad \text{Equation 1}$$

Ride Number (RN)

The algorithm developed to calculate the Ride Number (RN) is very similar to that of the IRI, with different coefficients, initialization and averaging. The main difference is that the accumulated suspension travel is calculated with the Root Mean Square (RMS) for the RN rather than the mean absolute since the RN decomposes the profile into three wavebands, represented by three profile index (PI) statistics. Specifically, the PI is calculated for each waveband, and for each wheel path, implementing the same quarter car simulation with different parameters; see Appendix B for the parameter values. The result is a single PI for the terrain set, determined by the RMS of the wavebands for each wheel path. Further details of the calculations of the RN can be found in Sayers (Sayers and Karamihas 1996). While the IRI and RN are very similar, it is suggested that the RN does a better job at identifying pavement conditions as compared with the IRI despite using a similar algorithm due to the waveband approach.

The computational effort is equivalent for both indices with the main equation for the RN presented in Equation 2. RN is an exponential transform of the RMS slope statistic PI with dimensionless units of slope [m/m] (Sayers and Karamihas 1996).

$$RN = 5e^{-160(PI)}$$
Equation 2

Roughness Index for Driving Expenditure (RIDE)

The Roughness Index for Driving Expenditure (RIDE) is based on the sprung mass acceleration response of a reference vehicle to the pavement profile. Similar to IRI, the index is based on results from a quarter car (QC) model where a transfer function (TF) is developed which relates the QC response to that of the reference vehicle. By using a QC model, the roughness index is insensitive to specific mechanical faults of a reference vehicle. RIDE is calculated from the power spectral density (PSD) of the sprung mass acceleration of the reference vehicle, as depicted in Equation 3, with full details of the calculations provided in Papagiannakis (Papagiannakis and Raveendran 1998).

$$RIDE = \frac{1}{L}\sqrt{\int_{-\infty}^{\infty} PSD(f)df}$$
Equation 3

The index is purported to have several advantages: it is directly related to riding comfort; it is related to the dynamic axle loads and reflects heavy vehicle ride, damage, and pavement damage; it is fully compatible with ISO standard 2631 (ISO 2004); and it is sensitive to pavement roughness excitation frequencies that are close to resonant frequencies of the sprung mass and is suited for identifying pavements that need rehabilitation (Papagiannakis and Raveendran 1998).

2.1.3.2   Five DOF Model Based Index

The index presented in this Section only considers vertical translations of a simplified full vehicle model, as shown in Figure 10. Roll, pitch and yaw are neglected for simplification of the analysis. The motivation for this index is that it better reflects the ride quality of a particular vehicle as compared with a QC model. This index also

allows for implementation of representative parameters true to the vehicle being studied, as opposed to the Golden Car parameters.
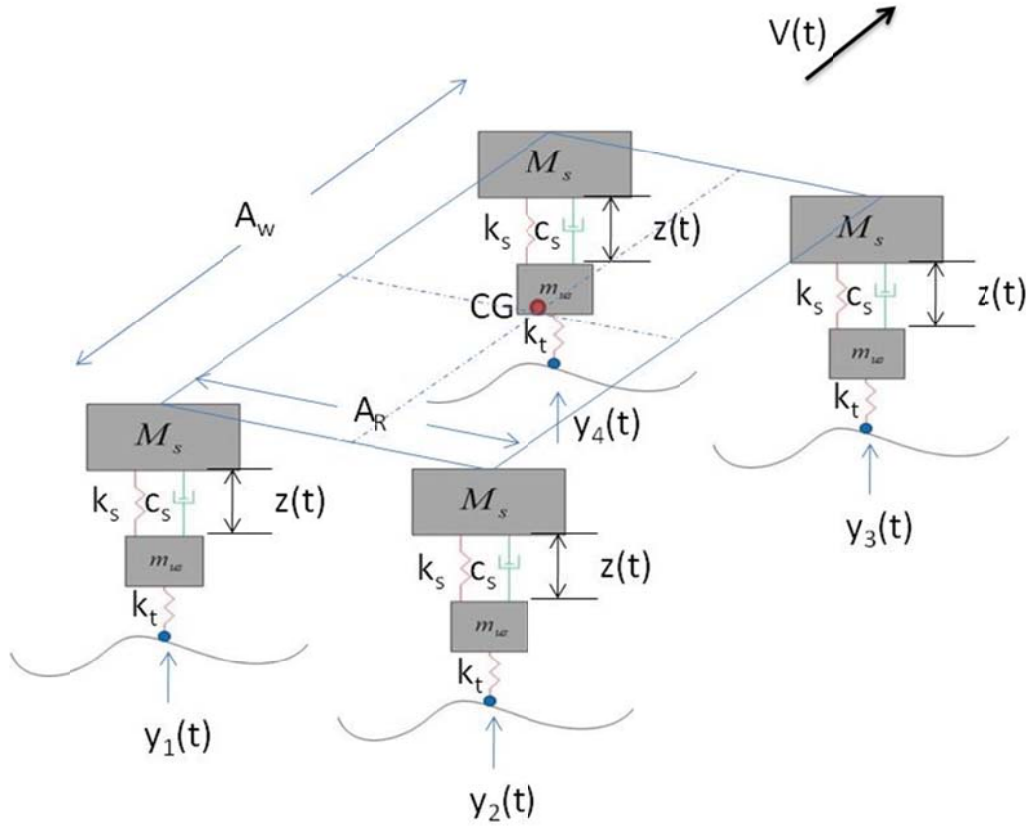
Figure 10. Full-car Roughness simulation model.

Full-car Roughness Index (FRI)

While IRI is the most popular single index, it does not consider high speed vehicles or trucks, and therefore, IRI should only be used in conjunction with another indicator if other criteria are desired to be studied. The Full-Car Roughness Index (FRI), developed by Capuruco, claims that the full-car model more accurately represents vehicle responses than a single QC model (Capuruço, Hegazy et al. 2005). The FRI combines four QC models connected by front, rear and longitudinal axles located on a rectangular frame, restricted to vertical displacements. Only the bounce forces and corresponding displacements are of interest in calculating pavement roughness, and thus only the vertical dynamics of the full-car model are studied. The major result of this study is that roughness indices are most affected by vehicle speed and tire footprint. The shortcomings of the FRI lie in the simplifications of a full vehicle model, for example, it

31

is assumed that the center of mass of the vehicle is at the centroid of the rectangular frame instead of closer to the powertrain location. Of course, when a specific vehicle is identified for simulation, a more representative result would be derived from a properly parameterized, higher fidelity, vehicle model (Capuruço, Hegazy et al. 2005). The FRI involves more computational effort, but results in higher repeatability of calculating the index over similar terrains, making it more useful than the IRI for higher fidelity simulations.

2.1.3.3   Sprung Mass Acceleration Roughness Indices

The indices presented in this section are based on the calculation of roughness indices using response-type measurement responses. Response-type profilers are typically equipped with an accelerometer located on the sprung mass of a vehicle or trailer. The measured responses are system specific (not transferrable) and temporal- the mechanical systems change with usage and time (Hajek, Kazmierowski et al. 1998). These types of systems are very limited in their capabilities of describing universal roughness, but are useful in understanding how a particular vehicle is affected by the terrain it is traversing. This method is a better for ride quality evaluations, rather than calculating a roughness index.
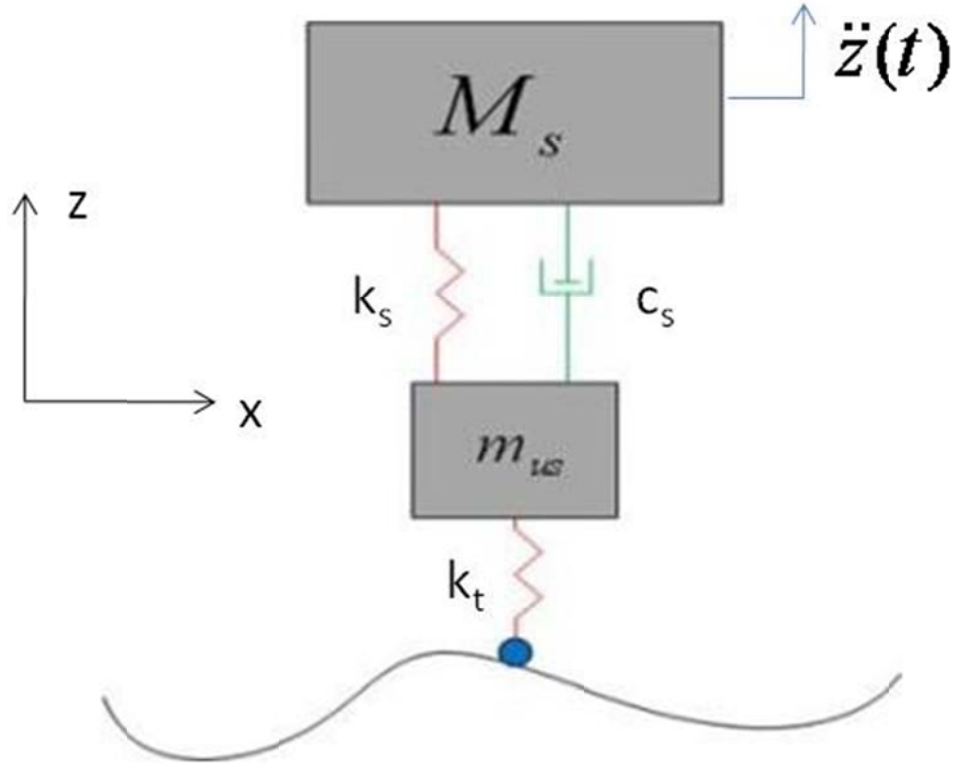
Figure 11. Quarter-car model used to calculate roughness indices based on response-type measurements.

Ride Condition Rating (RCR)

The Ontario Ministry of Transportation (MTO) uses a response-type road roughness measuring system that measures roughness in terms of Root Mean Square Vertical Acceleration (RMSVA) of a trailer axle (Hajek, Kazmierowski et al. 1998). The roughness data used by MTO for pavement management is expressed in terms of a Ride Condition Rating (RCR), based on a scale from zero to ten. A rating of ten represents the smoothest ride, and a rating of five or six indicates that the pavement needs resurfacing (Hajek, Kazmierowski et al. 1998). Unlike most organizations that are dubious of relying solely on the IRI, the MTO supports the IRI since it eliminates the application of response-type measurements for analyzing road roughness and replaces it with the application of longitudinal pavement profile measures.

Root Mean Square of Terrain Height (RMS$_H$)

The Root Mean Square of Terrain Height, RMS$_H$, analyzes the importance of short, medium and long road unevenness of the vehicle vibration (Kropáč and Múčka 2004; Kropáč 2008). RMS$_H$ is defined in Equation 4, where D$_H$ is a single measure of the combined variance of the random component of the road profile with the harmonic component of the road profile (Kropáč 2008). More information on the specifics of calculating RMS$_H$ can be found in the article by Kropáč and Múčka (Kropáč and Múčka 2005).

$$RMS_H = \sqrt{D_H}$$
<div align="right">Equation 4</div>

It is important to decompose the analysis of road unevenness to different wavelengths because the vehicle responds differently to certain frequencies. The vehicle response to road unevenness indicated that some of the tested roads yielded a much higher vertical acceleration than others, so much so that the allowable exposure to vibration for people in the vehicle would be reduced by several hours (ISO 2004). This study found that the IRI was insensitive to roughness caused by unevenness in certain wavebands. From these findings, Kropáč and Múčka proposed the use of RMS$_H$, as an alternative to IRI (Kropáč 2008).

Perceived Vehicle Roughness

Ahlin focused his studies on vertical vibration for a single wheel track to understand how well a synthesized terrain profile matches perceived vehicle wheel roughness, or more commonly known as vertical vibrations. For this test, vertical vibrations were measured on the right wheel of the front axle of a Volvo FL12 heavy truck using an accelerometer (Ahlin, Granlund et al. 2004). Through quarter car simulations, Ahlin compared measured vibrations to simulated vibrations to determine the correlation between models. He found that the higher fidelity measurement of the terrain used for simulated vibrations, the QC model yielded better correlation results as compared with the lower fidelity measurements. Specifically, Ahlin used three single-point lasers that were setup across the wheel path of the host vehicle to show that the accuracy of the data in representing vertical wheel vibration is better than a single, single-point laser placed in the wheel path (Ahlin, Granlund et al. 2004). With the increased

data set from the three-single-point lasers, the perceived vehicle wheel roughness of the synthesized terrain was found to be more representative of the measured terrain. To further Ahlin's study, roughness indices could be applied to the measured and synthesized terrain to determine how well the synthesized data truly captures the terrain characteristics. Ahlin's approach will be implemented as a 'check' to determine how well the synthesized terrain surface compares to the measured data in this project in Chapter 5 of this dissertation. The next section of this Chapter leverages empirical terrain measurements to develop analytical terrain.

2.2 Analytical Terrain

Analytical terrain can be separated into two categories, those based on a more global approach such as approximating large areas of land with surface fitting by contour lines (~ 1000km X 1000km) and those based on a more local approach by characterizing terrain profiles (~1 km). Both categories use parameter identification to parameterize the mathematical model chosen to characterize the terrain of interest where the key differences lie in what the mathematical models are characterizing. In the first category, global elevation is characterized based on surface estimation via elevation contour line measurements, typically represented by sets of polynomials or Fourier series. The purposes of such characterizations are typically for data compression and global terrain analysis, such as hydrological flow or erosion analysis (Moore, Grayson et al. 1991; Quinn, Beven et al. 1991). The second category is geared towards modeling segments of longitudinal profiles over relatively short distances. The purposes of these characterizations are to synthesize road segments based off the original terrain of any desired length, and also for data compression.

2.2.1 Global Terrain Approximation

Global terrain approximation starts with discretely measured terrain surface data and uses polynomials or power series to approximate the surface. In this way, a continuous curve is developed to represent the area of terrain. Most terrain surface approximation techniques originated in DTM. Two major drawbacks of the methods described are the coarse resolution and large amount of filtering which yield a basic

description of the surface, with flaws occurring in the representation of 'sharp' terrain features such as curbs, cliffs, rocks and others. Three terrain surface approximation techniques are described next.

Active Contour

Active contour mapping is used when modeling ground surface contours for creating DTM from airborne scanning data. Active contours have roots in the area of image processing. The shape of an active contour to a particular ground surface is the contour which minimizes an energy function. The energy function is typically parameterized by internal energy as described by the physical characteristics associated with the contour (material properties) and a potential energy field as described by the terrain height (Elmqvist 2001; Elmqvist, Jungert et al. 2001). One issue that may arise is that the solution may solve for a local minimum and not a global minimum. Another issue is properly identifying the material properties of the terrain surface in question. Interestingly, different terrain features can be identified or removed from the data depending on the requirements of the application. For example, for rocky terrain, it may be desired to estimate the underlying terrain and not the rocks, or it may be desired to include the rocks in the true terrain surface. This method can be adapted for either case such that the resulting terrain surface representation will be a continuous model of the desired true terrain surface. However, due to the nature of this method, it is prone to over-smoothing sharp-edged features such as rocks.

Linear Prediction

Pfeifer and Kraus developed an iterative robust interpolation algorithm based on linear prediction to estimate the digital terrain surface. This algorithm is comprised of two components, a stochastic model to define the weighting function and a functional model to describe the interpolation scheme, both of which will be described in the next paragraphs. As identified previously in this Chapter, measurements acquired from airborne laser scanners contain information about vegetation, buildings and the terrain. The ultimate goal of this approach is to remove vegetation and buildings from the data set and be left with the original terrain surface.

36

The first component of this algorithm determines the weighting function of each elevation point. Each point is assigned a weight from zero to one based on the classification of the point as a ground point or a vegetation point. For example, if the point is located on or near the ground plane, then a weighting close to one is assigned, and if a point is considered vegetation, or very far from the ground plane, then a weight of zero is assigned. The specific weighting function used in this work is shown in equation form in Equation 5 and in graphic form in Figure 12. The weight function, p_i, for robust estimation is presented in Equation 5, where the parameters a and b determine the steepness of the weight function, g is a shift value, and w is a 'ceiling' elevation value and is user defined, such that any value above $g+w$ is set to zero (Kraus and Pfeifer 1998; Pfeifer, Kostli et al. 1998). An average is computed on the surface to calculate the residuals, v_i between each height measurement and the average surface. The shift value, g forces the left branch of the weighting function equal to one with details on how to calculate the shift value, g, explained in Kraus and Pfeifer (1998) and Pfeifer et al. (1998) (Kraus and Pfeifer 1998; Pfeifer, Kostli et al. 1998). If the residuals are less than a shift value, g, they are typically negative and classified as terrain points with weighting set equal to one. If the residuals are between the shift value g and the shift value plus some offset, $g+w$, they are typically small negative or positive and are classified as lower lying vegetation or rocks, with a weighting value occurring between zero and one. If the residuals are greater than the shift value plus some offset, then the points are classified as vegetation or buildings and the weighting value is set equal to zero. Furthermore, the parameters presented in Pfeifer's work can be tuned to keep or eliminate specific terrain features.

$$p_i = \begin{cases} 1 & v_i \leq g \\ \dfrac{1}{1 + (a(v_i - g)^b)} & g < v_i \leq g + w \\ 0 & g + w < v_i \end{cases}$$
   Equation 5
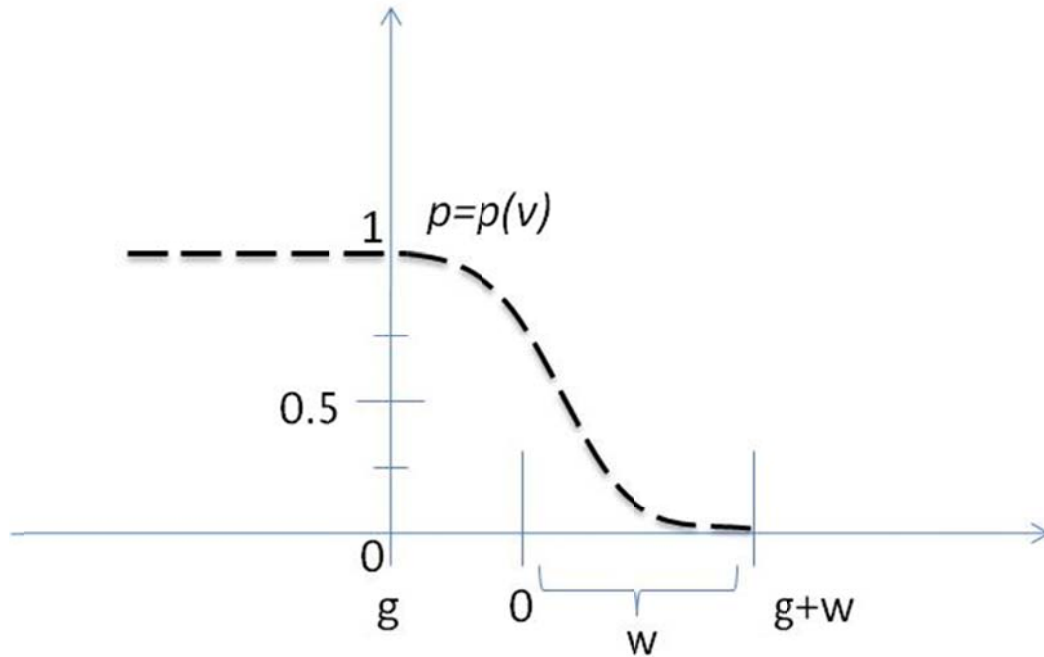
Figure 12. Residual weighting function.

In the second component of the algorithm, linear prediction is used to identify an accuracy value for each elevation grid point in a similar method to kriging and moving least squares.   This interpolation method is based on describing the surface with a covariance function (Pfeifer, Stadler et al. 2001). The covariance function measures how much two variables change together and describes the variance of a random variable process or field. In this case, the two variables are the X and Y elevation point locations. The end result is not a perfect terrain representation, as 'sharp' terrain features such as trenches and curbs are smoothed out and buildings or trees are not always eliminated, but it does provide an accurate representation of the terrain for a large land base.

Fixed Grid Polynomials

True terrain surfaces are continuous functions, which due to technological limitations are discretized by measurement systems. Measurement systems acquire data at different sampling rates and resolutions, as described in the earlier part of this chapter. The problem Segu sought to solve was to regenerate the continuous terrain functions with continuous polynomials based on discretized measured terrain data that has been gridded based on specific terrain features (Segu 1985). The first step is to take the measured

surface data and grid it based on terrain features. For example, if an area of the surface is rough, then the gridding will need to have more nodes than if the surface is relatively flat and uneventful. Then the discretized terrain height can be approximated as of function of x and y location as shown in Equation 6, where a_r are coefficients. Equation 6 can be abbreviated in power series form as shown in Equation 7. The importance of this representation is that Equation 7 takes on the same form as a double Chebyshev polynomial, which provides a least squares solution for surface fitting. The coefficients of double Chebyshev polynomial are determined from the measured data, and continuous contour lines describing the surface are developed. Of course the methods presented by Segu only considered the implementation of Chebyshev polynomials, but that is not to say that cubic splines, linear interpolation, or other polynomials or power series could not be used instead. Loucks suggests that instead of being concerned with the least squared solution only, grid spacing and desired resolution need to be taken into account. He suggests using cubic splines for finer resolution data to make a greater effort in capturing more abrupt terrain features such as cliffs and other drop offs and linear interpolation or Chebyshev polynomials for coarse resolution applications where an exact fit to the terrain data is not required (Loucks 1996). Regardless of the approximation tool used, the methods developed in Segu's work are used as a baseline for the definition of 3D terrain surfaces presented in Chapter 4 of this dissertation. The key differences are that this method is used for developing contour elevation lines, whereas the methods presented in Chapter 4 decompose the terrain surface to its principle components and then approximate the components with sets of polynomials. The next section presents different methods to characterize the terrain with mathematical models, that when synthesized, capture the inherent statistical properties of the host terrain.

$$z = a_0 + a_1 x + a_2 y + a_3 x^2 + a_4 xy + a_5 y^2 + \cdots \qquad \text{Equation 6}$$

$$z = \sum_{i=0}^{k} \sum_{j=0}^{i} a_r x^{i-j} y^j, \ \ r \in 0,1,2,\dots,m' \qquad \text{Equation 7}$$

2.2.2   Local Terrain Characterization

If the terrain is considered to be a particular realization of an underlying stochastic process, then measured terrain profiles can be used to develop the form and the

order of the mathematical description of the stochastic process. A large set of terrain profiles is then represented by a single stochastic process that has the same properties (e.g., stationarity) as the measured terrain. A measured terrain profile is used to parameterize the model, and the parameters are then defined as the terrain characterization. Once the model is properly parameterized, realizations of any length can be generated, and vehicle responses can be simulated that are accurate for the given application. Equation 8 defines the stochastic process Z_n as a family of random variables defined over a common probability space.

$$\{Z_n, n = 0,1,2, \dots\} \qquad \text{Equation 8}$$

These random variables correspond to discrete locations along a prescribed path in the horizontal plane and are indexed by a real parameter $n \geq 0$. Let S denote the state space so that realizations of the process at index n, Z_n, are elements of S.

Several methods have been examined for characterizing terrain profiles in this dissertation. Extensive work has been conducted in using the Power Spectral Density (PSD) of road profiles (Andren 2006; Kutay, Weaver et al. 2007) and road roughness metrics (Gorsich, Gunter et al. 2003). Rouillard, Bruscella, and Sek suggested a method based on PSD analysis and spatial acceleration to characterize and classify road profiles (Bruscella and Rouillard 1999; Rouillard, Bruscella et al. 2000; Rouillard, Sek et al. 2001). They used spatial acceleration to separate the stationary terrain profile characteristics from non-stationary transient events. Later, Ferris examined the viability of characterizing roads as an nth order Markov chain (Kuchar 2001; Ferris 2004). Recently, Kern, Wagner and Ferris demonstrated that an Autoregressive Integrated Moving Average (ARIMA) model could be used for modeling road data and generating synthetic road profiles that have similar statistical properties to those of measured roads by defining a linear (AR) model excited by a stationary non-Gaussian process (Kern and Ferris 2006; Kern and Ferris 2006; Kern and Ferris 2007; Li and Sandu 2007; Wagner and Ferris 2007; Khashei, Bijari et al. 2008; Wagner and Ferris 2008). These models, in addition to several others, have expanded the applicability of computational simulation in the ground vehicle design and development process. In the remainder of this section,

typical model properties are explained, and then several models of interest are investigated.

2.1.3.4   Model Properties

Terrain models are formulated based the statistical properties of the terrain, including one or more of the following: stationary, isotropic, homogenous, or Gaussian. Chaika states that rigorous statistical tests must be performed before a statistical model is created, but first, the properties must be defined (Chaika, Gorsich et al. 2004). When examining these common model properties on measured terrain, these results were found.

1. Long sections of terrain profiles are not stationary,
2. Short sections of terrain profiles may not be stationary depending on their content and characteristics.
3. Generally smooth roads are stationary, but rough roads are not.

Understanding the properties of measured terrain is essential for accurate terrain modeling.

Stationary and Homogeneity

Stationarity of a time series implies that all joint probability density functions are time invariant. For example, the variance of stationary signals will be time invariant. (Bruscella and Rouillard 1999). However, the term stationary is commonplace for time series analysis whereas terrain does not vary in time, rather location. The confusion arises in the fact that time series analysis techniques are applied to terrain analysis. The appropriate term to use in terrain analysis is homogeneous. A terrain profile is homogeneous if the joint probability density functions are invariant with horizontal translations in the coordinate frame (Bendat and Piersol 2000). This is analogous to being stationary in time, but it is in the spatial domain. Many studies have shown that smooth roads may be homogenous, but rough roads are not due to localized disturbances (Bruscella and Rouillard 1999; Chaika, Gorsich et al. 2004; Andren 2006).

Isotropic

A terrain profile is isotropic if the joint probability density functions are invariant with changes in direction or orientation of the coordinate frame (Frederick and Chang

1965). Therefore, this property examines the terrain in all directions. For example, a terrain surface with wheel rutting would not be considered isotropic, but rather anisotropic.

Gaussian

A Gaussian distribution is a normal distribution with the mean, μ, and variance, σ^2 (Bendat and Piersol 2000). Equation 9 and Figure 13 demonstrates a Gaussian distribution.

$$P(z) = \left(\sigma_z\sqrt{2\pi}\right)^{-1} exp\left[-\frac{(z-\mu_z)^2}{2\sigma_z^2}\right]$$
Equation 9
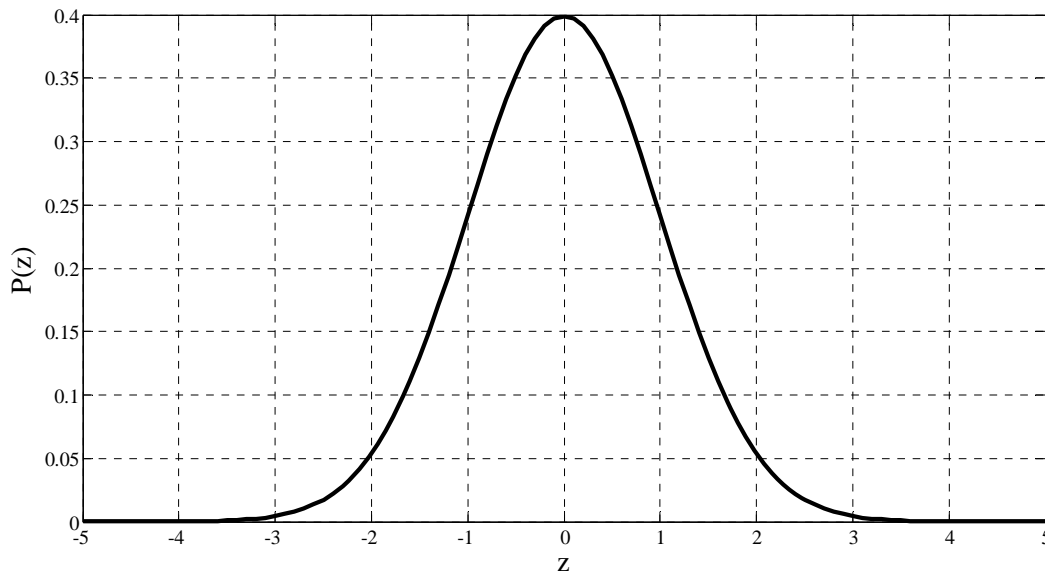
Figure 13: Gaussian Probability Distribution Function.

The Gaussian distribution is a commonly used assumption in terrain characterization and realization; however, this assumption may not be valid in all cases. If localized disturbances (e.g., potholes) are removed from terrain profiles, the residual profile may be nearly Gaussian (Andren 2006). Some studies have transformed the actual, non-Gaussian distribution to a Gaussian distribution, then performed the statistical analysis on the transformed data, and finally, applied an inverse transformation to return to the original data set (Bogsjo and Forsan 2004)**.**

2.1.3.5   Terrain Models

Several terrain models are presented for discussion of the implementation and pros and cons of each. Ultimately, one model will be selected as a proof of concept for modeling the principle components of terrain in Chapter 5 of this dissertation.

Power Spectral Density (PSD)

One of the more popular models is the Power Spectral Density (PSD) approximation of terrain profiles. The PSD identifies the frequency content of a stochastic process or a deterministic function of time to identify frequencies of interest and periodicities. Andren presents a literature survey of PSD approximations which implement the basic equation of this model, as shown in Equation 10, where G_d is the roughness level of the displacement PSD, C is the general roughness parameter with units of *meters*^{3}, λ is the wavelength with units of *meters*, and w is a dimensionless parameter called waviness, ranging from 1-5 on measured roads, and is related to the wavelength of the road (Andren 2006). Traditionally, w is approximately equal to 2 and C varies from road to road.

$$G_d(n) = C\lambda^{-w}$$
<div align="right">Equation 10</div>

This approximation represents a terrain profile by two numbers, C and w, the resulting PSD approximation is a straight line. This simple concept can be expanded by decomposing the spectrum into wavebands (commonly two or three). In this way, a common value for the roughness parameter C is assumed to remain constant and different values for w can be fit to each waveband.

Equation 11 is an example of a three waveband approximation.

$$G_d(n) = \begin{cases} C\lambda^{-w1} & 0 \leq \lambda \leq \lambda_1 \\ C\lambda^{-w2} & \lambda_1 \leq \lambda \leq \lambda_2 \\ C\lambda^{-w3} & \lambda_2 \leq \lambda \leq \infty \end{cases}$$
<div align="right">Equation 11</div>

For a spectral analysis estimate to be valid, the road should be a member of a stationary random process. When terrain is homogeneous and isotropic, a complete terrain surface can be modeled with a single terrain profile PSD. Several variations of the PSD are presented next. The data used in the analysis of the following PSDS were filtered such that long-wavelength disturbances (such as hills) were removed by applying a high-pass

filter to keep wavelengths between 0.1m and 100m (ISO8608 1995). It is important to note that the majority of terrain analysis using PSDs is completed incorrectly as it validates the stationary requirement.

Parametric Road Spectrum (PRS)

Bogsjo presented four different models to parameterize road spectra following a Gaussian distribution: parametric road spectrum, shifted spatial range spectrum, direct spectrum estimation, and transformed direct spectrum estimation. The Parametric Road Spectrum (PRS) model is a straight line fit in the log-log scale of a power spectrum. Equation 12 describes the PRS approach, where n_O is 0.1m^{-1}, C indicates the degree of road unevenness, varying around 10^{-5} and the exponential value w varies from 2-3. The parameter values were determined by a least-square fit over the spatial frequency range of $0.011\text{-}2.83\text{m}^{-1}$ (Bogsjo and Forsan 2004).

$$R_{PRS}(n) = \begin{cases} C*\left(\frac{n}{n_O}\right)^{-w} & 0.01 < n \leq 10 \\ 0 & otherwise \end{cases}$$

Equation 12

Shifted Spatial Range Spectrum (SSR)

The Shifted Spatial Range spectrum (SSR) model focuses on correctly approximating the low resonance frequency range of the vehicle, i.e. 1-2Hz. The SSR is defined over the least square fit of the range corresponding to the spatial frequency range for a vehicle traveling at constant velocity, defined as: $\frac{1}{v} < n \leq \frac{2}{v}$. The SSR spectrum is then defined over this range, in Equation 13, where n_O is 0.1m^{-1}, C indicates the degree of road unevenness, varying around 10^{-5} and the exponential value w varies from 2-3. The lower frequency limit of the spectrum is adjusted to preserve the variance of the measured profile as defined in Equation 14 (Bogsjo and Forsan 2004).

$$R_{SSR}(n) = \begin{cases} C*\left(\frac{n}{n_O}\right)^{-w} & n_{start} < n \leq 10 \\ 0 & otherwise \end{cases}$$

Equation 13

$$n_{start} = max \begin{cases} (10^{1-w} - \frac{1-w}{Cn_0^w}\sigma^2)^{\frac{1}{1-w}} \\ 0.01 \end{cases}$$

Equation 14

Direct Spectrum Estimation (DSE) & Transformed Direct Spectrum Estimation (TrDSE)

The Direct Spectrum Estimation (DSE) model smooths the power spectrum of the actual road; the smoothed road spectrum is then used in the model to develop synthetic road profiles. A slight variation of the DSE is the Transformed Direct Spectrum Estimation (TrDSE), which applies an empirical transformation function to the road profile, and the resulting transformed road spectrum is determined by the DSE model. This transformation procedure assumes strictly stationary data (Bogsjo and Forsan 2004).

When applied to a QC simulation model to analyze pseudo-damage, the PRS and DSE models underestimated the fatigue damage in all cases, the SSR model performed better at velocities that were not near vehicle resonance, and the non-Guassian TrDSE model overestimated the fatigue damage, resulting in very conservative fatigue estimates. From these studies, it was found that traditional stochastic analysis using Gaussian distributions is insufficient in characterizing the spectra of the terrain that is critical in predicting fatigue damage. This is likely due to the fact that fatigue is sensitive to random, localized disturbances (Bogsjo and Forsan 2004). The conclusions identified in this report will be considered for future ground vehicle reliability studies but in the case of this situation, PSDs will not be applied as a stochastic modeling technique.

Markov Chains

Measured road profiles can be considered a realization of a discrete-time, finite-state, closed, irreducible and aperiodic process (Ferris 2004). The Markov Property is written in Equation 15.

$$P(Z_{n+1} = z_{n+1}|Z_n = z_n, Z_{n-1} = z_{n-1}, \dots, Z_0 = z_0) = P(Z_{n+1} = z_{n+1}|Z_n = z_n) \qquad \text{Equation 15}$$

That is, conditioned on the present, the future and the past are independent. The Markov Chain has stationary transition probabilities when Equation 8 holds for all values of *n*. The process defined in Equation 15 forms a Markov Chain when these random variables, X_n, satisfy the Markov Property and they have stationary transition probabilities. The transition function of the chain is defined in Equation 16.

$$P(s_1, s_2) = P(Z_{n+1} = s_2|Z_n = s_1), \quad s_1, s_2 \in S, \ n \geq 0 \qquad \text{Equation 16}$$

The Markov Property implies that regardless of how the process arrived at state s_1, it has probability $P(s_1, s_2)$ of being in state s_2 at the next measurement point. The *n-step* transition function, $P^n(s_1, s_2)$ is the probability of going from state s_2 to state s_2 in *n-steps*. A transition matrix arises from the transition function, where each transition matrix represents a large set of road profiles that share the same characteristics (Ferris 2004). The Markov Chain representation of terrain will not be investigated in this work but may be applied to future work for its capabilities of modeling a time-invariant process.

Hilbert-Huang Transform and Empirical Mode Decomposition

The Hilbert-Huang Transform (HHT) is capable of analyzing and classifying nonstationary and nonlinear terrain data. The HHT consists of two steps, the Empirical Mode Decomposition (EMD) and the Hilbert Spectral Analysis (HSA). The EMD is a multi-resolution, local decomposition method which decomposes the terrain data into basis functions specific to the data. These basis functions are known as Intrinsic Mode Functions (IMF). IMFs have the following properties: they have instantaneous frequency defined at every point; they are almost orthogonal; they form a complete basis set (Ayenu-Prah and Attoh-Okine 2009). These basis functions are used to decompose a single longitudinal profile whereas the methods developed in this work decompose the entire terrain surface into characteristic basis vectors. Futhermore, the IMFs exist in the frequency domain, whereas the basis vectors in this work exist in the spatial domain. The EMD of the terrain data is completed when the residual process is a monotonic function, or a constant. Specific details of the EMD can be found in works by Ayenu-Prah and Attoh-Okine (Attoh-Okine, Ayenu-Prah Jr et al. 2006; Ayenu-Prah and Attoh-Okine 2009). It is important to note that during this decomposition, all information of the original signal is retained such that the summation of the IMF and the residuals return the original signal. The next step in the HHT is a Hilbert transform applied to each of the IMFs to identify instantaneous frequencies and amplitudes for use in plotting the Hilbert amplitude spectra (Attoh-Okine, Ayenu-Prah Jr et al. 2006; Ayenu-Prah and Attoh-Okine 2009). Identification of the instantaneous frequencies and amplitudes of each IMF allows for the analysis of the energy content of each IMF; the higher the energy content, the

greater effect that IMF has on the original signal. The impact of the energy associated with each IMF is not completely understood and remains under investigation. The HHT is similar to the method presented in Chapter 4 of this dissertation, in that the terrain surface is decomposed into a set of basis vectors, of which each basis vector is studied for its contribution to the overall terrain surface.

Autoregressive Integrated Moving Average (ARIMA)

Autoregressive Integrated Moving Average (ARIMA) models can be used to characterize road profiles through parameter identification of the Autoregressive coefficients. Uncertainty is captured in the resulting residual process, which, under certain circumstances, may be represented by a single probability density function. This stochastic residual process can be used to drive the creation of a distribution of synthetic terrain topology profiles. This distribution is then dependent on the probability distribution of the residuals as well as the ARIMA coefficients. ARIMA models are one method for characterizing road profiles which incorporates the uncertainty of terrain characteristics into the model structure. The ARIMA parameters are identified as *p*, *d* and *q*, where *p* is the order of the autoregressive model, *d* is the number of times the data was differenced and *q* is the order of the moving average model. It has been shown previously that terrain topology does not generally exhibit moving average characteristics, so the general form of the model becomes Equation 17, where φ is the autoregressive coefficient, a_i is the distribution of the residuals which is assumed to be independently and identically distributed for all time, and ∇^d is the backward difference operator.

$$\nabla^d z_i = \varphi_1 \nabla^d z_{i-1} + \varphi_2 \nabla^d z_{i-2} + \cdots + a_i \qquad \text{Equation 17}$$

The autoregressive coefficients traditionally are constants encompassing the correlation of the current profile point to the previous profile points (Wagner and Ferris 2008). This model will be investigated further for its capabilities to model road profiles.

2.3 Literature Review Summary

Two major categories of terrain measuring and modeling have been reviewed: terrestrial applications and airborne applications. A clear distinction should be made that

this work focuses on terrestrially acquired measurements and focuses on modeling terrain that is anisotropic (path specific). DTMs are concerned with modeling a large area at a lower resolution (elevation points every 3-15m (Maxwell 1970; Pfeifer, Kostli et al. 1998)) than what is captured with the terrestrial systems (elevation points every 4-25.4mm (Kern and Ferris 2007; Wagner, Kern et al. 2007)) due to the intended application and current limitations of airborne technology.

Furthermore, when processing airborne laser scanner data, it has been assumed that the data are sufficiently isotropically distributed (Pfeifer, Stadler et al. 2001). This is the key difference between the methods used in DTMs and the methods presented in this dissertation. The terrain measurement system used to acquire data for this research, the VTMS, directly samples the area of interest. Airborne laser scanning systems scan a broad area, and then require filtering techniques to remove unwanted vegetation or buildings before the data can be used. This filtering often results in over-smoothing of 'sharp' terrain features such as rocks, curbs, drop-offs, etc. All of these sharp, distinct terrain features are vital for vehicle durability and reliability analysis and cannot be ignored or filtered out. Additionally, the data acquired in this work is sufficiently anisotropic on the global scale so that vehicle paths are identifiable (e.g., >100mm), but on the local scale (e.g. <10mm), the data are assumed to be isotropic for grid node elevation identification.

Many terrain modeling discussions, in both the terrestrial and airborne applications, culminate in the need to model terrain on different scales (e.g., millimeter, centimeter, meter…). The methods proposed in this work propose decomposing the terrain surface into its principle characteristics, and then modeling each characteristic before reconstructing the surface, a method similar to empirical mode decomposition (EMD) (Attoh-Okine, Ayenu-Prah Jr et al. 2006). In this way terrain features such as elevation, banking and crowning can be modeled individually as opposed to directly modeling different scales of the terrain. In the limiting case, when the road is so rough that there is no clear path, then this method decomposes into DTM techniques used for forests and other unstructured environments: a very large number of basis vectors are required to represent the terrain.

Many roughness indices and terrain modeling methods have been discussed. At this time, there is no 'best' method for identifying the roughness of the road or modeling the terrain. The International Roughness Index (IRI) is the standard practice for identifying road roughness, but as can be seen by this literature review, many other methods exist and seek to change the standard. This work suggests a different implementation of the IRI from terrain profiles to evaluating the elevation component of terrain. Similarly, there is no 'best' method or suggested model for modeling specific types of terrain, as can be seen from the models presented in this literature review. This work makes use of the Autoregressive model due the capabilities of adding randomness to the synthesized profile from the residual process.

In summary, the methods developed in this work are based on data acquired from a terrestrial terrain measurement system, the VTMS. Modern terrain measurement systems use an Inertial Navigation System (INS) to measure and remove vehicle movement from laser measurements of the terrain surface. Instrumental and environmental errors inherent in the INS produce noise and drift errors in the resulting estimates of vehicle position and orientation. Chapter 3 develops a method to remove INS drift in terrain measurements. This method employs a principle component analysis through singular value decomposition (SVD) to define the 'global' and 'local subspaces. SVD of the terrain surface identifies the principle terrain characteristics (basis vectors), such as elevation, bank angle, and crowning and corresponding singular values. Next, the empirical basis vectors are approximated with a set of analytic polynomials by a Galerkin method in Chapter 4. Each set of empirical basis vectors corresponds to a specific type of terrain, and therefore a specific set of analytic polynomials are chosen to represent each terrain type. The types of terrain are identified by computing International Roughness Index (IR), the universally employed terrain roughness index values for the elevation component of terrain, which is defined as the projection of the terrain surface onto each basis vector. It is important to note that this cross-section representation of principle terrain components does not hinder the knowledge of the terrain as it had been considered to do in the past (Maxwell 1970). Previous methods assumed that the cross-section was an appropriate representation for data acquired with photogrammetrics, when it had been shown that the resolution of the data was too low to accurately describe the

surface in question. In Maxwell's work, cross-sections were taken every 3-15m. In this work, the basis vectors are projected onto the terrain surface to develop specific components of terrain which describe the magnitude of each basis vector along the length of the terrain surface at a much higher resolution, specifically every 0.0254m along the length of the terrain segment. Lastly, in Chapter 5, an Autoregressive model is employed on the components of terrain to model these components in the same manner as a terrain profile. This is one realization of a stochastic terrain surface, but as seen in this Chapter, many other modeling techniques exist. The next Chapter presents the methodology for removing INS drift in terrain measurements.

3. Correcting INS Drift in Terrain Surface Measurements

3.1 Chapter 3 Introduction

Modern terrain surface measurement systems incorporate a scanning laser (Herr 1996) that is rigidly mounted to the body of a host vehicle (Kern and Ferris 2007; Wagner, Kern et al. 2007; Liu, Dembski et al. 2008). This vehicle traverses the terrain while simultaneously acquiring terrain measurements. When the vehicle encounters a disturbance, the laser translates and rotates with the body of the host vehicle. To obtain accurate terrain measurements, the motion of the vehicle must be accurately measured so that it can be removed from the laser measurement. Modern systems use an Inertial Navigation System (INS) to measure the vehicle movement (Kennedy, Hamilton et al. 2006). The accuracy of the INS depends on the alignment of the Inertial Measurement Unit (IMU) to the laser and satellite coverage of the Global Positioning System (GPS). Typically, an INS is capable of establishing a geodetic position with two centimeter accuracy (Smith 2009). Much of the error is due to drift in both the position and orientation estimates of the INS. These artifacts of the INS drift skew the estimated height and orientation of the terrain surfaces in the global reference frame. The accuracy of the INS is dependent on instrumental biases and environmental biases, such as the number of satellites used by the receiver, the satellite-receiver distance, atmospheric effects, satellite and receiver clock offsets, phase ambiguities, and others (Doyle 1978; Sardon, Rius et al. 1994; Sardon and Zarraoa 1997; Pfeifer, Kostli et al. 1998; Briese and Pfeifer 2001; Pfeifer, Stadler et al. 2001; Kraus, Briese et al. 2004; Sarma 2008). This work does not seek to identify the origins of the error, but an overall technique to remove the INS drift error from the system as a whole.

INS drift error manifests itself in vehicle simulations when multiple lanes must be concatenated to form a surface. For example, if two adjacent lanes of a highway are measured at different times during the day, then any vertical misalignment of the two lanes resulting from the INS drift would produce an artificial vertical shift between lanes.

The misalignment between the two lanes would cause undesired excitations into a model undergoing a lane change maneuver. An example of INS drift in ellipsoidal height is shown in Figure 14A. Ellipsoidal height is the height above or below the WGS84 reference ellipsoid. The host vehicle was parked away from buildings, such that the antenna had a clear view of the sky. In this static position, several two-minute samples of INS data were collected over the course of five consecutive days where the curves in Figure 14A depict these samples of data. In a dynamic experiment, a section of road was measured 10 times at different times throughout the day. The height variation of the road can be seen in Figure 14B. It can be seen that the drift is consistent between static and dynamic experiments. These findings exemplify the results of several studies indicating that INS drift can be modeled as a random walk, zero mean stochastic process (Favey, Cerniar et al. 1999; Giremus, Doucet et al. 2004; Kim and Sukkarieh 2004; Sun, Fu et al. 2008). This model is used in the development of a new method to define, identify, and remove the INS drift to improve the accuracy of terrain surface measurements.
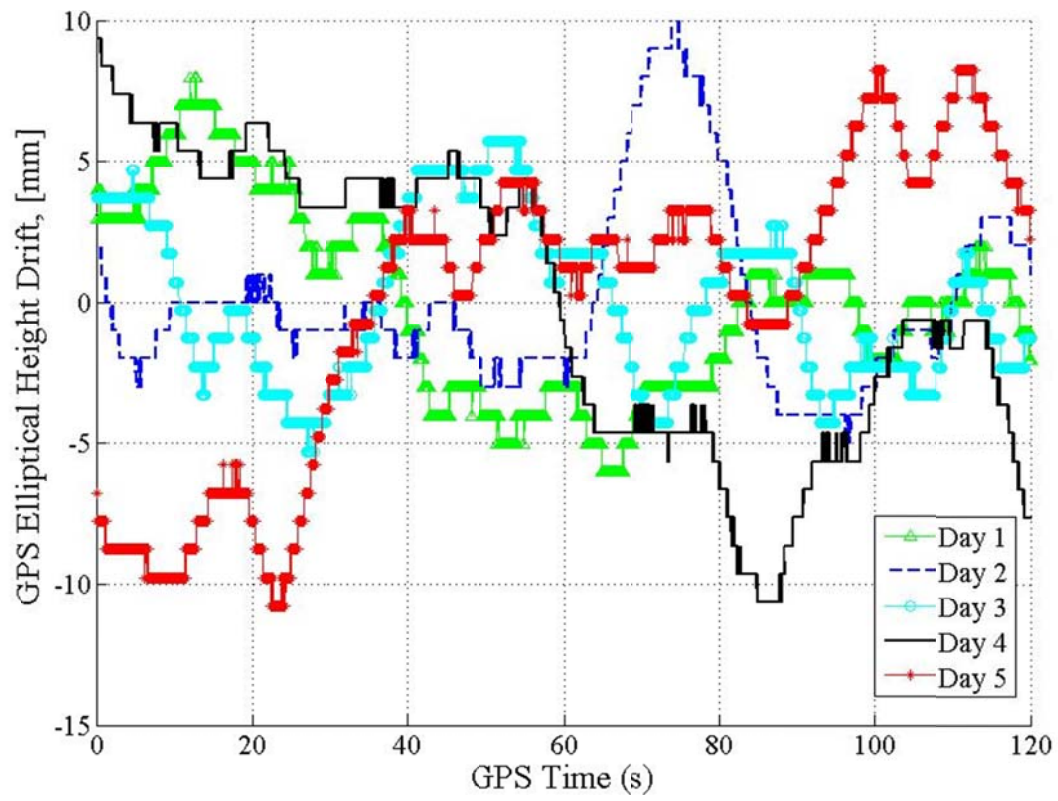
A)
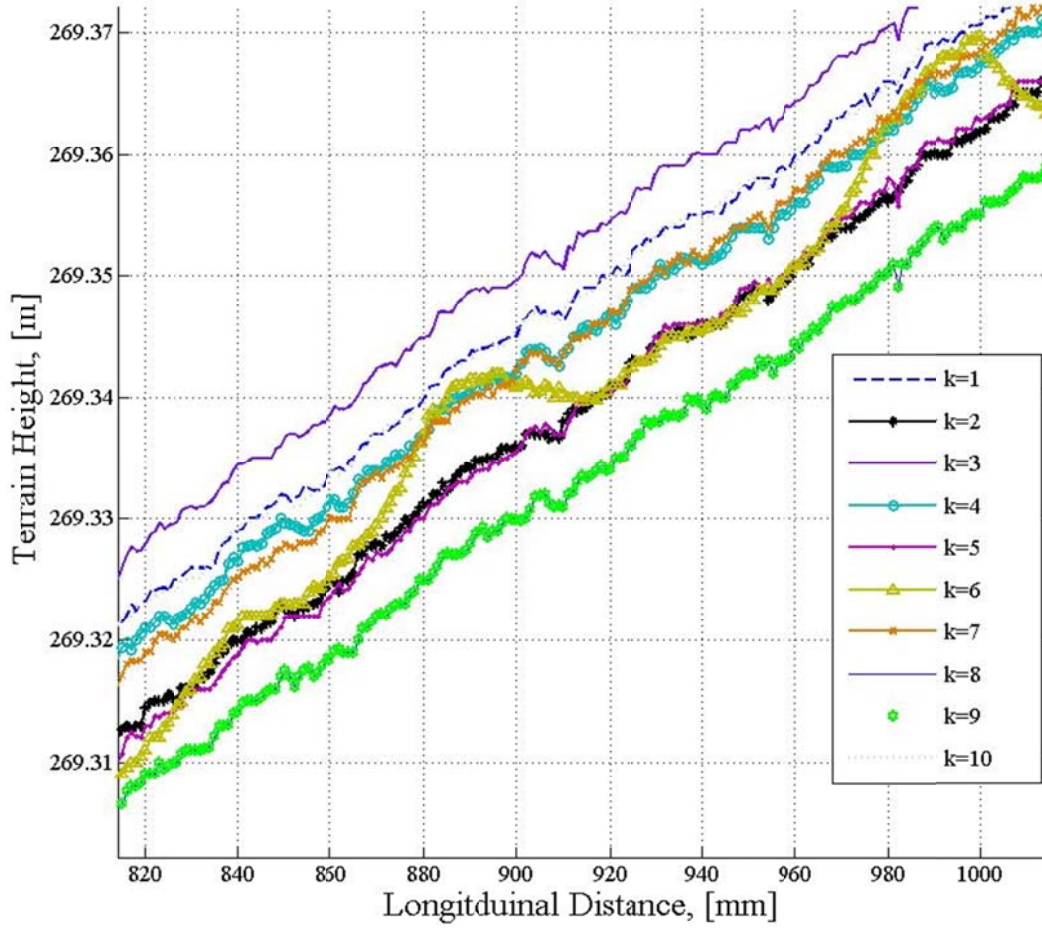
B)

Figure 14. Sample of GPS test results A) Static test results B) dynamic test results.

The objective of this Chapter is to develop a compensation method for INS drift in terrain surface measurements. The remainder of this Chapter is developed as follows. Background on body motion mitigation as associated with the Vehicle Terrain Measurement System (VTMS) is presented. The contribution of this work is then developed. A terrain surface measurement is considered as a combination of the true surface and a realization of a stochastic error surface process. Each measured surface is then a realization of a stochastic process and is defined as a sequence of vectors that are elements of a Hilbert space. The Hilbert space is composed of a global vector space and a complementary local vector space. A truncated set of orthonormal basis vectors are constructed to span the global space via a Singular Value Decomposition (SVD) of the set of measured vectors. Drift is defined as the global contribution to the error process

54

and the drift components of the surface are defined as the projection of the error vectors onto the basis vectors that span the global vector space. INS drift is then modeled as a random walk process. The true surface is developed by estimating the expectation of the sequence of measured vectors in the global vector space and the local vector space separately. The method is applied to experimental data to demonstrate the concept and to form a basis for discussion. A model order study and residual analysis is conducted on the set of terrain to support the assumptions presented in this paper, and to show that the fidelity of the resulting representation is retained to within the resolution of the measurement system. Further application of this concept is discussed in the context of INS errors, followed by concluding remarks.

3.2 Chapter 3 Background

Terrain measurement systems have evolved considerably from the early vehicle-response systems (Spangler 1962; Spangler and Kelly 1966; Gillespie, Sayers et al. 1980) to vehicle-independent measurement systems (Dembski, Rizzoni et al. 2006; Kern and Ferris 2007; Wagner, Kern et al. 2007). Early vehicle-independent measurement systems relied on accelerometer data to remove unwanted body motion from post-processed data files. The reliability of these traditional accelerometer-based systems suffers when vehicle speed falls below 5 m/s (Walker and Becker 2006) and in other low frequency environments. More recent vehicle-independent measurement systems rely on some combination of an INS, accelerometers, a distance measurement instrument, and inclinometers to remove the host vehicle's body motion from the laser data. Small misalignment between the scanning laser and the IMU will compound the error that can be anticipated from any DGPS (Smith and Ferris 2008). Recent advances have produced 3D terrain measurement systems that are capable of scanning a wide transverse path with millimeter resolution. A more detailed description of terrain measurement systems can be found in Chapter 2, Section 2.1.1 of this dissertation. The VTMS (Kern and Ferris 2007; Wagner, Kern et al. 2007), was used to acquire the data for this work.

The VTMS is equipped with a scanning laser and an INS. The scanning laser, affixed at the rear of the vehicle, acquires 941 data samples, at millimeter resolution. The data are collected transversely across a 4.2 meter wide path each millisecond, which

results in approximately one-million data points each second. The INS consists of a GPS receiver mounted to the host vehicle (rover GPS), an IMU rigidly mounted to the same platform as the laser, and a GPS remote receiver that is fixed to a stationary point near the section of terrain being measured (base station) (Wagner, Kern et al. 2007). The rover GPS estimates the location of the vehicle in time and space, with some error due to the distortion of the satellite signal from the Earth's atmosphere (Lanyi and Roth 1988; Sardon, Rius et al. 1994; Sardon and Zarraoa 1997; Mannucci, Hajj et al. 2004; Sarma 2008). The rover estimate of the vehicle location in time can be considered an estimate of the true vehicle position plus some error vector. On the contrary, due to the fixed position of the GPS base station, the true location can be determined accurately. Combining the solutions from the rover and base station GPS solutions results in a differential GPS (DGPS), which is more effective at minimizing the error in the vehicle position. The IMU is comprised of a three-axis gyroscope to measure the roll, pitch and yaw of the vehicle to estimate the orientation of the vehicle in space and three-axis accelerometers to measure the translational accelerations of the vehicle. Once the raw data have been collected, the measurements from the subsystems of the INS are combined using proprietary post-processing software to obtain an accurate inertial navigation solution. The information defining the position and orientation of the vehicle is denoted as the number 1 in the schematic of Figure 15 with the information defining the laser measurement denoted as the number 2 in the schematic. Coordinate transformations convert the INS-centered data to laser-centered data, and then both signals are transformed to ground-centered coordinates, known as the global (X,Y) coordinate system. As described in Chapter 2, Section 2.1.4, the global coordinate system with non-uniform spacing of the data is converted to a uniform grid in CRG format. Additional details of the system and data processing can be found in Appendix A, or other work by Kern and Wagner (Kern and Ferris 2007; Wagner, Kern et al. 2007).
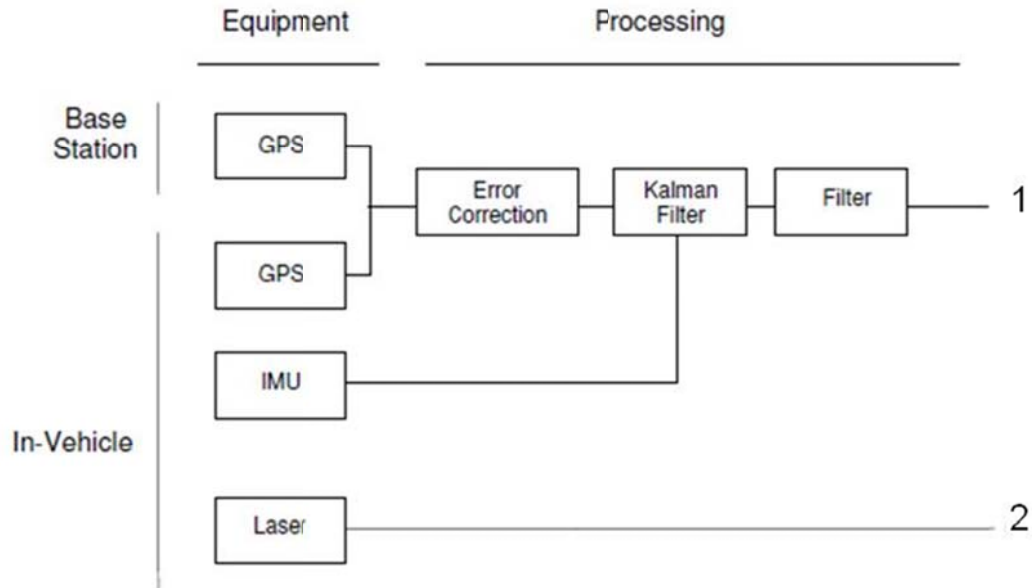
Figure 15. Signal processing overview: Data collection and filtering.

An example is used to show the importance of accurately measuring and removing the body motion of the vehicle from the terrain measurements: if the host vehicle body motion is not removed from the laser measurements, then the measurement from the scanning laser is not useful for vehicle simulations. Figure 16 illustrates this concept. Figure 16A is a photograph of a sidewalk curb that was measured with the VTMS. The data was processed with results shown in Figure 16B and Figure 16C. As can be seen in Figure 16B, the vehicle body motion has not been properly removed during processing, and if implemented in a vehicle simulation would produce unrealistic response results in addition to other errors. Figure 16C depicts the surface rendering of the same curb with the host vehicle body motion properly removed from the laser measurement using the information acquired from the INS.

A)
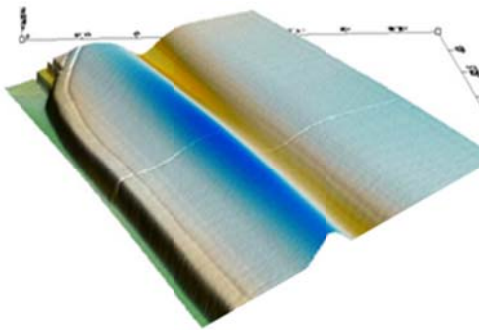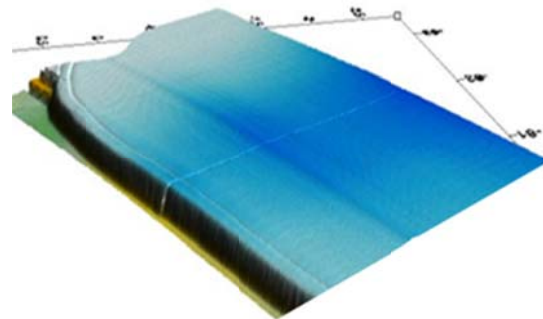
B)   C)

Figure 16. Body motion removal with VTMS.

Now that terrain surface has been nominally corrected for body motion, the terrain can be represented as a sequence of transverse scans, perpendicular to the direction of vehicle travel as described in Chapter 2, Section 2.1.4 of this dissertation. The formulation of the terrain surface as this series of vectors in the (u, v) coordinate system is important for the development of the INS drift removal method presented in the next Section of this chapter. It is important to note that this nominal correction for body motion relies on the ability to accurately measure the position and orientation of the vehicle, typically using an INS that is prone to small amounts of drift (Wagner, Kern et al. 2007). This drift manifests itself in several directions, predominantly in the vertical direction, but also in directions such as roll. The duration of each scan from a typical scanning laser is approximately 1 millisecond; it is assumed, therefore, that changes in the drift are negligible during the duration of each individual scan such that the drift for a particular scan is considered to be constant.

3.3 Developing the INS Drift Compensation Technique

3.3.1 Defining the Coordinate System

The terrain coordinate system used to develop this method is described in Chapter 2, Section 2.1.4 of this dissertation.

3.3.2 Defining the Vector Space

Each realization of the terrain surface comprises a sequence of vectors, $z_{i,k}$. Each vector in this realization has two components: the true surface vector, s_i, and the error surface vector, $e_{i,k}$. Since there is exactly one true surface, s_i, for any set of realizations, it is not indexed by k. Each of these sets of vectors, $\{z_{i,k}\}, \{s_i\}, \{e_{i,k}\}$, span the same Hilbert space **H**, such that

$$z_{i,k} = s_i + e_{i,k} \ \text{where } e_{i,k} \in \textbf{H} \qquad \text{Equation 18}$$

where the Hilbert space, **H**, is defined as the span of the set of error vectors $\{e_{i,k}\}$. Then the difference in any two vectors in the set $\{e_{i,k}\}$, must also be elements of the space as shown in Equation 19.

$$(e_{i,k} - e_{i,h}) \in \textbf{H} \quad \forall i \ \in [1,2,\dots,m] \ and \ \forall k, h \in [1,2,\dots r] \qquad \text{Equation 19}$$

Rearranging Equation 18 and substituting into Equation 19 shows that a difference in any two realizations, $z_{i,k}$ and $z_{i,h}$, must also span **H**.

$$e_{i,k} - e_{i,h} = z_{i,k} - z_{i,h} \in \textbf{H} \qquad \text{Equation 20}$$

That is, the difference in the measured vectors and all linear combinations of the vector sets must also span **H**. A Singular Value Decomposition (SVD) of differences between realizations $\{z_{i,k} - z_{i,h}\}$, is used to determine a set of orthonormal basis vectors $\{b_l\}$ that span **H** (Sylvester ; Lenzen and Waller 1997).

3.3.3 Decomposing the Vector Space

In the general sense, terrain can be described in terms of global and local components. Global components have a large and sustained effect on the surface such as hills and grades, bank angles, and ruts. Local terrain components are considered to be small, localized disturbances such as pot holes, speed bumps, chatter bumps etc. It is

important to establish the difference between the global and local concepts pertaining to terrain and the global and local subspaces pertaining to this work. The space $\mathbf{H}$ is decomposed into two subspaces: the global subspace, $\mathbf{G}$, and its complement, $\mathbf{G^c}$ which captures the local features of the space. In this sense, the global subspace, $\mathbf{G}$, pertains to the primary drift components, specifically the vectors that define the principle directions of $\mathbf{H}$, and the local subspace, $\mathbf{G^c}$, pertains to the noise, or residual of the drift components. The global subspace, $\mathbf{G}$, is defined by the span of a truncated set of these basis vectors, $\{b_l\}$. The selection criteria for these basis vectors are developed in the Error Modeling Section of this Chapter.

Vectors comprising the true surface, the realizations of each measured surface and the error surface all have components in both the global and local subspaces. To illustrate this concept, these vectors are decomposed into global and local subspaces according to Equation 21, Equation 22, and Equation 23 respectively. The first summation term in each Equation pertains to the contributions of the true surface, realization or error surface to the global subspace. The terms in the parentheses in each Equation pertain to the contributions of the true surface, realization or error surface to the local subspace. As can be seen by these Equations, the local subspace is defined as the residuals of the global subspace.

$$\boldsymbol{s_i} = (\boldsymbol{s_i})_G + (\boldsymbol{s_i})_{G^c} = \sum_{l=1}^{q} \langle \boldsymbol{b_l}, \boldsymbol{s_i} \rangle \, \boldsymbol{b_l} + \left(\boldsymbol{s_i} - \sum_{l=1}^{q} \langle \boldsymbol{b_l}, \boldsymbol{s_i} \rangle \, \boldsymbol{b_l} \right) \qquad \text{Equation 21}$$

$$\boldsymbol{z_{i,k}} = \left(\boldsymbol{z_{i,k}}\right)_G + \left(\boldsymbol{z_{i,k}}\right)_{G^c} = \sum_{l=1}^{q} \langle \boldsymbol{b_l}, \boldsymbol{z_{i,k}} \rangle \, \boldsymbol{b_l} + \left(\boldsymbol{z_{i,k}} - \sum_{l=1}^{q} \langle \boldsymbol{b_l}, \boldsymbol{z_{i,k}} \rangle \, \boldsymbol{b_l} \right) \qquad \text{Equation 22}$$

$$\boldsymbol{e_{i,k}} = \left(\boldsymbol{e_{i,k}}\right)_G + \left(\boldsymbol{e_{i,k}}\right)_{G^c} = \sum_{l=1}^{q} \langle \boldsymbol{b_l}, \boldsymbol{e_{i,k}} \rangle \, \boldsymbol{b_l} + \left(\boldsymbol{e_{i,k}} - \sum_{l=1}^{q} \langle \boldsymbol{b_l}, \boldsymbol{e_{i,k}} \rangle \, \boldsymbol{b_l} \right) \qquad \text{Equation 23}$$

3.3.4   Error Modeling

This section focuses on the development of the INS error model. Recall that the goal of this work is to remove the drift component of error, and thus, the primary focus of these developments occurs in the global subspace, $\mathbf{G}$. Again recall that the global

subspace, $\mathbf{G}$, is defined by the span of the truncated set of basis vectors, $\{b_l\}$, arising from a SVD of the realizations $z_{i,k}$. The error surface in the global subspace, $(e_{i,k})_G$, is defined as the summation of the projection of the error vectors in the global subspace, as defined in Equation 24.

$$\left(e_{i,k}\right)_G = \sum_{l=1}^{q} \langle b_l, e_{i,k}\rangle\, b_l \qquad\qquad \text{Equation 24}$$

The components of drift for each realization can occur in each of the q principle directions defined by the set of basis vectors $\{b_l\}$, but not in the local subspace, $\mathbf{G^c}$. The component of the drift, $\delta_{i,k,l}$, in each principle direction is the projection of the error vector in that direction as defined in Equation 25.

$$\delta_{i,k,l} \triangleq \langle b_l, e_{i,k}\rangle \qquad\qquad \text{Equation 25}$$

The components of the error vectors that do not lie in the global subspace, $\mathbf{G}$ are defined as residual noise vectors, $n_{i,k}$, and span the local subspace, $\mathbf{G^c}$. The decomposition of the error surface, $e_{i,k}$, is developed in Equation 26, where each error vector is defined as the summation of the projection of the error vectors, $e_{i,k}$, in the global directions, b_l, and the residual noise vectors, $n_{i,k}$.

$$e_{i,k} = \left(e_{i,k}\right)_G + \left(e_{i,k}\right)_{G^c} = \sum_{l=1}^{q} \delta_{i,k,l} b_l + n_{i,k} \quad \forall\, i \in [1,2,\dots,m]\ and\ \forall k$$
$$\in [1,2,\dots r] \qquad\qquad \text{Equation 26}$$

The number of basis vectors defining the global subspace, q, is incremented until the residual noise vectors can be considered realizations of a zero-mean process. The order in which basis vectors are added to the truncated set is determined by the magnitude of their corresponding singular value, in descending order. Then, by construction, the expectation of the noise vectors, $n_{i,k}$, is zero, as shown in Equation 27.

$$E\left[n_{i,k}\right] = 0\ \ \forall\, i \in [1,2,\dots,m]\ and\ \forall k \in [1,2,\dots r] \qquad\qquad \text{Equation 27}$$

Again, by construction, there is no noise component in the global subspace as shown in Equation 28.

$$\langle b_l, n_{i,k} \rangle = 0 \quad \forall \, l \in [1, 2, \ldots, q], \forall i \, [1,2, \ldots, m] \; and \; \forall k \in [1,2, \ldots r] \qquad \text{Equation 28}$$

The magnitude of drift, $\delta_{i,k,l}$, is defined for each direction in the global subspace, **G**. The drift is modeled herein as a random walk process (Favey, Cerniar et al. 1999; Giremus, Doucet et al. 2004; Kim and Sukkarieh 2004; Sun, Fu et al. 2008) defined in Equation 29.

$$\delta_{i,k,l} = \delta_{i-1,k,l} + \xi_{i,k,l}, where \; \xi_{i,k,l} \, \epsilon \, a \; zero \; mean \; process \qquad \text{Equation 29}$$

Drift is further decomposed into two components, $\alpha_{i,k,l}$ and γ_l, as defined in Equation 30, where γ_l is the average drift in each principle direction across all realizations and all locations and $\alpha_{i,k,l}$ is the difference from γ_l in the ith longitudinal location, for the kth realization.

$$\delta_{i,k,l} \triangleq \alpha_{i,k,l} + \gamma_l \qquad \text{Equation 30}$$

Drift is modeled as a random walk process and presently it is shown that $\alpha_{i,k,l}$ must be a zero mean process. Consider differences in the expected drift between sequential longitudinal locations, developed from Equation 29 and Equation 30, as shown in Equation 31.

$$E[\delta_{i,k,l}] - E[\delta_{i-1,k,l}] = E[\alpha_{i,k,l}] - E[\alpha_{i-1,k,l}] = E[\xi_{i,k,l}] = 0 \qquad \text{Equation 31}$$

Since there can be no difference in the expected drift between sequential longitudinal locations, the expected drift, $E[\delta_{i,k,l}]$, and the expected difference from the average drift, $E[\alpha_{i,k,l}]$, must be constant for all longitudinal locations, i. Now consider drift at a particular longitudinal location

$$\delta_{i,k,l} = \sum_{p=1}^{i} \xi_{p,k,l} + \delta_{0,k,l} = \sum_{p=1}^{i} \xi_{p,k,l} + \alpha_{0,k,l} + \gamma_l \qquad \text{Equation 32}$$

The expectation across all realizations is then

$$E[\delta_{i,k,l}] = \sum_{p=1}^{i} E[\xi_{p,k,l}] + E[\alpha_{0,k,l}] + E[\gamma_l] = E[\alpha_{0,k,l}] + \gamma_l \qquad \text{Equation 33}$$

Define γ_l to be the expectation of $\delta_{i,k,l}$ (which must be constant for all i) so that

$$\gamma_l \triangleq \mathrm{E}\left[\delta_{i,k,1}\right] \text{ and therefore } \mathrm{E}\left[\alpha_{0,k,l}\right] = 0 \qquad \text{Equation 34}$$

Since the expected difference from the average drift, $\mathrm{E}\left[\alpha_{i,k,l}\right]$, must be constant for all longitudinal locations and $\mathrm{E}\left[\alpha_{0,k,l}\right] = 0$, then by induction

$$\mathrm{E}\left[\alpha_{i,k,l}\right] = 0 \; \forall \; i, l \qquad \text{Equation 35}$$

Substituting the decomposed model of drift as defined in Equation 30, into the definition of error, as defined in Equation 26, yields the total error model as shown in Equation 36.

$$\boldsymbol{e}_{i,k} = \sum_{l=1}^{q} \delta_{i,k,l} \boldsymbol{b}_l + \boldsymbol{n}_{i,k} = \sum_{l=1}^{q} \left(\alpha_{i,k,l} + \gamma_l\right) \boldsymbol{b}_l + \boldsymbol{n}_{i,k} \qquad \text{Equation 36}$$

The expected error is then the summation of γ_l, the average drift in each principle direction across all realizations and all locations multiplied by each basis vector, as shown in Equation 37.

$$\mathrm{E}\left[\boldsymbol{e}_{i,k}\right] = \sum_{l=1}^{q} \gamma_l \boldsymbol{b}_l \qquad \text{Equation 37}$$

3.3.5   True Surface Estimation

The true surface must be estimated from a set of realizations, or measured surfaces, that contain some components of error in the form of drift and noise. The true surface is decomposed into the global and local subspaces according to Equation 38 and Equation 39.

$$\boldsymbol{s}_i = (\boldsymbol{s}_i)_{\mathrm{G}} + (\boldsymbol{s}_i)_{\mathrm{G}^c} = (\boldsymbol{z}_{i,k} - \boldsymbol{e}_{i,k})_{\mathrm{G}} + (\boldsymbol{z}_{i,k} - \boldsymbol{e}_{i,k})_{\mathrm{G}^c} \qquad \text{Equation 38}$$

$$\boldsymbol{s}_i = \left(\sum_{l=1}^{q} \beta_{i,k,l} \boldsymbol{b}_l - \sum_{l=1}^{q} \delta_{i,k,l} \boldsymbol{b}_l\right) + \left(\boldsymbol{z}_{i,k} - \sum_{l=1}^{q} \beta_{i,k,l} \boldsymbol{b}_l - \boldsymbol{n}_{i,k}\right) \qquad \text{Equation 39}$$

The projection of each realization, $\boldsymbol{z}_{i,k}$, in each global direction is defined as $\beta_{i,k,l}$.

$$\beta_{i,k,l} \triangleq \langle \boldsymbol{b}_l, \boldsymbol{z}_{i,k} \rangle \qquad \text{Equation 40}$$

It should be clear that the expectation of the true surface is simply the true surface itself, and taking the expectation operator across Equation 39 yields the expectation of the contributions from the global and local subspaces.

$$s_i = E[s_i] = \left(\sum_{l=1}^{q} E[\beta_{i,k,l}] \boldsymbol{b}_l - \sum_{l=1}^{q} E[\delta_{i,k,l}] \boldsymbol{b}_l \right)$$

$$+ \left(E[\boldsymbol{z}_{i,k}] - \sum_{l=1}^{q} E[\beta_{i,k,l}] \boldsymbol{b}_l - E[\boldsymbol{n}_{i,k}] \right)$$

Equation 41

$$= E\left[\sum_{l=1}^{q} \beta_{i,k,l} \boldsymbol{b}_l \right] - \sum_{l=1}^{q} \gamma_l \boldsymbol{b}_l + E\left[\boldsymbol{z}_{i,k} - \sum_{l=1}^{q} \beta_{i,k,l} \boldsymbol{b}_l \right]$$

Two issues arise with this definition of the true surface. First, γ_l is unknowable from the measured data. It is the offset, or bias, in each of the principle directions for all of the measurements. Second, in practice the true expectation cannot be known exactly, but must be estimated from a finite number of available measurements, as such the estimated expectation is introduced and the estimate of the true surface is defined by Equation 42. Detweiler (Detweiler and Ferris 2009) showed that nonlinear estimators of expectation (e.g., median, mode, trimmed mean, inverse distance to a power, and kriging) are most appropriate due in part to their insensitivity to outliers in the data. The 50% trimmed mean is implemented in the remainder of this work as the estimator of location for the estimated expectation

$$\hat{s}_i \triangleq \widehat{E}\left[\sum_{l=1}^{q} \beta_{i,k,l} \boldsymbol{b}_l \right] + \widehat{E}\left[\boldsymbol{z}_{i,k} - \sum_{l=1}^{q} \beta_{i,k,l} \boldsymbol{b}_l \right]$$

Equation 42

This estimation of the true surface is cast in terms of the true surface and the residual error terms that arise from the measurement offset, γ_l, and the error in estimating the expectation operator in Equation 43. Measurement techniques that mitigate the effects of γ_l and estimation techniques for the expectation operator will be addressed in the Discussion section of this Chapter.

$$\hat{s}_i \triangleq s_i + \sum_{l=1}^{q} \gamma_l \boldsymbol{b}_l + \hat{\mathrm{E}}\left[\sum_{l=1}^{q} \beta_{i,k,l}\boldsymbol{b}_l\right] - \mathrm{E}\left[\sum_{l=1}^{q} \beta_{i,k,l}\boldsymbol{b}_l\right]$$

$$+ \hat{\mathrm{E}}\left[\boldsymbol{z}_{i,k} - \sum_{l=1}^{q} \beta_{i,k,l}\boldsymbol{b}_l\right] - \mathrm{E}\left[\boldsymbol{z}_{i,k} - \sum_{l=1}^{q} \beta_{i,k,l}\boldsymbol{b}_l\right] \qquad \text{Equation 43}$$

$$= s_i + \sum_{l=1}^{q} \gamma_l \boldsymbol{b}_l + \hat{\mathrm{E}}\left[\sum_{l=1}^{q} \beta_{i,k,l}\boldsymbol{b}_l\right] + \hat{\mathrm{E}}\left[\boldsymbol{z}_{i,k} - \sum_{l=1}^{q} \beta_{i,k,l}\boldsymbol{b}_l\right]$$

$$- \mathrm{E}\left[\boldsymbol{z}_{i,k}\right]$$

Once the true surface has been estimated, it is possible, and instructive, to calculate the estimated error,

$$\hat{\boldsymbol{e}}_{i,k} = \boldsymbol{z}_{i,k} - \hat{\boldsymbol{s}}_i \qquad \text{Equation 44}$$

estimated drift,

$$\left(\hat{\boldsymbol{e}}_{i,k}\right)_G = \hat{\delta}_{i,k,l} = \sum_{l=1}^{q} \langle \boldsymbol{b}_l, \hat{\boldsymbol{e}}_{i,k}\rangle \boldsymbol{b}_l \qquad \text{Equation 45}$$

estimated noise,

$$\left(\hat{\boldsymbol{e}}_{i,k}\right)_{G^c} = \hat{\boldsymbol{n}}_{i,k} = \hat{\boldsymbol{e}}_{i,k} - \sum_{l=1}^{q} \langle \boldsymbol{b}_l, \hat{\boldsymbol{e}}_{i,k}\rangle \boldsymbol{b}_l \qquad \text{Equation 46}$$

and the estimated drift-free measurement of the surface.

$$\boldsymbol{z}_{i,k} - \left(\hat{\boldsymbol{e}}_{i,k}\right)_G = \boldsymbol{z}_{i,k} - \sum_{l=1}^{q} \langle \boldsymbol{b}_l, \hat{\boldsymbol{e}}_{i,k}\rangle \boldsymbol{b}_l \qquad \text{Equation 47}$$

3.3.6   Model Order Study

This section develops an error analysis on the residuals to prove the retention of high-fidelity information of the data set. In this case, the residuals are the noise process associated with the local subspace. Furthermore, the phrase "retaining high-fidelity information" is defined such that the noise floor of the residuals is at or below the system resolution and accuracy, specifically that the mean and standard deviation of the noise are

below 1mm. Thus, the standard deviation and mean of the noise surface are computed after the implementation of each basis vector and compared to these desired values to determine if the criteria have been satisfied.

For each horizontal grid location (i, j) consider r realizations (measurements) of the terrain surface and the corresponding estimate of noise $n_{i,j,k}$. These realizations represent a distribution of the estimated noise at that grid location from which the standard deviation and mean noise are calculated. Together, these statistics form surfaces describing the noise for each added basis vector. These surfaces indicate how the residuals change over the terrain surface as more basis vectors are added. The mean is calculated to demonstrate that the assumption of a zero-mean process is valid for this example, as shown in Equation 48. An estimation of the mean of the entire noise surface is calculated with Equation 49. The standard deviation is computed with the Equation 50, where $\overline{n_{i,j}}$ is the sample mean at each grid location. In order to estimate the standard deviation of the entire surface, the square root of the average of variances of each location is calculated, as shown in Equation 51, where m is the length of the terrain sample, and n is the width of the terrain sample.   These two measures representing the entire noise surface are used to perform a model order study and determine exactly how many basis vectors are required to satisfy the assumptions presented in the theory. An example using the INS drift compensation method is presented next.

$$\overline{n_{i,j}} = \frac{\sum_{k=1}^{r} n_{i,j,k}}{r}$$ Equation 48

$$\overline{n} = \frac{\sum_{i=1}^{m} \sum_{j=1}^{n} \overline{n_{i,j}}}{(m * n)}$$ Equation 49

$$\sigma_{i,j} = \sqrt{\frac{\sum_{k=1}^{r} \left(n_{i,j,k} - \overline{n_{i,j}}\right)^2}{r - 1}}$$ Equation 50

$$\sigma = \sqrt{\frac{\sum_{i=1}^{m} \sum_{j=1}^{n} \sigma_{i,j}^2}{(m * n)}}$$ Equation 51

3.4 Chapter 3 Proof of Concept

3.4.1   Design of Experiment

An example is presented to demonstrate the effectiveness of the INS drift compensation method developed in this work. By identifying and eliminating the INS drift component of the surface, a large contribution of the system error can be removed and thus, improve the correlation between multiple measurements of a single terrain surface. A single lane of longitudinally tined concrete road, approximately 151.1m long, was selected at the MnRoad testing facility in Albertville, Minnesota for this experiment. The pavement was measured ten times, providing ten realizations of the surface for analysis (r=10). The lane was closed to traffic and the data were collected in alternating directions. The data were acquired at a vehicle speed of approximately 10 m/s, so that a transverse scan was collected every 10mm longitudinally.   The measurements were nominally corrected for body motion (Wagner, Kern et al. 2007) and mapped to a regularly spaced grid as described in Section 2.1.4. The longitudinal spacing of 10mm resulted in 15110 longitudinal locations and a transverse spacing of 25.4mm at 48 transverse locations resulted in a 1.22 m wide terrain surface.

3.4.2   Application of Theory

SVD was performed on a set of ten realizations of longitudinally tined concrete terrain surfaces yielding a set of corresponding singular values and basis vectors. The magnitude of the ordered set of singular values corresponding to each basis vector is represented graphically in Figure 17. As indicated by the Figure, the first singular value has a magnitude of 6.63, the second singular value has a magnitude of 1.76, and subsequent singular values are negligible. Additionally, the shapes of the first two basis vectors are plotted in Figure 18.   In general terms, it appears that the first basis vector accounts for a constant offset in elevation and the second basis vector describes the bank angle (i.e., 'cross-slope' or 'road camber'). This agrees with intuition that the primary changes in the road surface are in elevation and banking. The results of the magnitude of the singular values suggest that the first two singular values will have the greatest effect on the terrain surface and may be classified as the global subspace, **G**. However, in practice, the number of basis vectors, q, sufficient to define the global subspace, **G**, must

be incremented until the noise vectors, $\boldsymbol{n}_{i,k}$, have a mean that is not statistically different than zero. The order in which these basis vectors are added is determined by the magnitude of their corresponding singular value. The required number of basis vectors to result in a zero-mean noise process for this example is studied later in this Section.
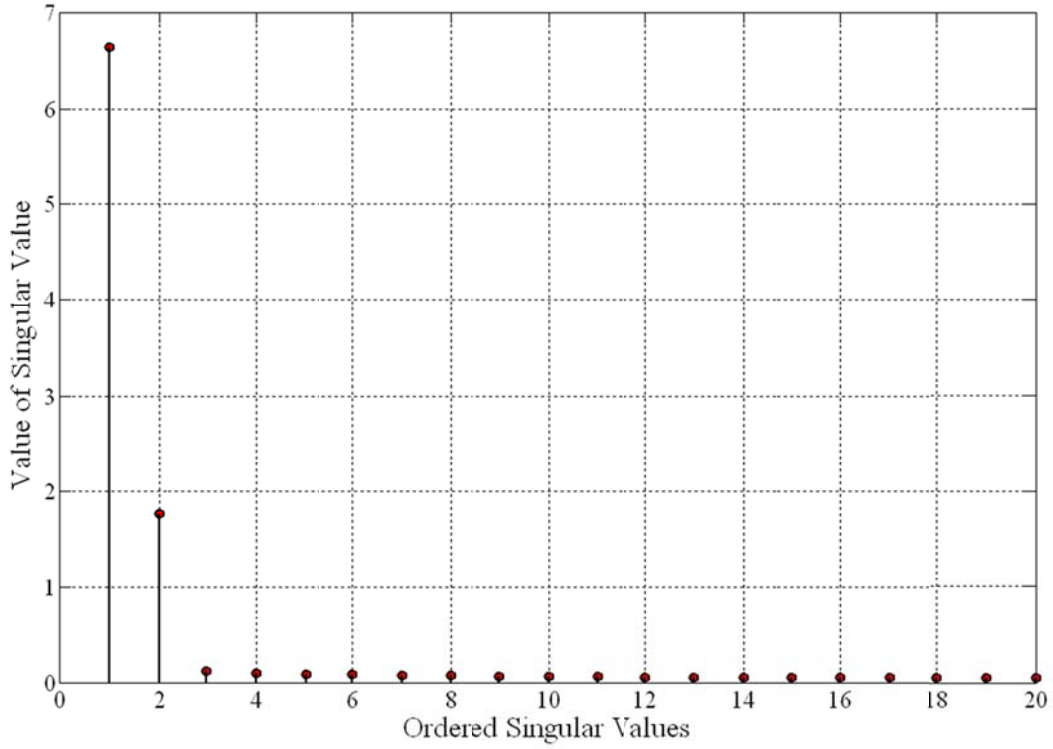
Figure 17. Plot of first 20 Singular Values.
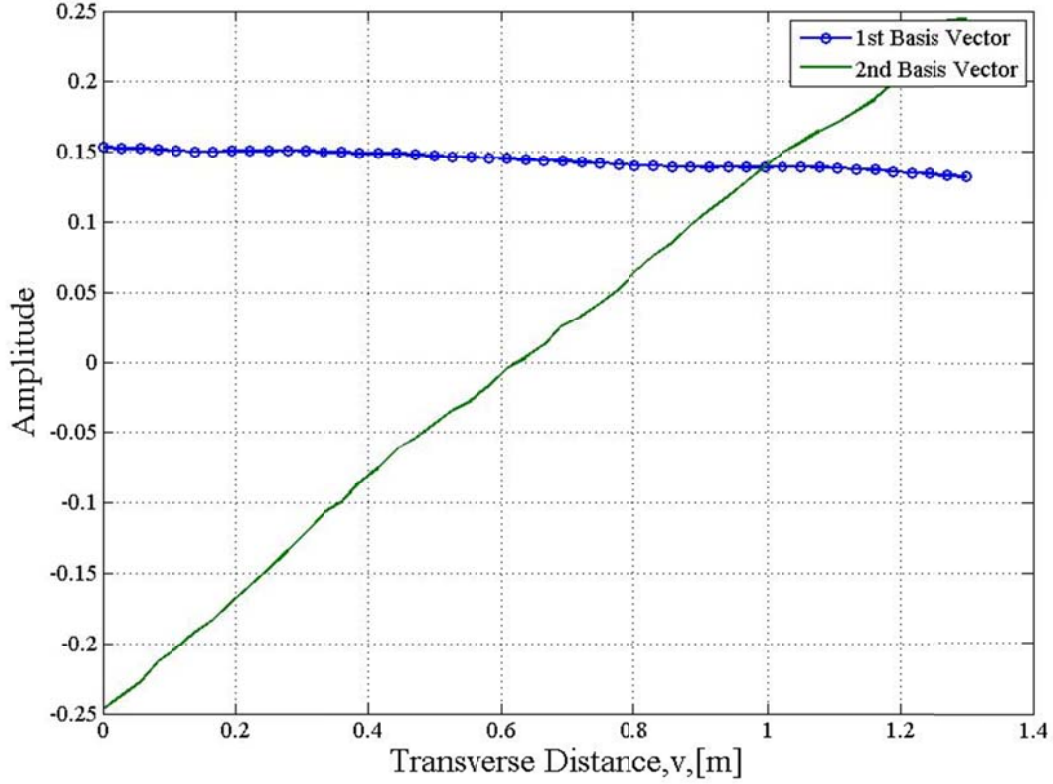
Figure 18. Plot of first two basis vectors.

The implications of adding basis vectors to the span of the global subspace, **G** are shown qualitatively by examining the estimated noise vectors (refer to Equation 46) and the estimated drift-free measurements of the surface (Equation 47). This qualitative examination is considered in two 'views': first, transverse noise vectors are considered at a particular longitudinal location; and second, longitudinal profiles of the measurements, with the drift components removed, at a particular transverse location. For the second view, consider each estimated drift-free measure of the surface to be the concatenation of longitudinal profiles; each longitudinal profile is then parallel to the coordinate u in Figure 8 and located at a particular transverse location, indexed by j.

First, consider the case when no basis vectors have been introduced and no drift has been removed. A typical set of noise vectors (viewed from the transverse direction at an i index location of 300), as shown in Figure 19, and longitudinal profiles of measured terrain (for a j index location of 25), shown in Figure 20. Since no basis vectors have been introduced, the noise vectors in this instance represent the total error for this set of realizations. It is clear that significant vertical offset and banking (seen in the estimated

noise vectors) exist in the data. The longitudinal profiles show +/-10 mm of vertical drift, which is consistent with the static test results shown in Figure 14. Additionally, it can be seen that one of the profiles, indicated by yellow triangles, does not follow the same trends as the other nine profiles. If these data were merely averaged, then this profile would have a significant effect on the results of the estimated true surface. Since each profile is individually corrected for drift error, and a nonlinear expectation estimator that is insensitive to outliers (i.e., the trimmed mean) is selected, then this outlier has minimal effect on the calculation of the estimated true surface.
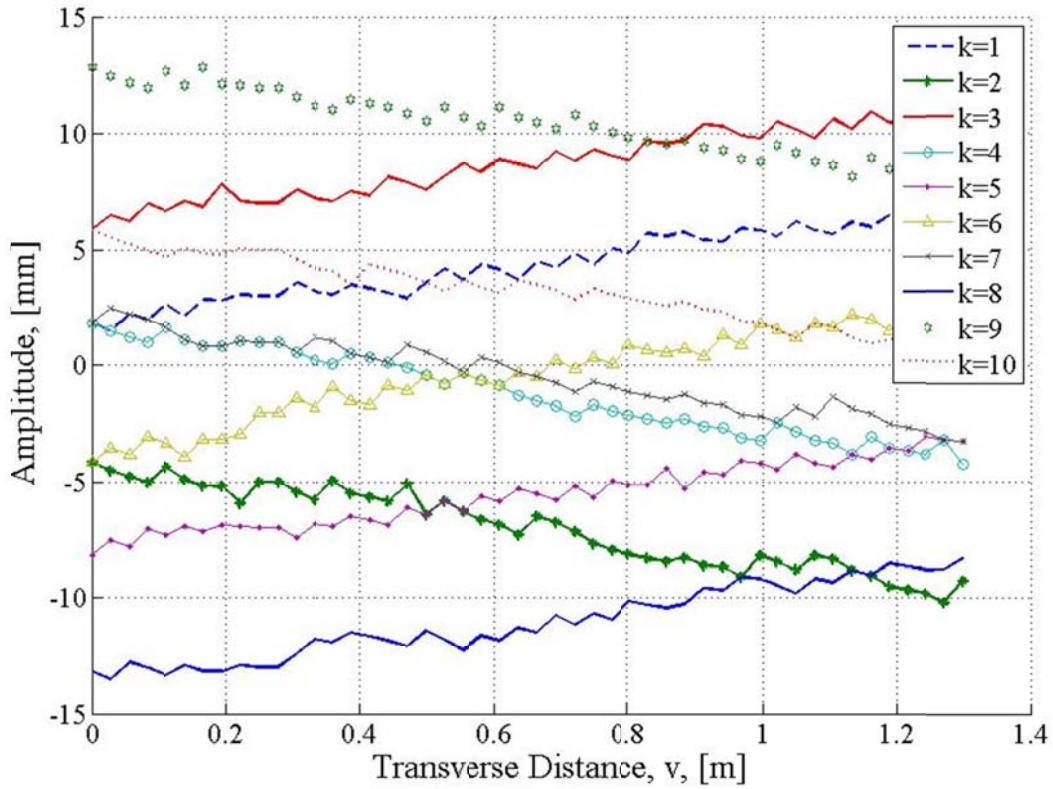
Figure 19. Estimated noise vectors (q=0 and i=300).
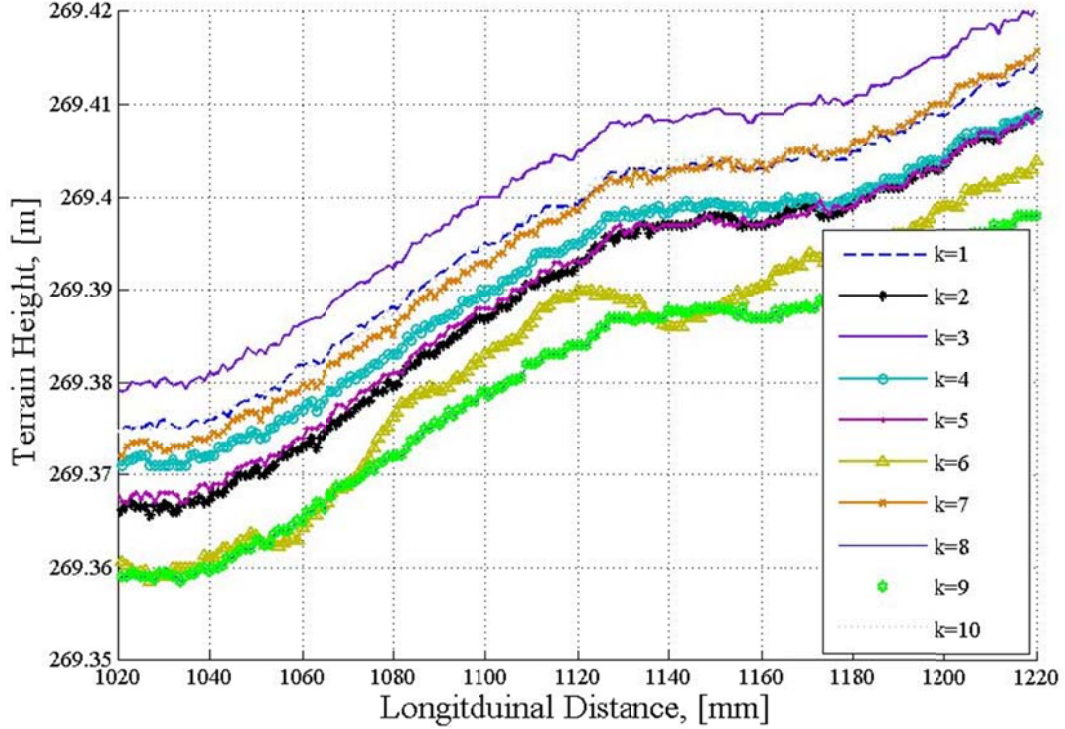
Figure 20. Longitudinal profiles (*q*=0 and j=25).

After implementing the first basis vector (i.e., *q*=1), the resulting noise vector (Figure 21), shows that a significant trend in the banking remains, however the vertical offset has been eliminated. It is also clear from both the noise vectors and the longitudinal terrain profiles that the measured data are segregated into two groups depending on the direction the vehicle was traveling when the data were collected. In the transverse direction, roll drift exists in both directions. Recall that the data were collected in opposite directions. This data acquisition strategy allows this roll drift to be detected in the measurements, as indicated by the two distinct groups: one group exhibiting a positive slope that was collected in one direction, and the second group exhibiting a negative slope that was collected in the opposite direction. Similarly, these two groups are characterized by distinct offsets in the longitudinal direction. This longitudinal profile is taken near the right side of the pavement, so that the grouping with a positive slope has a positive offset from the group with the negative slope. This indicates that the terrain surface itself does not contain significant banking, but rather that the system has some drift (or bias) in the body roll estimate. Although a distinct offset remains between

71

these groups, the spread of the data as shown in Figure 22 is decreased to +/-5 mm, which is a significant improvement over the original data shown in Figure 20.
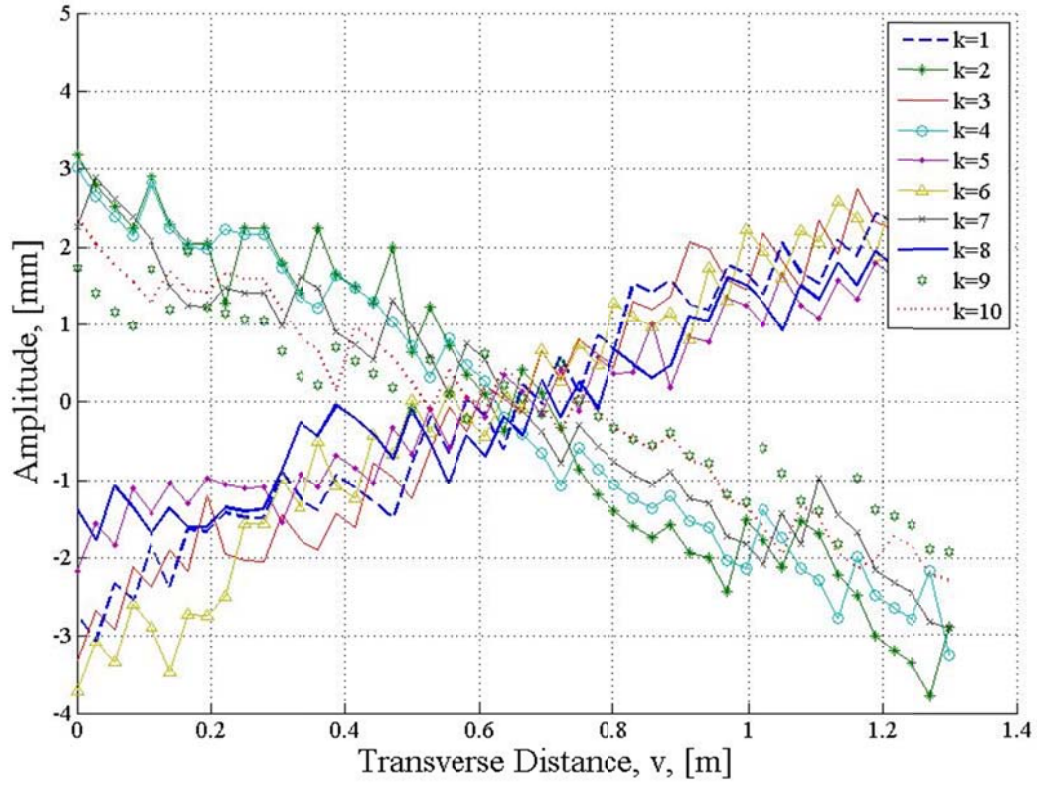
Figure 21. Estimated noise vectors (*q*=1 and i=300).
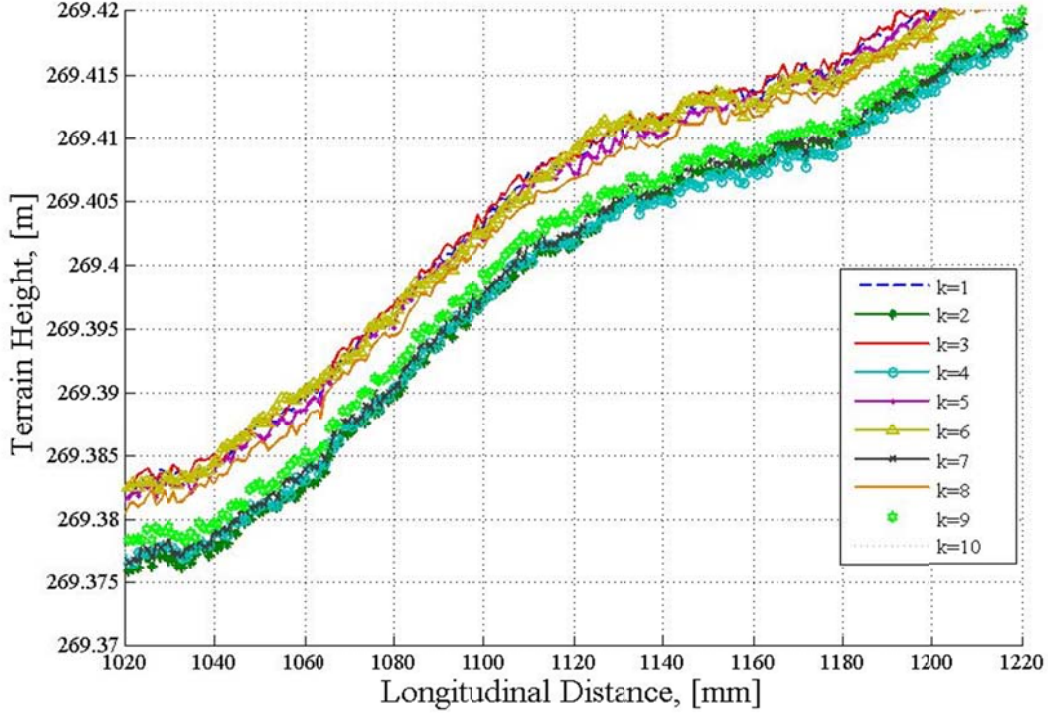
Figure 22. Longitudinal profiles (q=1 and j=25).

Implementing the first two basis vectors (i.e., q=2) of drift compensation results in the estimated noise vectors shown in Figure 23 and in the longitudinal profiles of drift-free terrain shown in Figure 24. From Figure 23, it can be seen that two basis vectors suffice to reduce the estimated noise to a zero-mean process. The constant offset in elevation and the cross-slope have been eliminated from the estimated noise; the remaining vertical variation is within +/-1 mm, which is consistent with the resolution of the scanning laser (1 mm). Consistently, the variation in the drift-free surfaces results in a vertical variation of +/-1mm. An enlarged segment of the profiles is shown in Figure 24, where the estimated true profile is shown as a dashed line. The estimated true profile depends on the method selected to estimate the expected value of the profile at that location, and as such, the trimmed mean was used as the estimation of the expectation to determine the true surface.
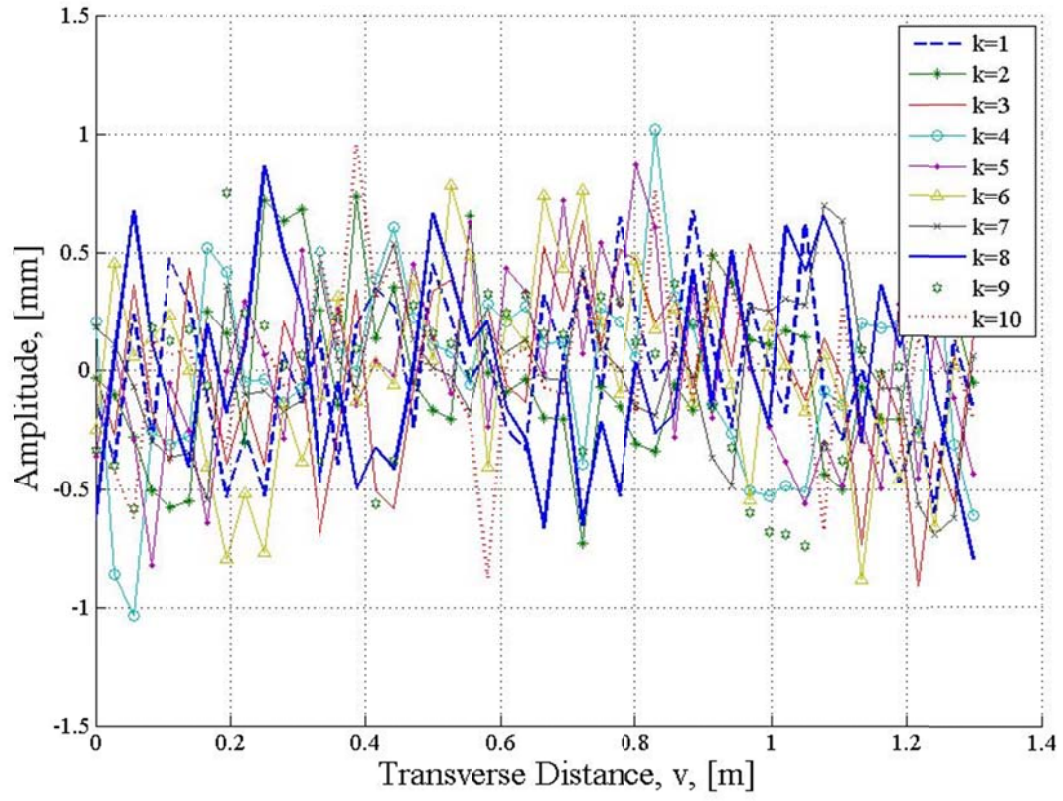
Figure 23. Estimated noise vectors (q=2 and i=300).
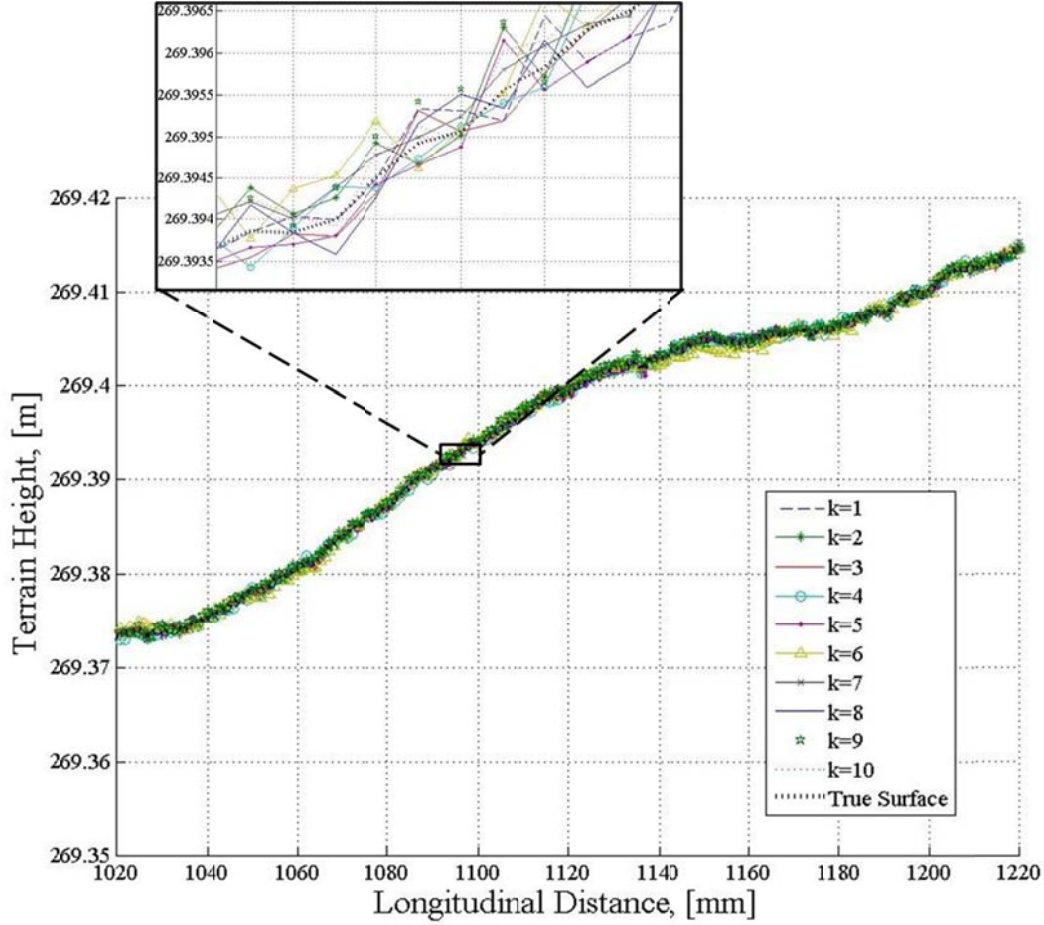
Figure 24. Longitudinal profiles (q=2 and j=25) and estimated True Surface.

3.4.3   Residual Analysis and Model Order Study

The previous Section qualitatively showed that the implementation of two basis vectors appears to be sufficient in classifying the global subspace. An error analysis on the residuals is performed in this section to demonstrate the retention of high-fidelity information of the data set. Recall that the phrase "retaining high-fidelity information" is defined such that the noise floor of the residuals is at or below the system resolution and accuracy, or that the mean and standard deviation of the noise are below 1mm. Thus, the standard deviation and mean of the noise surface are computed after the implementation of each basis vector and the results of this study are presented.

First consider the mean noise amplitude surface with no basis vectors, as shown in Figure 25A. Since no basis vectors have been introduced, the mean noise surface represents the mean error for this set of realizations. It can be seen that significant

75

vertical offset and cross-slope exist in the data. Referring to Figure 25B, transverse mean noise vectors have been plotted along a sample of longitudinal locations along the terrain surface to clearly show the variation of the surface. This figure also identifies the bounds specified by the measurement system of +/-1mm as indicated by red dashed lines. While the variation of each realization is on the order of +/-10mm, the variation of the average of the realizations is within the desired bounds of +/-1mm. This suggests that another measure, the standard deviation, needs to be investigated.
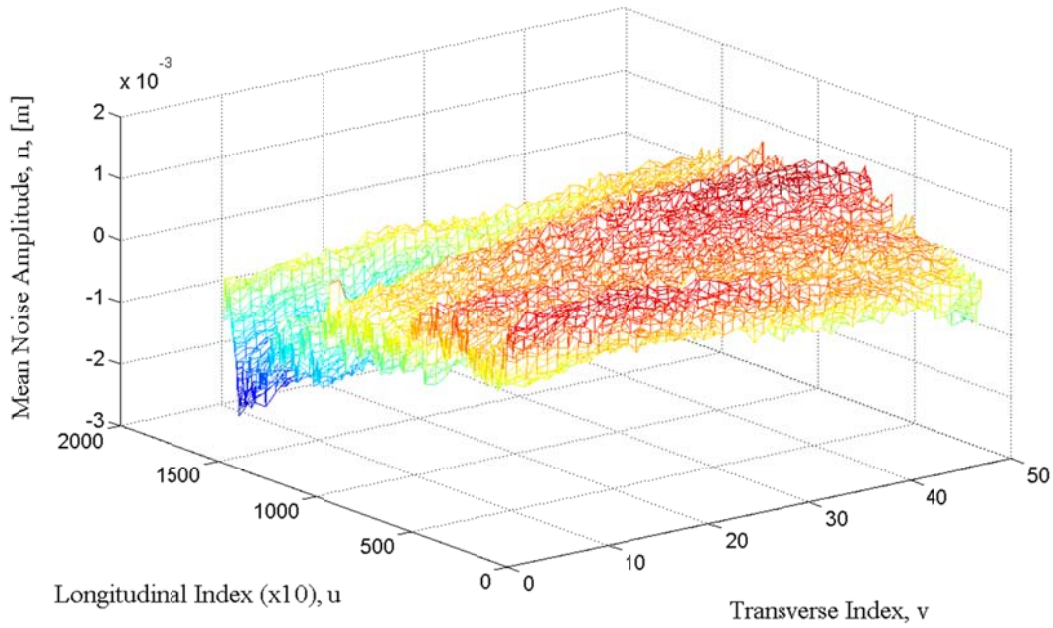
A)
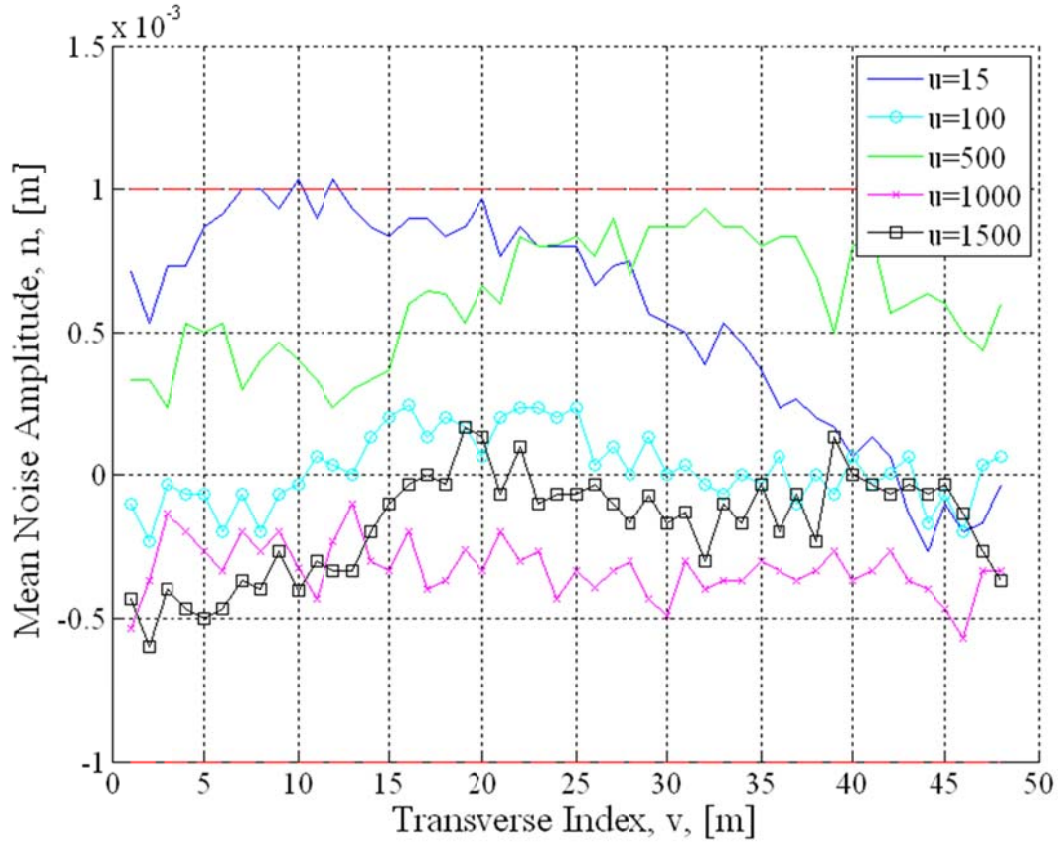
B)

Figure 25. Amplitude of mean noise surface A) 3D plot B) Cross-sections from 3D Plot for 0 basis vectors.

Next, the standard deviation surface of the terrain realizations is calculated and presented in Figure 26A. The corresponding transverse standard deviation vectors have been plotted in Figure 26B along a sample of longitudinal locations along the terrain surface to clearly show the variation across the surface. Again, since no basis vectors have been introduced, the standard deviation surface represents the standard deviation of the total error for this set of realizations. Figure 26B clearly identifies the desired std. deviation of 1mm with a red dashed line. It can be seen that the standard deviation of the noise surface for 0 basis vectors does not meet the 1mm or less requirement.
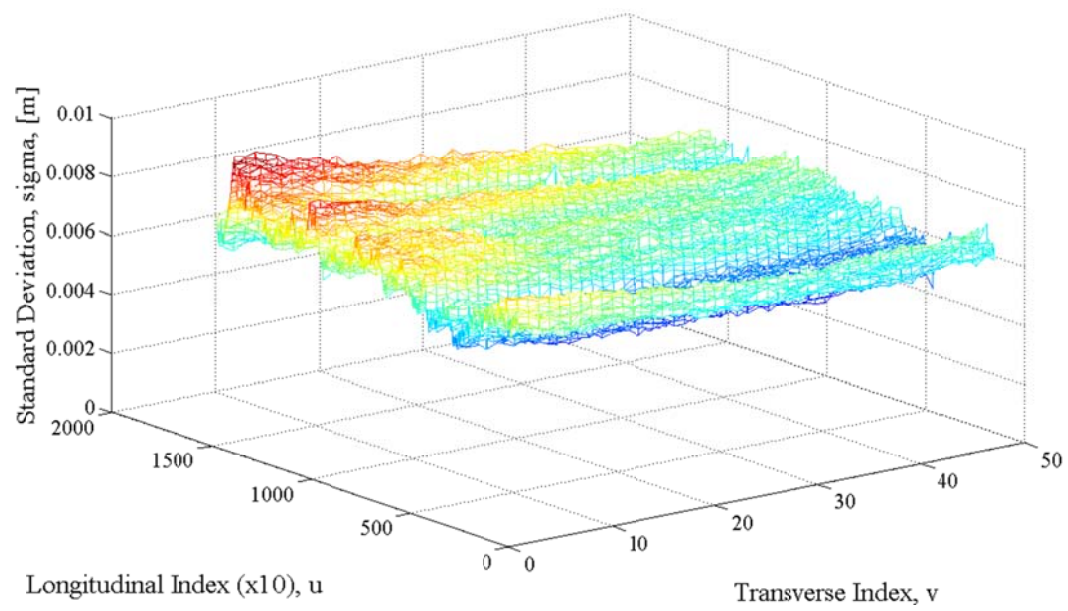
A)
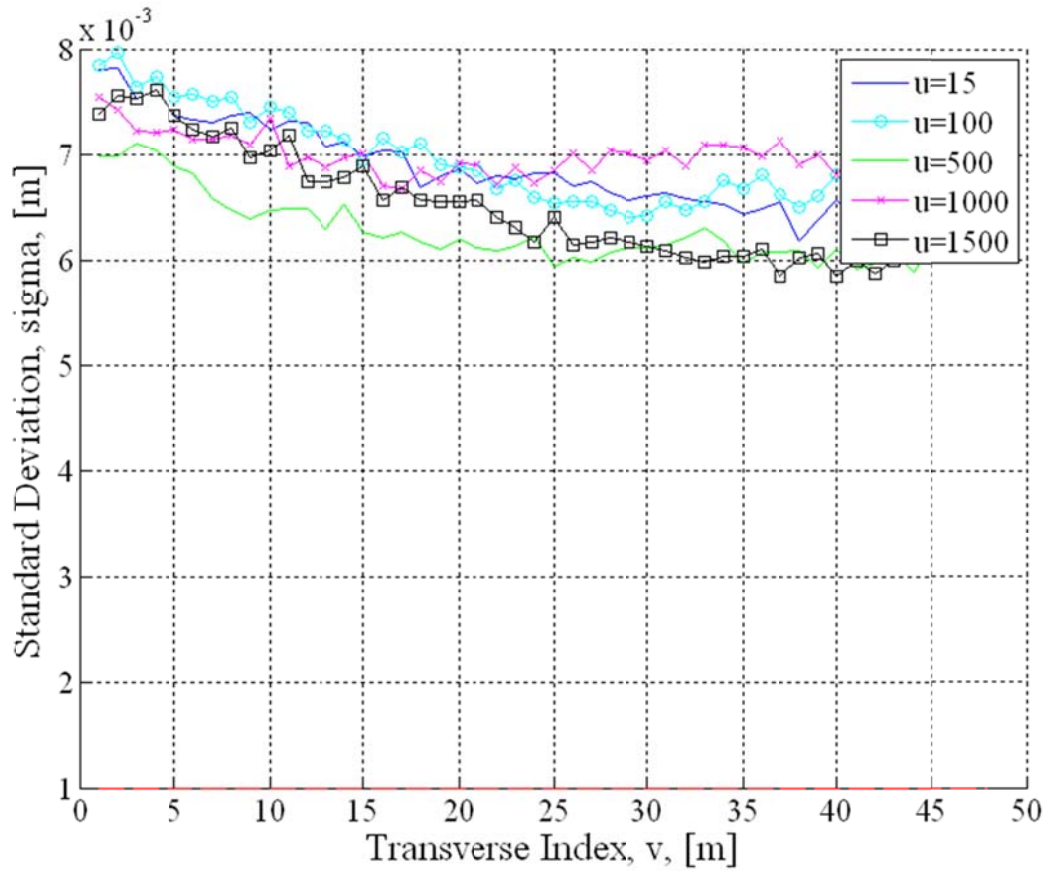
B)

Figure 26. Standard deviation of noise surface A) 3D plot B) Cross-sections from 3D plot for 0 basis vectors.

After implementing the first basis vector (i.e., $q=1$), the resulting mean noise amplitude surface (Figure 27A), shows that the major vertical fluctuations of the surface have been removed. Referring to the transverse mean noise vectors in Figure 27B, it can be seen that the tolerance on the mean has been reduced to +/-0.2mm from +/-1mm. Again, the standard deviation of the noise surface must be investigated to identify if the implementation of one basis vector is sufficient to meet the requirement of a standard deviation of 1mm or less.
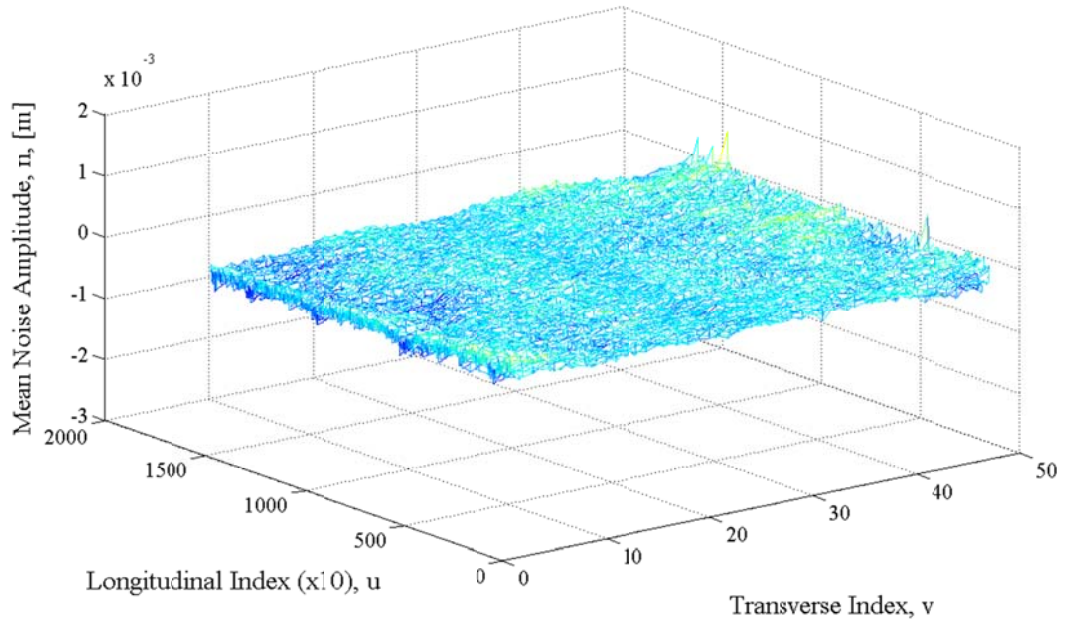
A)
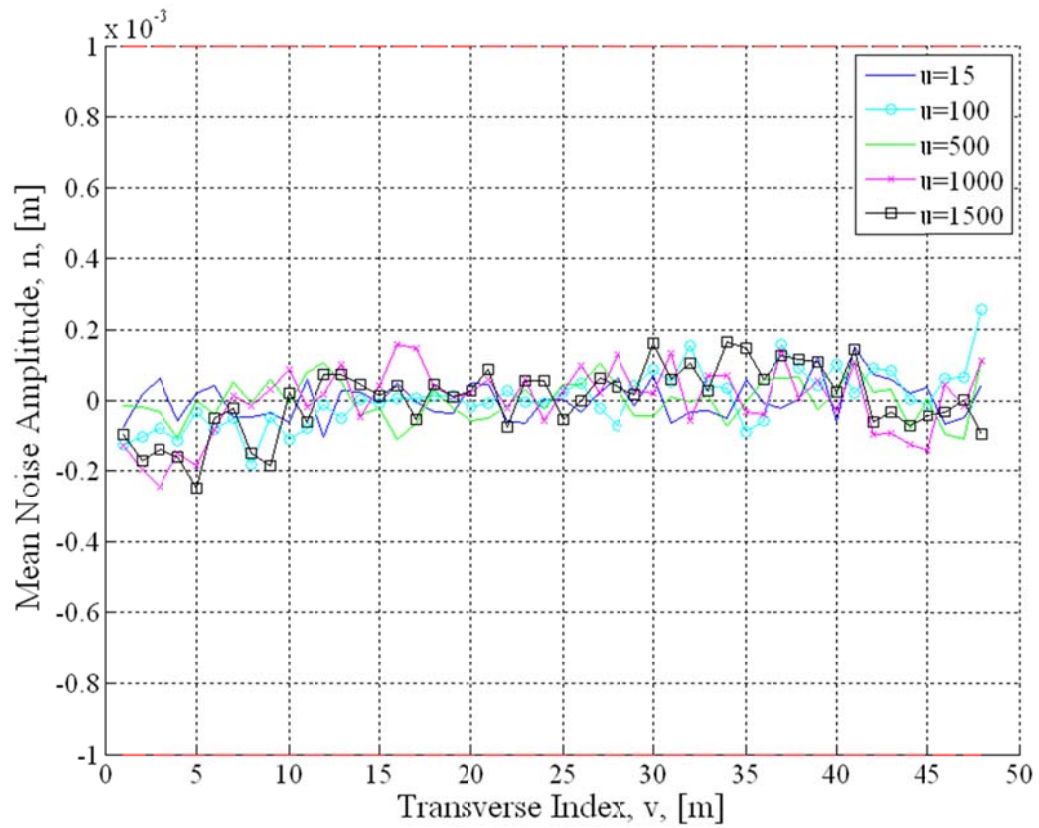
B)

Figure 27. Amplitude of mean noise surface A) 3D plot B) Cross-sections from 3D Plot for 1 basis vector.

The standard deviation surface of the terrain realizations is calculated and presented in Figure 28A after the implementation of one basis vector (q=1). The corresponding transverse standard deviation vectors have been plotted in Figure 28B along a sample of longitudinal locations along the terrain surface to clearly show the variation across the surface. The effect of the roll component of error is evident in the fact that the standard deviation takes on a 'v-shape.' This 'v-shape' corresponds with Figure 21, where the vertical error component has been removed and categorized as the first component of drift, but the roll component of error remains. From Figure 28B, it can also be seen that with the implementation of one basis vector, the standard deviation is not within the desired bounds of less than 1mm.

A)

B)
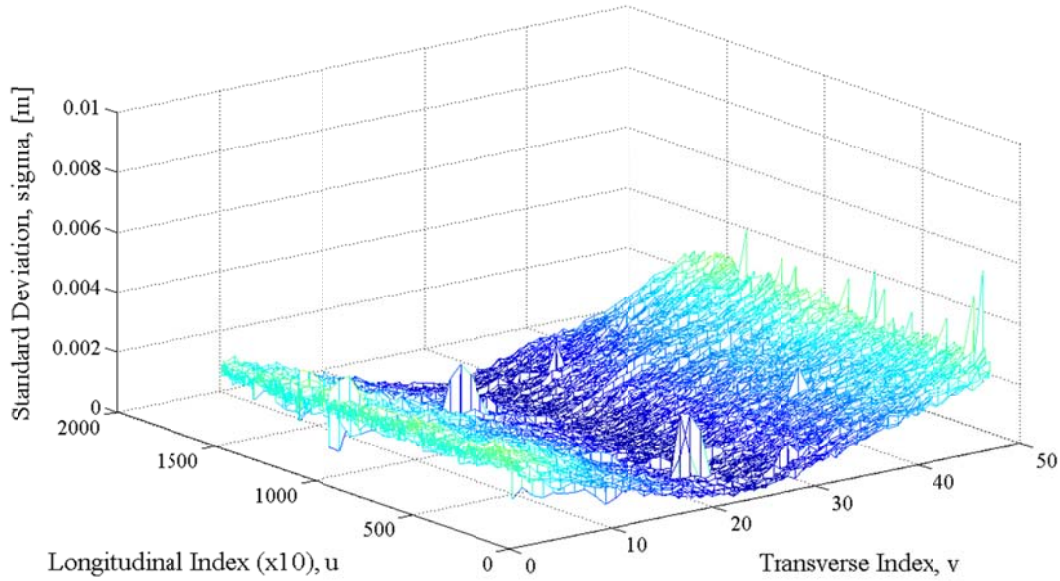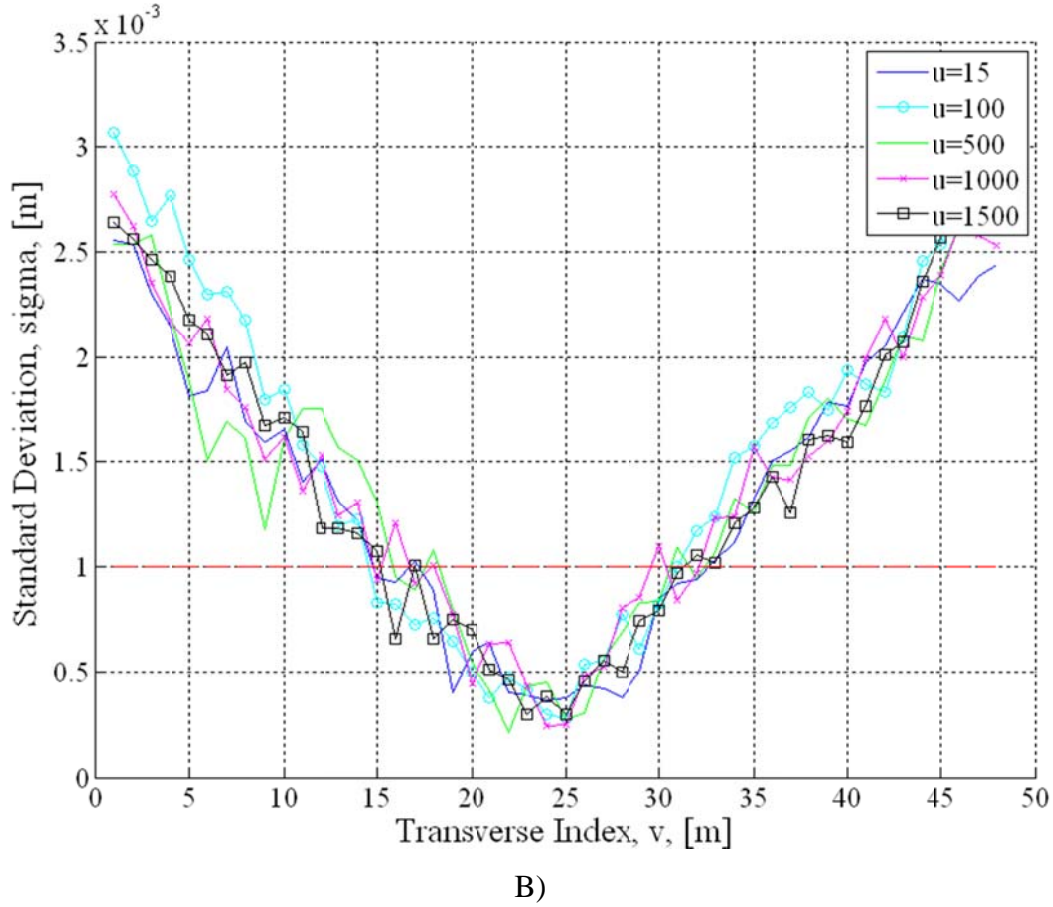
Figure 28. Standard deviation of noise surface A) 3D plot B) Cross-sections from 3D plot for 1 basis vector.

Implementing the second basis vector (i.e., $q=2$) of drift compensation results in the mean noise amplitude surface shown in Figure 29A and the transverse mean noise vectors in Figure 29B. The reduction of the amplitude of the mean noise surface is negligible from the implementation of two basis vectors. Again, the standard deviation of the noise surface must be investigated to identify if the implementation of two basis vectors is sufficient to meet the requirements of a standard deviation of 1mm or less.
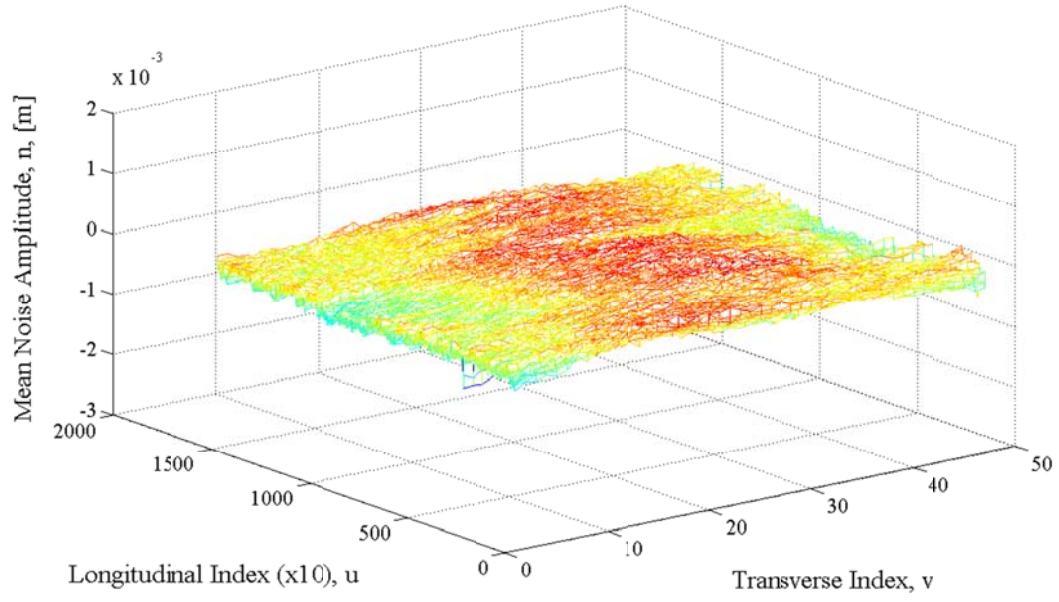
A)
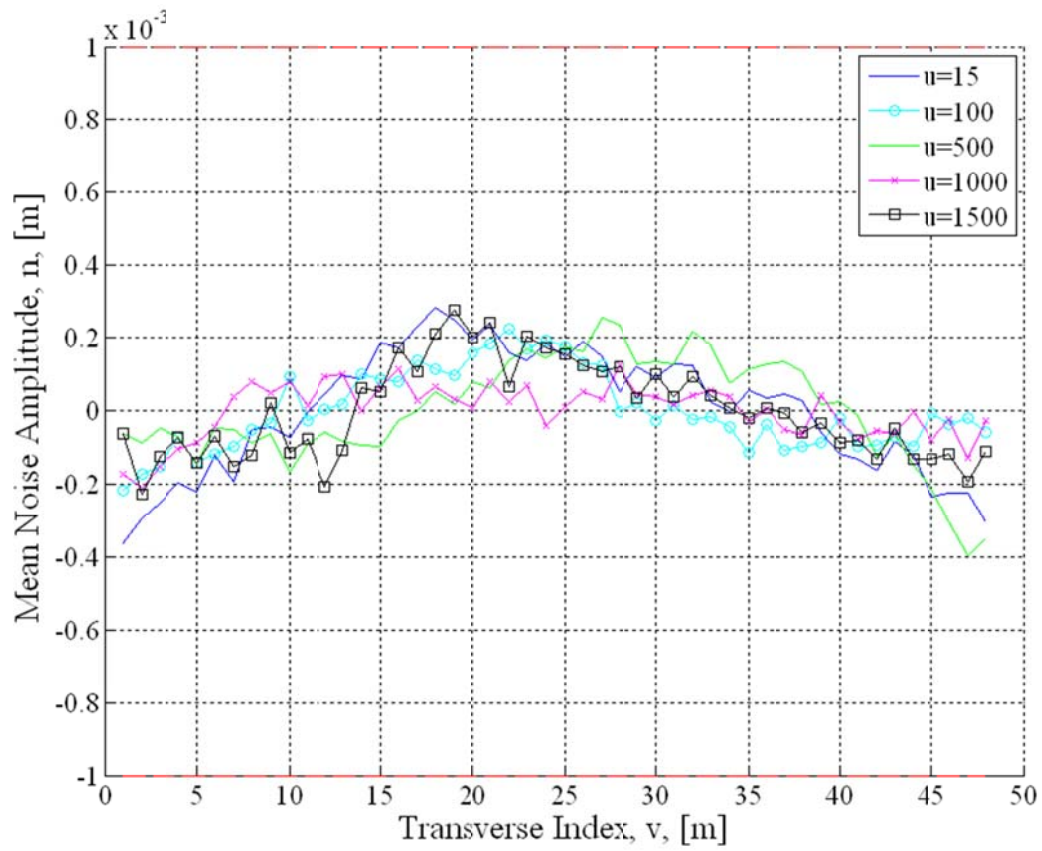
B)

Figure 29. Amplitude of mean noise surface A) 3D plot B) Cross-sections from 3D Plot
for 2 basis vectors.

Implementing the second basis vector (i.e., $q=2$) of drift compensation results in the standard deviation surface shown in Figure 30A and the transverse standard deviation vectors in Figure 30B. From Figure 30B, it can be seen that two basis vectors suffice to reduce the standard deviation of the noise surface to less than 1mm, except for a few isolated locations where the maximum standard deviation is on the order of 2mm. These isolated locations can be attributed to discrepancies between the realizations at those locations. In this case, the discrepancies can be attributed to cracks between the concrete slabs. This exemplifies one of the limitations of high-speed terrain measurements systems in the fact that these systems have difficulties in accurately recording the depth of cracks in terrain segments. One important feature of this method is that localized events (such as cracks) are identified through this noise analysis. Without the ability to remove the drift component of the error, the error due to drift would dominate the error analysis, masking the presence of localized anisotropicities. Once the method has been applied to the data, and the drift has been removed, these localized events produce deviations from the average standard deviation of the noise surface, in this case on the order of 2mm. Since this deviation is greater than the resolution of the laser system, it can be used to identify the significance of the localized event. It should be clear, however, that the method developed in this work is able to identify localized events and also produce a noise surface that is not so significant that it will hinder the vehicle response over this surface at this location. Implementation of two basis vectors produces an average standard deviation of the noise that is less than half of the desired value of 1mm. Again,, the maximum standard deviation can be used to identify localized events (local anisotropicities). The constant offset in elevation and the cross-slope are statistically insignificant (Refer to Figure 29B); the remaining vertical variation is within +/-0.2 mm, which is much less than the resolution of the scanning laser (1 mm). These results are consistent with the qualitative analysis completed in the previous section where the effect of each basis vector on each realization was studied.
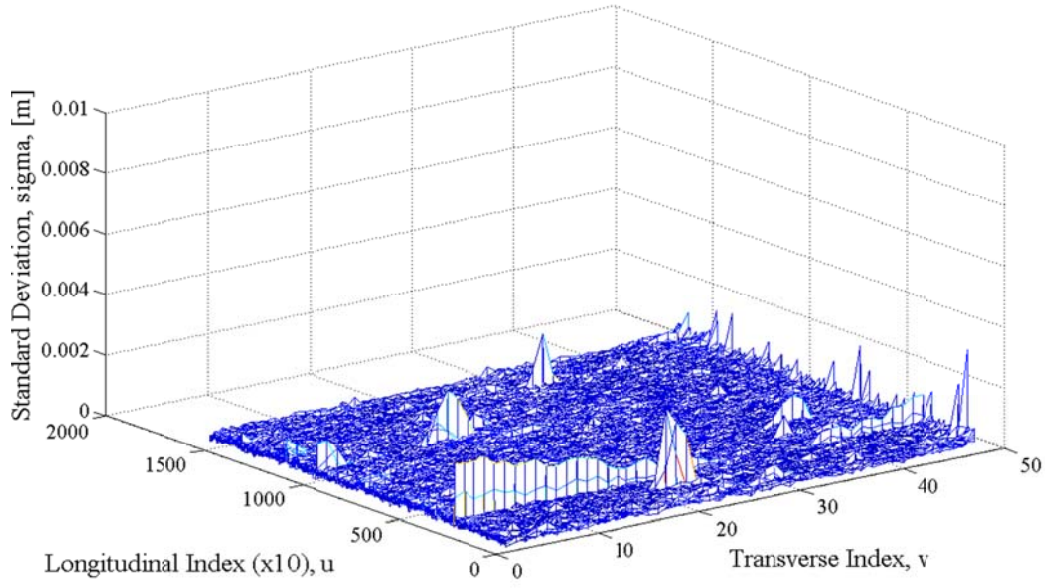
A)
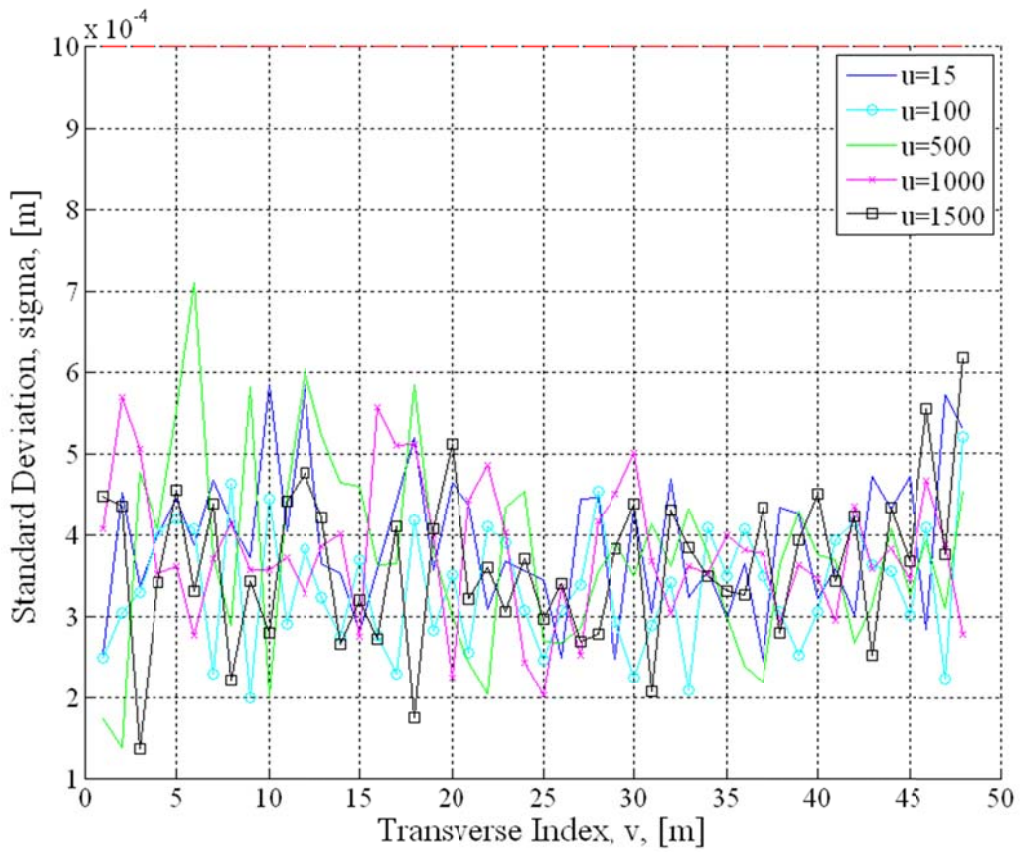
B)

Figure 30. Standard deviation of noise surface A) 3D plot B) Cross-sections from 3D plot for 2 basis vectors.

To summarize the results of the residual analysis on the noise surface, Figure 31 and Figure 32 were developed. While there are as many basis vectors as the dimension of the space (in this case 48), the general trend in the mean noise and standard deviation model order studies can be identified after the implementation of two basis vectors. Figure 31 is the result of a scalar representation of the mean noise amplitude surface after the implementation of each basis vector, as calculated by Equation 51. The desired value of 1mm corresponding to the resolution of the laser measurement system is identified as a red dashed line. It can be seen that the mean of the noise surface is on the order of 10^{-5} meters before any basis vectors have been implemented. Once one basis vector (and subsequent basis vectors) has been implemented, the mean of the noise surface drops down to the order of 10^{-8} meters. This correlates to the previous findings in this section that after the implementation of the first basis vector, the mean of the noise does not change. This scalar representation of the mean noise amplitude suggests that the residuals are a zero-mean process via the nonlinear expectation estimate, without any implementation of basis vectors ($q=0$). However, it is clear that the standard deviation of the noise needs to be considered to indicate a measurement of confidence that the mean noise surface is truly a zero-mean process.
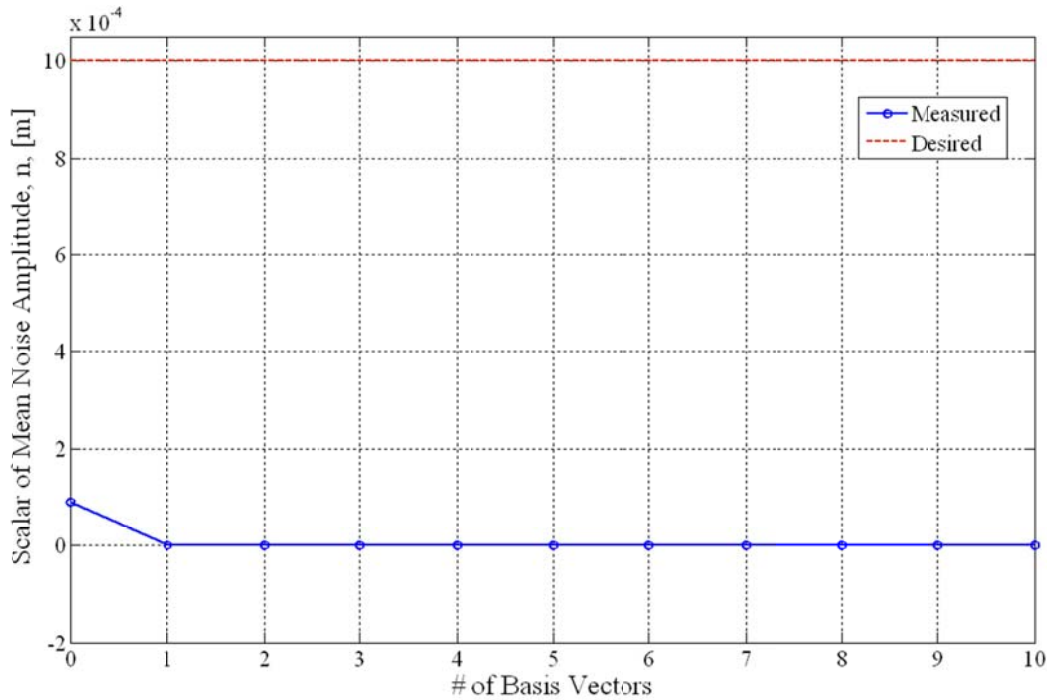
Figure 31. Model order study of summary measure of mean noise amplitude.

Figure 32 is the result of a scalar representation of the standard deviation of the noise surface after the implementation of each basis vector, as calculated by Equation 49. The desired value of 1mm corresponding to the resolution of the laser measurement system is identified as a red dashed line. It can be seen that the standard deviation of the mean noise surface (calculated from Equation 51) is approximately equal to 7mm before any basis vectors have been implemented. Implementing one basis vector brings the standard deviation down to approximately 1.5mm, which is still greater than the desired value of 1mm. Implementation of two basis vectors (and subsequent basis vectors) brings the standard deviation down to approximately 0.35mm, so that approximately three standard deviations lie within the 1mm bound, and is well within the resolution of the measurement equipment.
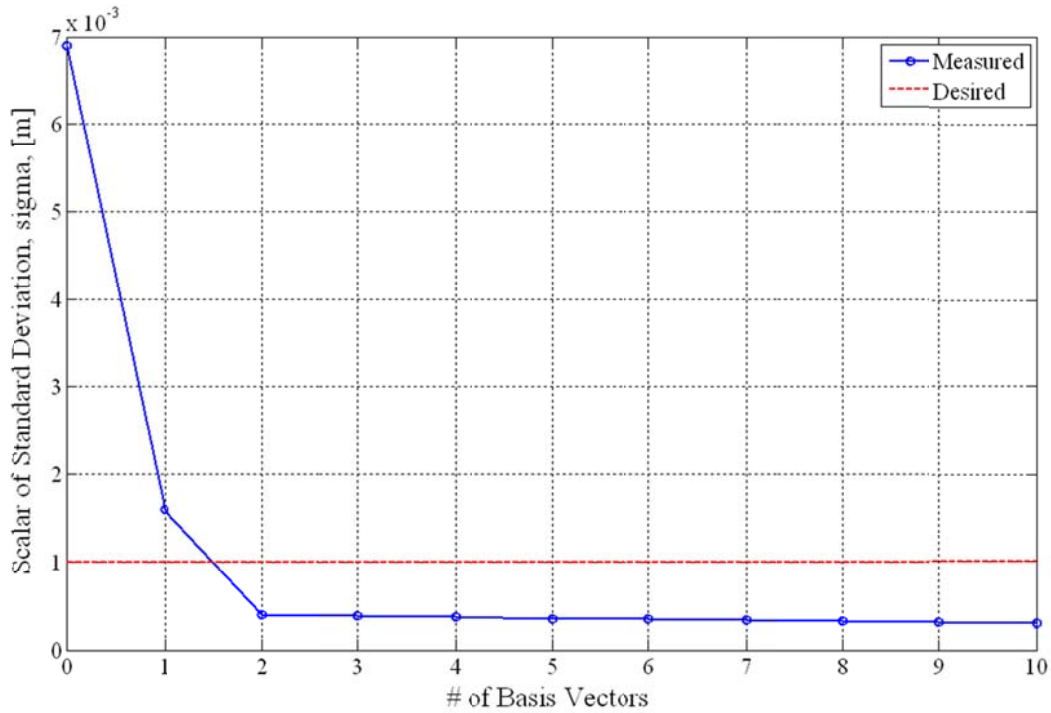
Figure 32. Model order study of summary measure of standard deviation as compared with desired value.

The overall results of the residual analysis and model order study indicate that the fidelity of the measurements have been maintained after implementation of the INS drift removal method. Based on the analysis conducted in this Section, the roll component of drift was identified and removed while still preserving the true road bank angle. Further

discussion on methods to mitigate the effects of γ_l is presented in the Discussion Section of this Chapter.

3.5 Chapter 3 Discussion

This work develops a method to address the INS errors occurring in the height measurements of terrain surface measurements. It was assumed that the data being considered had been discretized along a path-centered coordinate, forming a curved regular grid (CRG). Inherent in this transformation is the spatial interpolation required to estimate the height of the surface at each grid point. Clearly the uncertainties in the pitch and yaw measurements will affect the accuracy of the measurement of the horizontal locations of the measured heights. This uncertainty can affect the estimation of the height estimate at each grid location to varying degrees depending on the choice of spatial interpolation method. For example, the height estimate will be highly sensitive to uncertainty in the horizontal location if a nearest-neighbor or inverse-distance-to-a-power method is implemented. In every case, the search area around each grid point should exceed the uncertainty in horizontal location and this uncertainty should be minimized (Detweiler and Ferris 2008; Detweiler and Ferris 2009). This concept is further discussed in Chapter 4 of this dissertation.

The estimation of the true surface is affected by the offset, or bias, in each of the principle directions for all of the measurements, γ_l, and the error in estimating the expectation operator, as shown in Equation 43. The measurement offset, γ_l, is unknowable from the measured data, but can be mitigated by prudent experimental design. This was made evident in the example, where the data were collected in opposing directions. In this way, the drift in the roll measurement was distinguishable from the cross-slope in the terrain surface. If all the data were collected in the same direction, then drift in the roll direction would have been indistinguishable from actual cross-slope in the surface, and the resulting contribution of, γ_2, would have been significantly larger, yet unseen in the estimate of the true surface. In contrast, the error contributable to the first drift direction, the vertical measurement offset γ_1, cannot be mitigated by experimental design. For example, one cannot drive on the surface of the terrain, then turn the vehicle upside-down and perform subterranean driving on the

bottom of the surface. If this were possible, then the drift in the direction of vertical bias in the measurement could be distinguished from the actual height changes in the terrain. A suggestion to alleviate the unobservable component of γ_1 would be to compare the measured data with surveyed data to identify the 'true' variations of height of the terrain at specific locations (landmarks). It should be clear that if data can only be collected in one direction (as is the case when a road is not closed to traffic) that a static survey of the road bank angle at specific locations would be required to determine γ_2. The additional information provided by the surveyed landmarks is a collocation enhancement to the method developed in this section. That is, the offset, or bias, that is unknowable from the measured scanning laser data can be identified using a collocation approach with additional survey data, reducing the error in the estimate of the true surface. However, these survey data are not generally available and the effects of the bias can be identified and removed through careful experimental design and technique.

With careful planning, γ_2 can be mitigated by driving on each section of road in alternating directions. In this way the drift component in the roll direction can be differentiated from the road banking. Similarly, higher order effects such as road crowning and rutting can be identified by acquiring data at different lateral locations in each lane. For example, road crowning could be differentiated by a parallax issue by acquiring terrain data with the vehicle centered on the lane, +0.5m of center of the lane and -0.5m of center of the lane. In this way, the drift component due to parallax could be differentiated from true road crowning. Similar data acquisition techniques could also identify and differentiate any drift components associated with road rutting.

The methodology developed in this work examines one variable at a time, such that each component of drift is identified as decoupled from the next component due to the enforcement of orthogonality. As a result, each drift component and the effect exhibited on the error surface are studied individually. Another approach to error modeling is the implementation of response surface methodology. Such an approach is aimed at modeling the response of the system based on the coupling of two or more variables and studying the response. In the model presented in this Chapter, it seems logical that elevation is decoupled from bank angle, because the elevation does not change based on changes in bank angle. However, bank angle could affect the amount of

crowning in the terrain, or asymmetry may affect and even skew the amount of rutting that exists in the terrain surface. By implementing response surface methodology to further develop the INS drift error model, the results of each error model can be studied to identify better experimental methods to identify and remove error associated with coupled variables. This is similar to the method described previously where acquiring data in opposing directions on the same section of road identifies the roll drift component of error to be removed from the measurement. Similarly, coupled terms could be identified and removed from the realizations using response surface methodology. Future work will focus on and study the effects of coupling certain variables to improve the overall error model.

As identified in the theory development of this work, the true expectation cannot be known exactly, but must be estimated from a finite number of available measurements. The interpolation scheme that is selected to estimate the expectation operator is consequential to computational efficiency, outlier sensitivity, location sensitivity, and trend sensitivity (Detweiler and Ferris 2009). Recall the plot of the longitudinal data with no error compensation (i.e., q=0) shown in Figure 20. The ten measured profiles show +/-10 mm of vertical drift, which is consistent with the static test results shown in Figure 14. This may suggest that error contributions from the IMU can be considered negligible as it can be seen from Figure 20 that the DGPS is the primary contributor to the system drift as the drift error is identical in magnitude to the stationary test. Additionally, one of the profiles follows a slightly different trend then the other nine profiles, demonstrating the importance of using an estimator that is robust with respect to outliers in the data. Although the mean is the least squares estimator, it is based on the assumption that all realizations should be treated equally and therefore cannot be used to identify and mitigate the effects of outliers. Therefore, the mean estimator of expectation was not used in this work since outliers would skew this estimate of the true surface. This choice of the proper interpolation method was addressed by Detweiler, Stein and Wackernagel (Stein 1999; Stein, Meer et al. 2002; Detweiler and Ferris 2008; Detweiler and Ferris 2009; Wackernagel 2010). For this study, the trimmed mean interpolation method was selected: the mean of the inner two quartiles of data was used.

When a more comprehensive set of terrain surfaces have been studied, it may be possible to form an archetypal set of basis vectors for use in all terrain models. The results of this study are promising; it appears that the first basis vector can be defined as a constant vertical offset and the second basis vector can be defined as a the 'bank angle'. A third basis vector may be defined to describe the "crowning' of the road. Proper experimental technique, such as acquiring some measurements of the surface at varying lateral locations (e.g., to the left and to the right of the lane center), would then be used to differentiate the crowning of the road from any parallax issues in the measurement system. Similarly, an additional basis vector may define "road rutting". The method developed in this work highlights the importance of understanding the interdependence of the analysis techniques applied to measured data and proper data acquisition techniques, and how this understanding allows the identification of instrumental drift and the ability to minimize the effect on the estimate of the true terrain surface.

3.6 Chapter 3 Conclusions

Terrain measurement systems equipped with an INS have drift that skew the estimated height and orientation of the terrain surfaces in the global reference frame. A compensation method is developed to correct this INS drift in terrain measurements. Each terrain surface measurement is considered a combination of the true surface and the error surface in which the error is decomposed into drift (global error) and noise (local error). The global and local subspaces are constructed such that the drift is modeled as a random walk process while the noise is a zero-mean process. This theoretical development is coupled with careful experimental design to develop a method to identify the drift component of error and discriminate it from true terrain surface features, and correct for the INS drift. It is shown through an example that this new compensation method dramatically reduces the variation in the measured surfaces from +/- 10mm to within the resolution of the measurement system itself: +/- 1mm. This example required two basis vectors- elevation and bank angle, to establish the global subspace and reduce the noise to a zero-mean process with a standard deviation of 0.35mm. The implementation of this method has implications for a wide range of applications in which accurate terrain surfaces are required. In terms of ride quality, this method provides a

means to appropriately stitch multiple lanes of road together without the effects of unwanted discontinuities to perform various handling simulations. For example, a simulated vehicle lane change, where the two lanes are concatenated from separate terrain surfaces, will provide a representative surface with negligible discontinuity. An archetypal set of basis vectors is developed in the next Chapter for use in classifying terrain types for the development of terrain models, and of the development of models of the principle terrain surface features.

4. Using a Galerkin Approach to Define Terrain Surfaces

4.1 Chapter 4 Introduction

It is clear that the terrain is the main excitation to the vehicle (Aurell and Edlund 1989). Knowledge of this excitation, when applied in conjunction with high-fidelity tire and vehicle models, would allow chassis loads to be accurately predicted in vehicle simulations. Therefore, throughout the design process the system response to this consistent excitation can be calculated and compared for each tire and chassis design considered. Accurate terrain models would then provide the tire and chassis designers with a powerful tool to make informed design decisions early in the design process while changes are relatively inexpensive to implement. This will result in shortened vehicle development time and reduced overall development costs.

The current practice for vehicle and tire simulations is to use terrain profile data as the input to tire and vehicle models. The terrain profile is an indexed set of terrain heights extending longitudinally along each wheel path. These profiles can be considered signals and in many cases can be modeled as a stochastic process, thus creating a compact representation of the terrain profile. Although computationally efficient, this compact representation limits the available fidelity of data for use in tire models. It would be advantageous to the design engineer to implement terrain surface data as the input to tire and vehicle models since terrain surfaces capture more detailed information about the tire contact patch.

The terrain surface can be represented as an indexed set of transverse vectors of terrain height, represented in matrix form. That is, each vector of the matrix comprises a set of terrain heights located perpendicular to the direction of travel of the vehicle. Although terrain surface data are more computationally expensive to use in simulation, they typically provide better estimates of the tire-terrain interaction. Toward this end, this Chapter proposes a method to compactly represent terrain surfaces through analytic functions.

The objective of this Chapter is to develop a terrain surface representation that retains sufficiently high-fidelity information for detailed vehicle simulation, possesses the simplicity of terrain profiles, and is being insensitive to the choice of the terrain measurement system from which the data were acquired. The remainder of this work is developed as follows. Background on several gridding techniques used across various areas of studies is introduced. Then, the method by which currently available terrain measurement systems acquire data is provided to clarify the concept of terrain surfaces as a sequence of vectors. A vector space is formed by the span of these vectors and a corresponding set of empirical basis vectors is defined. This background is used as the foundation for the developments in this work. A range of possible analytic basis vectors are then generated using known sets of polynomials. The polynomial sets are parameterized to provide the minimized root mean square error (RMSE) fit between the resulting analytic basis vectors and the empirical basis vectors. A weighted inner-product is defined to form a Hilbert space and is use to project the terrain surface vectors onto the set of analytic basis vectors. The weighting matrix is developed such that these projections are insensitive to the number and placement of the discrete transverse locations at which the terrain heights are defined. In this way, a single set of analytic basis vectors are determined for a general class of terrain (such as all U.S. Highways), thus eliminating the need to calculate empirical basis vectors for individual surfaces in this class. This representation also allows less capable terrain topology measurement systems to estimate the main features of terrain surfaces. The method is applied to experimental data to demonstrate the concept and to form a basis for discussion, followed by concluding remarks.

4.2 Chapter 4 Background

4.2.1 Terrain Measurement Approaches Classified as Collocation or Galerkin Methods

Terrestrial terrain measurement systems sample the terrain surface based on the number and types of optical lasers implemented. Refer to Chapter 2, Section 2.1.1 for a complete review of terrestrial terrain measurement systems. Two main approaches exist to optically sample the terrain; the first is to use point lasers to measure longitudinal

profiles at specific transverse locations, and the second is to use a scanning laser that is capable of measuring a swept path at a higher concentration of sampled locations along the measured path. Consider the following demonstrative example, where it is assumed that the data are acquired from a scanning laser and mapped to a uniform grid spacing. Let s_i be the vector representing the true surface of the section of terrain at the longitudinal location indexed by i. Let vector s_i contain elements indexed 1 to 80 corresponding to transverse points in the $v(j)$ direction, such that the terrain section has a width of 80 points, as shown in Figure 33. Furthermore, let the set of vectors $\{s_i\}$ span 10,000 locations in the $u(i)$ direction, such that the terrain section has a length of 10,000 points. Using this representation, $\{s_i\}$ contains 800,000 elevation points to describe the section of terrain. The remainder of this section will leverage this example and discuss a collocation method and a Galerkin method to define the terrain surface.
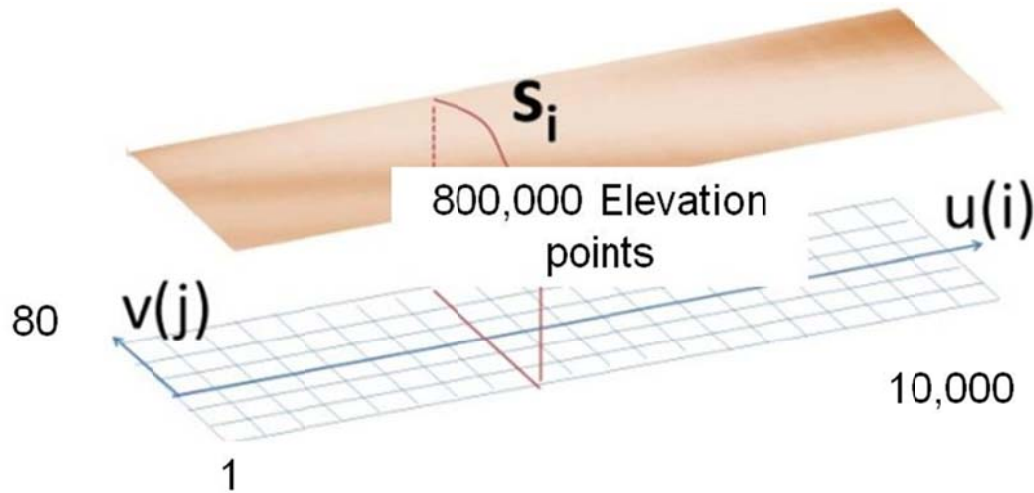
Figure 33. Example of discretized terrain surface in Curved Regular Grid Format.

The use of point lasers to sample the data can be considered a collocation method to characterize the terrain. The collocation method is based on a finite-dimensional space of candidate solutions, such as longitudinal terrain profiles, and a number of collocation points in the domain of the space, such as grid points. It is then desired to select the solution which satisfies the given equation at the collocation points, or in this case, select the solution which exactly represents the terrain profile at the each grid point of interest. In the case of point lasers, as can be seen from Figure 34, if lasers are placed along the wheel paths of a vehicle and sampled at the required grid points, then any spatial

interpolation method between these grid points must, by definition, be a collocation method provided that the equations exactly represent the grid points themselves.. By implementing a collocation approach and referring to the example presented earlier in this section, it would require 10,000 points to exactly represent each wheel path profile.
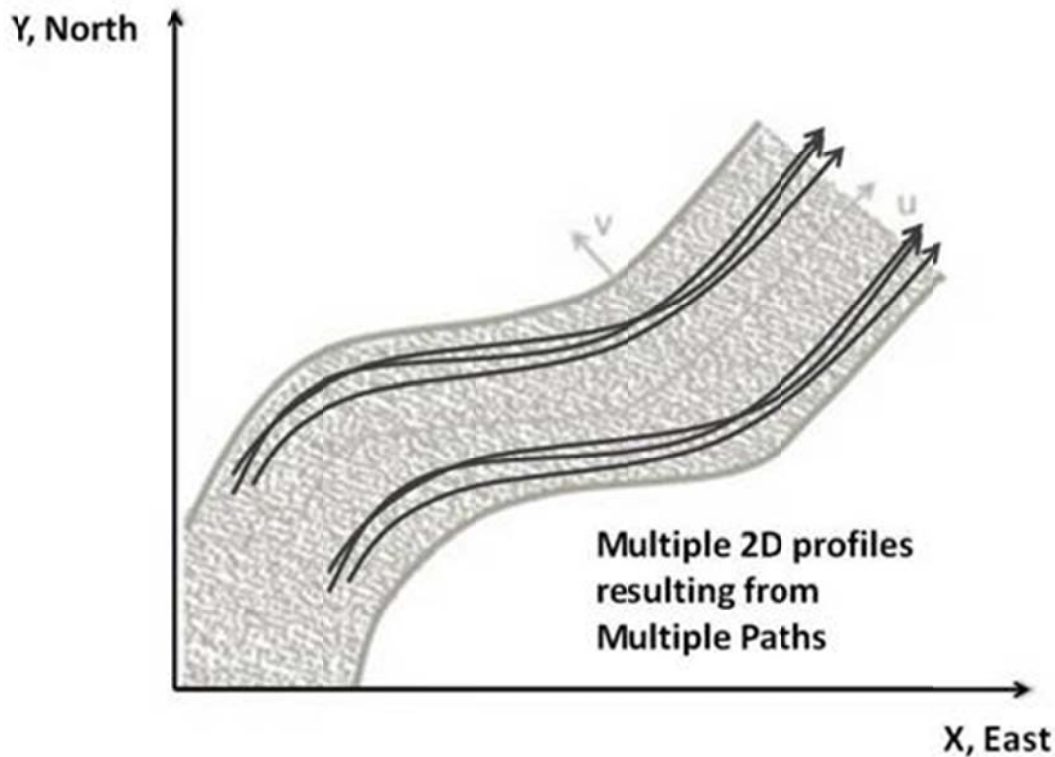
Figure 34. 2D terrain profiles resulting from point lasers used to characterize the surface by a Collocation method.

On the contrary, scanning lasers are capable of measuring the full width of a typical highway lane as depicted by Figure 35, which allows the entire surface to be characterized as opposed to just wheel path profiles. This method of approximating the terrain surface can be considered a Galerkin method of characterizing terrain. The key difference between a Galerkin method and a collocation method is that the Galerkin method satisfies conditions in an integral or average sense over all points, rather than an exact solution at every grid point, thus allowing the entire surface to be approximated. With the Galerkin method, it is acceptable to formulate a set of continuous analytic basis functions that can be discretized based on the capabilities of the measurement system for specific types of terrain. Once the analytic basis functions are identified for a sample set

96

of terrain using scanning lasers capable of measuring the full width of a typical highway lane, the effects of fewer measured samples of the terrain surface can be analyzed.
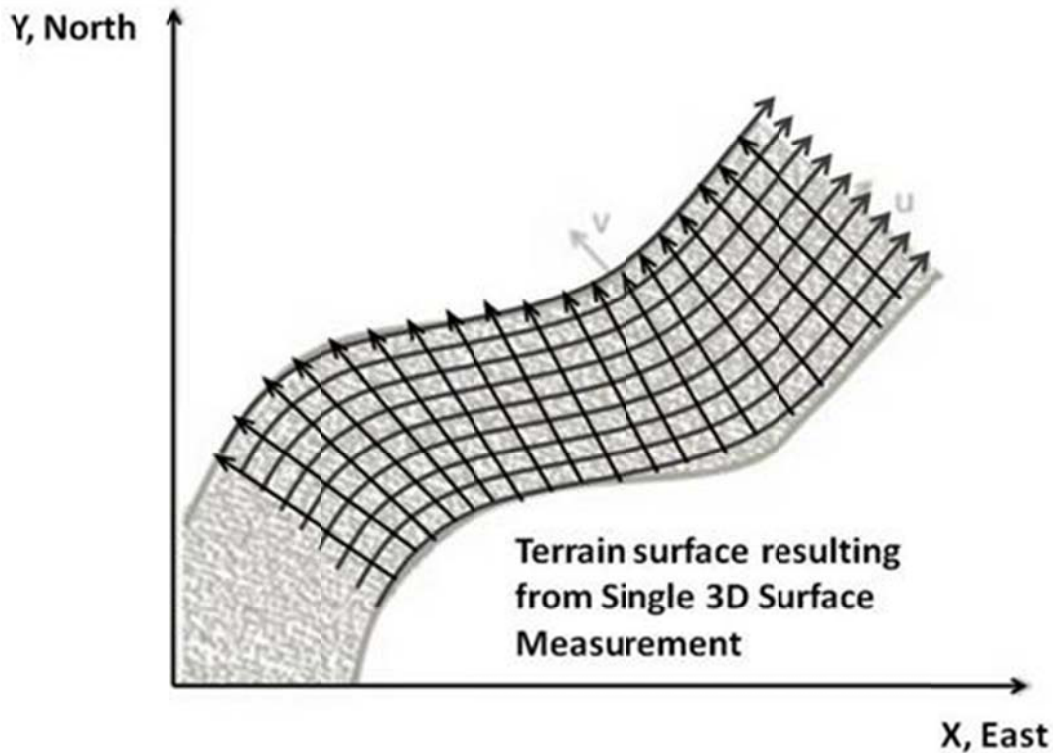
Figure 35. 3D terrain surface resulting from a scanning laser used to characterize the terrain by a Galerkin method.

4.2.2   Terrain Analysis Methods

Many methods are available for use in discretizing a continuous surface for analytical purposes and vary based on the application. In the application of terrain surface analysis, the surface is already discretized based on the terrain measurement system used to acquire the data. The data are acquired in the form of point-cloud data and thus, the issue of analysis technique arises based on the spatial interpolation used to map to a regularly spaced grid. Several methods of formulating a uniform grid are reviewed to identify key differences, and clearly describe the method developed in this dissertation.

Finite Element Analysis

Finite Element Analysis (FEA) is used to numerically solve field problems. A field problem is classified as a differential equation that requires the formulation of finite elements to produce a solution in terms of a spatial distribution of one or more dependent variables (Cook, Malkus et al. 2002). These finite elements are defined by grid nodes. Various elements exist to better solve different problems, the details of which are beyond the scope of this work. The collection of elements and grid nodes is commonly referred to as a mesh. The mesh can be adjusted to better fit the contours of the continuum being analyzed. The main advantage of FEA is that this method represents and solves for the dependent variable of interest locally, (between nodes), to approximate how the continuum will respond to specific inputs and boundary conditions. Consider the following example to demonstrate how the underlying principles of FEA work. An element comprised of four nodes is presented in Figure 36, implementing the ξ-η coordinate system. The shape functions, N_i, also known as basis functions, associated with each node are defined in Equation 52, where i corresponds to the node number. The displacement field, u, corresponding to this example is presented in Equation 53. Referring to the example presented earlier in this section, if the FEA method was implemented on the surface in its current configuration, then there would be 800,000 nodes and 800,000 shape functions required to define the displacement field of the sample terrain surface. Of course the beauty of using FEA is the mesh can be refined (coarser or finer) to appropriately represent the continuum under study. Furthermore, material properties are associated with the mesh so isotropic or anisotropic properties are defined prior to analysis. This method is applicable to a large general class of problems, including capturing and characterizing the details of local terrain features such as potholes or bumps. It is computationally expensive, however, when applied to anisotropic terrain surfaces that define the scope of applicability of the developments in this work. Specifically, the scope is terrain having an inherent path defined over which vehicles travel and where that travel, over time, has defined obvious paths or lanes.

$$N_i(\xi)= \tfrac{1}{2}(1-\xi)$$

$$N_i(\eta)= \tfrac{1}{2}(1-\eta)$$
<div align="right">Equation 52</div>

$$N_i(\xi,\eta)= \tfrac{1}{4}(1-\eta)(1-\xi)$$

$$u(\xi,\eta)=N_1(\xi,\eta)u_1+ N_2(\xi,\eta)u_2+ N_3(\xi,\eta)u_3+ N_4(\xi,\eta)u_4$$
<div align="right">Equation 53</div>
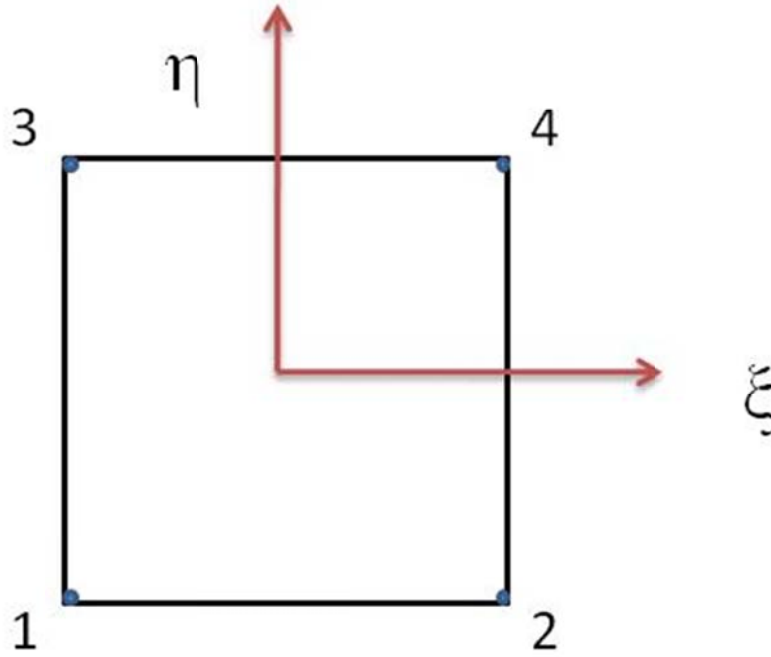
Figure 36. Simple four node FEA example.

Digital Terrain Model Analysis

Digital Terrain Models (DTM) are discussed in detail in Section 2.1.2, and the points that are important to this Chapter are briefly reviewed here. A Cartesian coordinate set is established in the horizontal plane (defined by X, and Y), and polynomials are formed that may include coupling terms for X and Y due the assumption that the terrain being studied is isotropic. The polynomials are parameterized in one of several ways. The order of the polynomial (as well as the number of the corresponding parameters) can be increased such that the measured points are matched exactly (a collocation approach) or lower order polynomials may be parameterized to minimize some integral error for all measured points (a Galerkin approach). In either case, the

order and form of the polynomial is developed without *a priori* knowledge of the specific anisotropicity addressed in this work. Similar in this way to the FEA method, the DTM method is applicable to a large general class of problems, but is computationally expensive when applied to anisotropic terrain having an inherent path defined over which vehicles travel and where that travel, over time, has defined obvious paths or lanes. In general, the DTM method is appropriate for modeling terrain on the global scale where localized terrain features such as potholes and bumps are inconsequential.

Principle Component Analysis

The terrain studied in this work has well defined paths. This specific anisotropicity is exploited, allowing the longitudinal and transverse directions to be decoupled and a single set of vectors (discretized polynomials) to be used to define the transverse direction at all longitudinal locations. That is, the entire terrain surface can be represented by several basis vectors as opposed to thousands of shape functions as with the FEA method, or with high order polynomials as with the DTM method. A Singular Value Decomposition (SVD) algorithm is used to develop these basis vectors since the SVD algorithm minimizes error based on the error type chosen, has high numerical precision (minimal round-off error), and controls linear dependence among the columns of the matrix by setting very small singular values approximated to zero (Haykin 2003). It can be applied to square or rectangular matrices. As a result of implementing SVD on the terrain sets, the principle components, or basis vectors are identified. In the case where the path is indistinguishable, more general methods of parameterizing the polynomials should be applied. The method proposed in this work would require a large number of basis vectors to be developed and would simply converge to an uncoupled formulation of the DTMs in the limiting case.

4.2.3 Defining the Hilbert Space and True Surface

Each vector composing the terrain surface measurement, $z_{i,k}$, is considered a combination of the true surface, s_i, and a realization of a stochastic error process, $e_{i,k}$. The error is decomposed into drift (global error) and noise (local error); drift is modeled as a random walk process while the noise is a zero-mean process (Favey, Cerniar et al. 1999; Sun, Fu et al. 2008). Each measured surface is then a realization of a stochastic

process and is defined as a sequence of vectors that are elements of a Hilbert space, **H**. Since there is exactly one true surface for any number of realizations, s_i is not indexed by k. It is shown that the true surface must also span the Hilbert space **H** (Chemistruck, Binns et al. 2010).

The Hilbert space is decomposed into principle directions. A set of orthonormal, empirical basis vectors, b_l, are constructed to span the space via Singular Value Decomposition (SVD) of the set of measured vectors. Basis vectors are added to the composition of the true surface until the residual noise vectors have a mean that is not statistically different than zero; this number of basis vectors is defined as q. The span of these q basis vectors defines the global subspace, **G**. The order in which the basis vectors are added is determined by the magnitude of their corresponding singular value, in descending order. Currently, the global subspace is defined by the span of the truncated set of empirical basis vectors. In this Chapter, the global subspace is redefined by the span of a truncated set of analytic basis vectors.

Due to the finite number of realizations (measurements), the true surface, s_i, is determined by considering the estimated expectation of the sequence of measured vectors projected in the global vector space, **G**, and the complementary vector space separately. The vector defining the estimated true terrain elevation at the longitudinal location indexed by i is then defined as $\hat{s}_i$. This estimation of the true surface must be coupled with careful experimental design. Specifically, the drift component of error is identified and removed from the measured terrain and the expected value of this drift-free surface is the estimated true terrain surface, thus correcting for the INS drift (Chemistruck, Binns et al. 2010). Note that

Equation 54 would reduce to the expected height, $E[z_{i,k}]$, if the expectation were known exactly or a linear estimator were used.

$$\hat{s}_i \triangleq \widehat{E}\left[\sum_{l=1}^{q} \langle b_l, z_{i,k}\rangle b_l\right] + \widehat{E}\left[z_{i,k} - \sum_{l=1}^{q} \langle b_l, z_{i,k}\rangle b_l\right] \qquad \text{Equation 54}$$

Samples of paved U.S. Highway data present very similar trends in their resulting empirical basis vectors; a typical set of vectors is shown in Figure 37. Generally, the first

basis vector accounts for a vertical offset in elevation and the second basis vector describes the 'cross-slope' or 'bank angle'. This agrees with intuition that the primary changes of a paved surface are in elevation and banking. The fourth basis vector approximates the shape of 'rutting,' a condition that arises from road wear due to tire interaction. Due to this consistency in road construction of paved surfaces and interaction with vehicles, the shapes of the empirical basis vectors are correspondingly similar for all terrain surfaces that have a distinct path. This work develops a method to capture these characteristics with analytic functions. The correlation between empirical basis vectors and analytic basis functions is exploited to develop a single set of analytic basis vectors that can be used for a specific type of terrain. These analytic basis vectors are developed to provide a definition of terrain surfaces that is insensitive to the choice of terrain measurement system and can be applied to terrain that have similar topology. In this way, new empirical basis vectors do not have to be developed and implemented in each characterization of terrain and direct comparison of terrain of the same terrain type can be performed. To achieve this, the empirical basis vectors are approximated by a set of analytic functions for specific types of terrain; the analytic functions are sampled at discrete locations defined by measurement locations provided by the measurement system or the prescribed gridding locations.
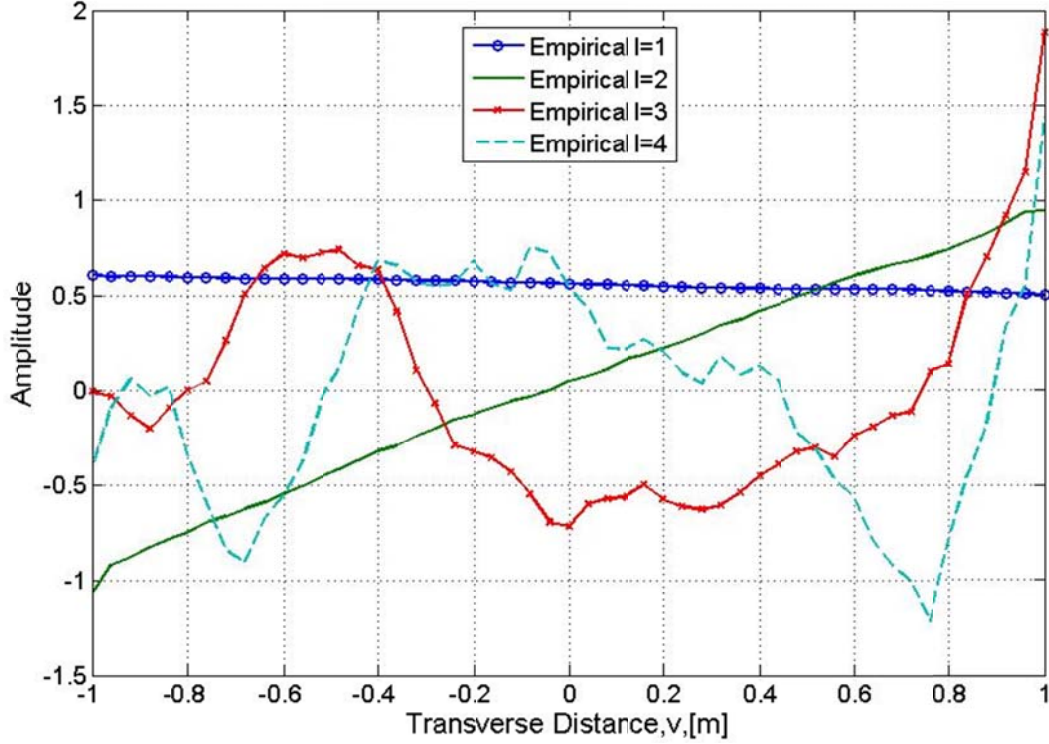
Figure 37. First four empirical basis vectors based on a sample of measured U.S. Highway data.

4.2.4 Analytic Functions

Several polynomial sets are introduced in this section and briefly defined as potential candidates to approximate the empirical basis vectors. Well-defined and commonly used sets of polynomials are used instead of curve-fitting polynomials to the empirical data so that definitions of the terrain surface are portable and cross-platform terrain measurements can be implemented.

Gegenbauer Polynomials

The Gegenbauer polynomials, more commonly known as ultraspherical functions, belong to the class of orthogonal polynomials. Normality is enforced after the functions are discretized. A recurrence relationship for the Gegenbauer polynomial is presented in

Equation 55, with parameter λ (0,1] and are defined over the domain [-1, 1]. The Legendre polynomial results when λ is equal to 0.5. The shapes of the first five Legendre polynomials are presented in Figure 38 and upon inspection it can be seen that these

103

analytic functions approximate the empirical basis vectors shown in Figure 37. The tails are set to +1 or -1, but the local minimums and maximums vary based on the order of the polynomial. Additional detail on the Gegenbauer Polynomials and the Legendre Polynomials can be found in Appendix C and Appendix D respectively.

$$nC_n^{(\lambda)}(x) = 2(n + \lambda - 1)xC_{n-1}^{(\lambda)}(x)$$

$$- (n + 2\lambda - 2)C_{n-2}^{(\lambda)}(x), \quad \forall n = 2,3 \dots$$

Equation 55
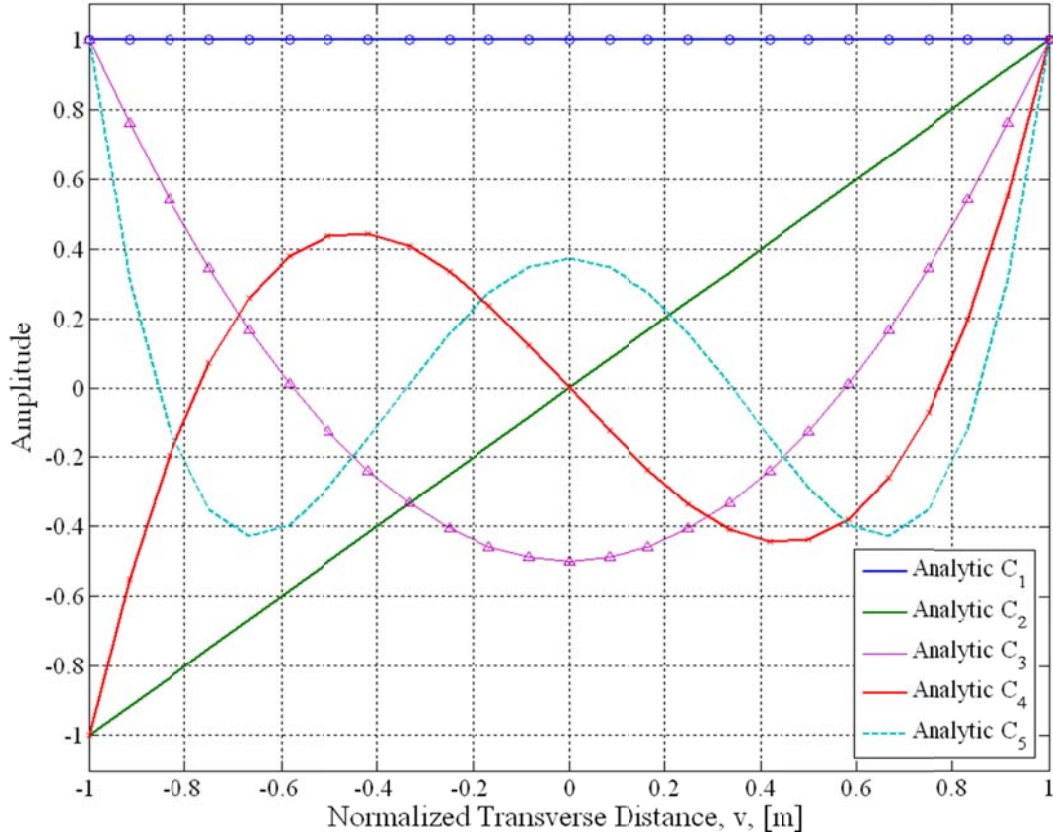
Figure 38. First five Gegenbauer polynomials with λ=0.5.

Chebyshev Polynomials

The Chebyshev polynomials also belong to the class of orthogonal polynomials. Normality of the polynomials is enforced after the functions are discretized. A recurrence relationship for the Chebyshev polynomials is presented in Equation 56, after setting T_0 to 1 and T_1 to x over the range of [-1, 1]. The shapes of the first five Chebyshev polynomials are presented in Figure 39. One of the properties of the Chebyshev polynomials is that the local minimums and maximums are set to one, and all of the tails

start or end at -1 or +1. Additional detail on the Chebyshev Polynomials can be found in Appendix E.

$$T_k(x) = 2xT_{k-1}(x) - T_{k-2}(x) \ \forall k = 2,3,\dots \hspace{2cm} \text{Equation 56}$$
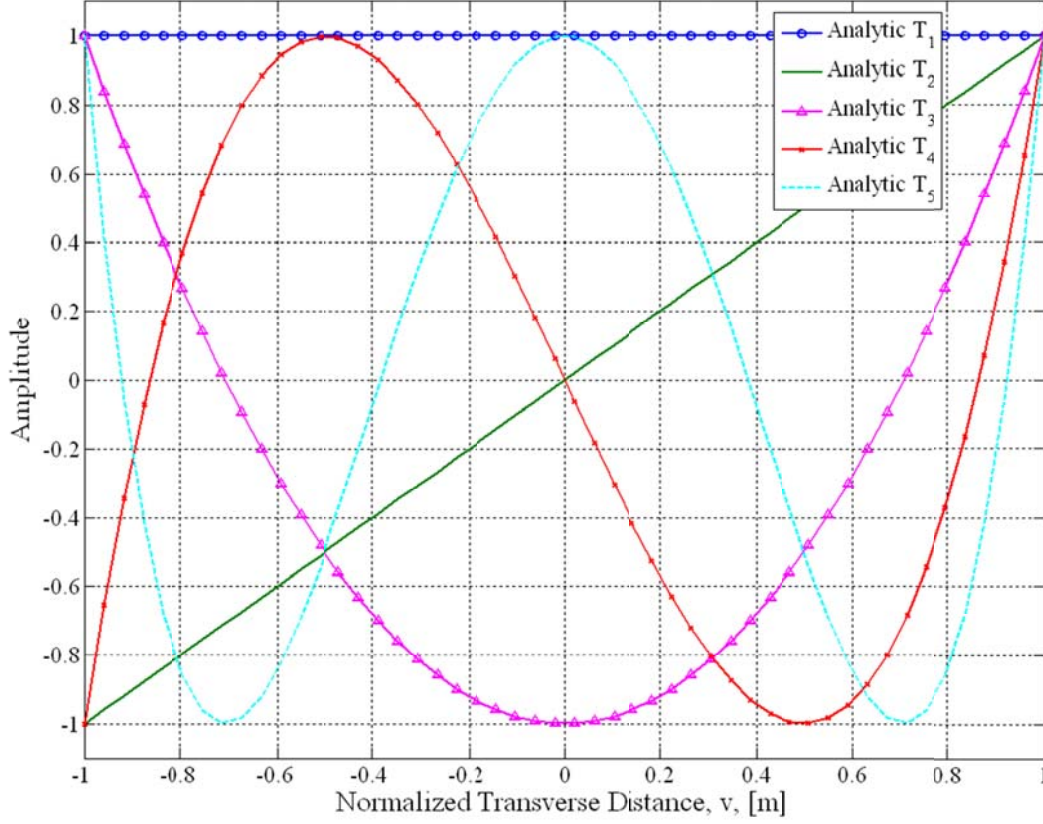
Figure 39. First five Chebyshev Polynomials.

Fourier Series Basis Functions

The Fourier series basis functions make use of the orthogonality relationships of the sine and cosine functions. The Fourier series form a complete orthogonal system over [-π, π] and are defined by the function presented in Equation 57. The coefficients a_0, a_n and b_n are defined in Appendix F. The shapes of the first four Fourier series basis functions are presented in Figure 40. Only the first four basis vectors are presented because the Fourier series basis functions do not include a linear (sloped) line as do the Gegenbauer and Chebyshev polynomials; thus, to match the shapes of these polynomials without the sloped line, only four basis functions are required. Since the Fourier series are based on sines and cosines, the tails of the basis functions will start or end at +1 or 0, and the local

minimum and maximums are equal to +1 or -1. Additional detail on the Fourier Series Polynomials can be found in Appendix F.

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n cos \frac{n\pi}{l} x + \sum_{n=1}^{\infty} b_n sin \frac{n\pi}{l} x \qquad \text{Equation 57}$$
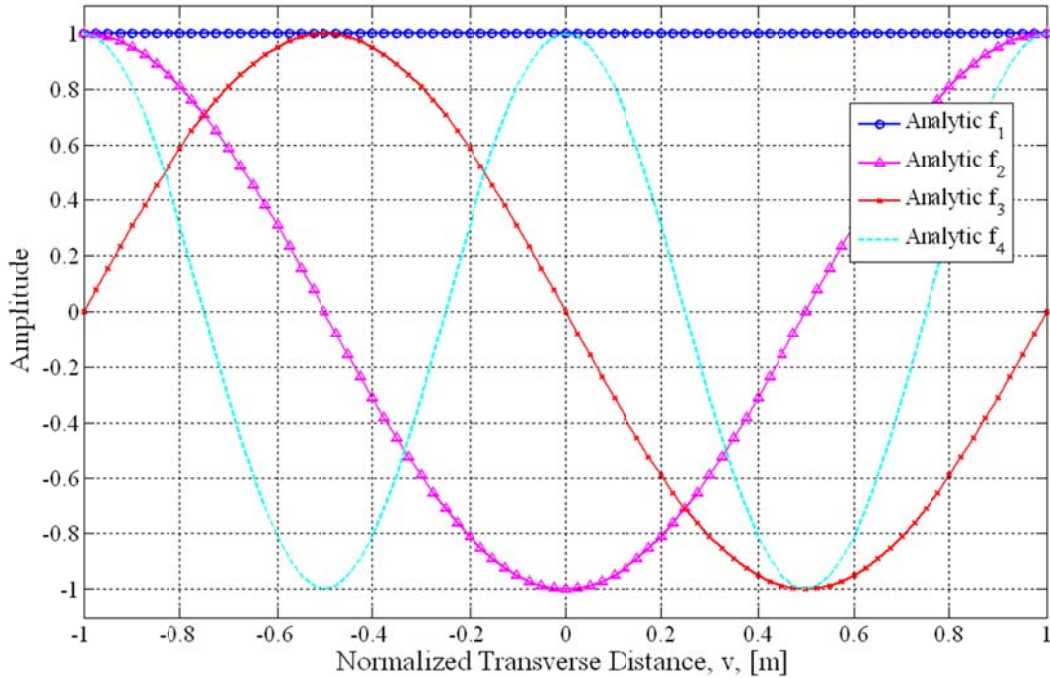
Figure 40. First four Fourier series basis functions.

4.3 A Galerkin Approach to Decomposing Terrain Surfaces

The estimated true terrain surface can be decomposed into contributions in the principle directions defined by the empirical basis vectors, b_l, (Equation 54). In this work, the estimated true surface is decomposed via analytic basis vectors, p_l, that capture the fundamental shapes of the empirical basis vectors. The analytic basis vectors, p_l, are defined from continuous analytic basis functions (shown in

Equation 55) using a Galerkin approach, thus redefining the global subspace, **G**, with a finite set of analytic basis vectors. The definition of the estimated true surface is redefined in Equation 58. Each component of the estimated true terrain surface is defined in terms of the principle directions, $\sigma_{i,l}$, as shown in Equation 59. The magnitude of each

principle direction along the length of the terrain surface is defined as $\sigma_{i,l}$, and is termed the component of terrain in the lth principle direction.

$$\hat{s}_i \triangleq \widehat{E}\left[\sum_{l=1}^{q} \langle p_l, z_{i,k} \rangle_w p_l\right] + \widehat{E}\left[z_{i,k} - \sum_{l=1}^{q} \langle p_l, z_{i,k} \rangle_w p_l\right]$$

Equation 58

$$\sigma_{i,l} \triangleq \langle p_l, \hat{s}_i \rangle_w$$

Equation 59

Several criteria are imposed in the development of the analytic basis vectors. Specifically, the analytical basis vectors must satisfy the properties enumerated below:

1. Spanning the same *n* dimensional vector space as the empirical basis vectors.

2. Orthogonal with respect to a weighted inner-product. This is a fundamental requirement of the Galerkin approach, which enables simple mapping of the estimated true terrain surface onto the analytic basis function.

3. Normalized with respect to a weighted inner-product, simplifying subsequent calculations.

4. Closely correlated to the empirical basis vectors.

5. Insensitive to the selection of the transverse locations composing *v*.

The remainder of this section develops the method to determine the analytic basis vectors.

It is clear that if the complete set of *n* analytic basis vectors is linearly independent, then they span the same space as the empirical basis vectors. However, a different Hilbert space must be formed. The redefined Hilbert space, **H**, has a weighted inner-product and induced norm. The analytic functions are mapped on **H** through Galerkin discretization so that the discretization error is orthogonal to **H**. When the continuous polynomials are discretized, the error associated with this discretization method is orthogonal to **H**, and no discretization error exists in **H**. The global subspace **G**, is redefined as the span of the truncated set of analytic basis vectors, $\{p_l\}$. The drift is then the projection of the error vectors onto the global subspace and the noise is defined as the error projected onto the complement of the global subspace. Again, basis vectors are added to the definition of the global subspace until the residual noise vectors have a

mean that is not statistically different than zero. With the Hilbert space redefined, the parameterization of the polynomials can be determined.

Recall v is the continuous coordinate (shown in Figure 7A) and note that any continuous polynomial can be written in the form presented in Equation 60. The order of the polynomial is defined by (a-1).

$$P(v) = c_0 + c_1 v + c_2 v^2 + \cdots + c_n v^{a-1} \qquad \text{Equation 60}$$

Next the continuous coordinate v is sampled at a sequence of n discrete points to form a vector of monotonically increasing values, $\boldsymbol{v}$. This vector represents the transverse locations at which the terrain heights are defined either through measurement or spatial interpolation. The transverse locations, $\boldsymbol{v}$, are scaled and centered on zero so that they exist on the interval [-1, 1], corresponding to the domain of the continuous polynomials. In this way, each discrete polynomial vector, $\boldsymbol{p}$, is represented in matrix form as shown in Equation 61. The matrix of polynomial terms for discretized values of $\boldsymbol{v}$ is defined as $\boldsymbol{V}$. The vector of polynomial coefficients, $\boldsymbol{c}$, is determined by the choice of the generating polynomial (e.g., Gegenbauer, Fourier, Chebyshev, or others…). The remainder of the theoretical developments in this Chapter are based on the set of Gegenbauer polynomials to avoid redundancy and without loss of generality. The Gegenbauer polynomial is parameterized by λ according to Equation 55.

$$\boldsymbol{p} = \boldsymbol{p}(\boldsymbol{v}, \lambda) = \begin{bmatrix} 1 & v_1 & v_1^2 & \cdots & v_1^{a-1} \\ 1 & v_2 & v_2^2 & \cdots & v_2^{a-1} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & v_n & v_n^2 & \cdots & v_n^{a-1} \end{bmatrix} \begin{Bmatrix} c_1 \\ c_2 \\ \vdots \\ c_a \end{Bmatrix} = \boldsymbol{V}(\boldsymbol{v})\boldsymbol{c}(\lambda) \qquad \text{Equation 61}$$

4.3.1 Vandermonde Matrix

In general, the matrix $\boldsymbol{V}$ is an n x a matrix where the number of polynomial terms, a, will be larger than the number of transverse locations, n. To formulate a Vandermonde matrix for $\boldsymbol{V}$, consider a truncated polynomial of order n. Since $\boldsymbol{v}$ is defined to be monotonically increasing (thereby having distinct values), then $\boldsymbol{V}^{\boldsymbol{1}}$ exists, $\boldsymbol{V}$ is positive definite, and the mapping from the polynomial coefficients to the solution of the

polynomial along v is a one-to-one correspondence (i.e., the polynomial interpolation problem is solvable with a unique solution).

Consider the complete set of n analytic basis vectors, $\{p_l\}$ where $l \in \{1, 2, \ldots, n\}$, defined by Equation 62. The analytic basis vectors are linearly independent and span the same space as the empirical basis vectors. Each vector in the set is a function of the particular measurement locations (defined by v) and a particular set of coefficients, $c_l(\lambda)$. A concatenation of these coefficient vectors is defined as the coefficient matrix, C, as shown in Equation 63. The Vandermonde matrix, V, and the coefficient matrix, C, are used to develop the weighted inner product and the resulting Hilbert space.

$$p_l(v, \lambda) = V(v)c_l(\lambda) \qquad\qquad \text{Equation 62}$$

$$C = [\, c_1 \ldots c_n \,] \qquad\qquad \text{Equation 63}$$

A weighted inner product is developed to produce a set of analytic basis vectors that are orthonormal with respect to the weighting matrix, W, such that $c_k^T V^T W V c_l$ is the Kroenecker delta function, as shown in Equation 64. Note that this definition satisfies the requirements that an inner product is positive definite since a Vandermonde matrix, V, is positive definite and its inverse must exist. More detail on inner products is provided in Appendix G. Furthermore, the coefficients of the polynomial are defined to be linearly independent due to the orthogonality of the polynmomials, and the inverse of the coefficient matrix, C, must exist. The definition of the weighting matrix that produces this desired orthonormality is defined in Equation 65.

$$\langle p_l, p_k \rangle_W = p_k^T W p_l = c_k^T V^T W V c_l = \delta_{kl} \qquad\qquad \text{Equation 64}$$

$$W \triangleq (V^T)^{-1}(C^T)^{-1}(C)^{-1}V^{-1} = (P^T)^{-1}P^{-1} \qquad\qquad \text{Equation 65}$$

Analytic basis vectors are defined as the orthonormal discrete basis vectors from the Gegenbauer polynomials. Consider the first basis vector, p_1, corresponding to the zero-th order Gegenbauer polynomial. Since p_1 is a vector of ones, the corresponding vector of coefficients, c_1, is given by Equation 66. The second basis vector, p_2, is derived from the first order Gegenbauer polynomial and the corresponding vector of coefficients is given by Equation 67. The coefficient matrix, C, has the form shown in Equation 68.

It is clear that all coefficient vectors except the first are functions of the parameter λ. This parameter is used to adjust the shape of the analytic basis vectors so that they are closely correlated to the empirical basis vectors (satisfying the fourth design criterion). This process is demonstrated for U.S. Highways in the Proof of Concept Section of this Chapter. Next, the general formulation of this method is presented in terms of a weighted grammian.

$$\boldsymbol{p_1} = p_1(\boldsymbol{v}) = \boldsymbol{1} = \boldsymbol{V}\boldsymbol{c_1} \ where \ \boldsymbol{c_1} = \begin{Bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{Bmatrix} \qquad \text{Equation 66}$$

$$\boldsymbol{p_2} = p_2(\boldsymbol{v}) = \boldsymbol{v} = \boldsymbol{V}\boldsymbol{c_2} \ where \ \boldsymbol{c_2} = \begin{Bmatrix} 0 \\ 2\lambda \\ 0 \\ \vdots \\ 0 \end{Bmatrix} \qquad \text{Equation 67}$$

$$\boldsymbol{C} = \begin{bmatrix} 1 & 0 & -\lambda & \\ 0 & 2\lambda & 0 & \\ 0 & 0 & 2\lambda(1+\lambda) & \cdots \\ \vdots & \vdots & 0 & \\ 0 & 0 & \vdots & \end{bmatrix} \qquad \text{Equation 68}$$

4.3.2 Underdetermined Systems

In general, the number of polynomial terms, a, will be larger than the number of transverse measurement locations n, such that $a > n$. This results in an underdetermined system of equations, which, if there is any solution, then there are an infinite number of solutions which form an affine space. It is desired to identify the "simplest" solution to this problem, which is of the form $y=Ax$. The goal is to identify the least squared error solution in this affine space, that is, the solution that minimizes the norm of the error vector (Donoho, Kakavand et al. 2007). Recall the formulation of the problem defined in matrix form in Equation 69. In the general sense, $\boldsymbol{V(v)}$ is an $n \ x \ a$ matrix of polynomial terms for discretized values of $\boldsymbol{v}$ and $\boldsymbol{C}$ is an $a \ x \ n$ matrix of vectors of polynomial coefficients. In the case of Gegenbauer polynomials, the matrix of coefficients is parameterized by λ and written $\boldsymbol{C(\lambda)}$

$$P = P(v, \lambda) = \begin{bmatrix} 1 & v_1 & v_1^2 & \dots & v_1^{a-1} \\ 1 & v_2 & v_2^2 & \dots & v_2^{a-1} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & v_n & v_n^2 & \dots & v_n^{a-1} \end{bmatrix} [c_1(\lambda) \quad \dots \quad c_n(\lambda)] \qquad \text{Equation 69}$$

$$= V(v)C(\lambda)$$

The weighted inner product is reformulated for the general case in which the polynomials are not truncated to $n = a$ as in the case of the Vandermonde matrix. A weighted grammian, R, is introduced to aide in the formulation of the reformulated weighted inner product in Equation 70. Specifically, a weighting matrix is developed such that the weighted inner product of the element $\langle p_l, p_k \rangle_W$ is equal to the Kroenecker delta function. A weighted inner product in matrix form is developed to produce a matrix of analytic basis vectors that are orthonormal with respect to the weighting matrix, W, such that the weighted grammian equals the identity matrix, as shown in Equation 71. Note that this definition satisfies the requirements that an inner product is positive definite since a grammian is always positive definite. The definition of the weighting matrix that produces this desired orthonormality is defined in Equation 72, which is the same as Equation 65. This more general formulation retains all of the higher order polynomial terms, and reduces to the Vandermonde formulation when the order of the polynomial is truncated to the number of measurement locations (i.e., when $a=n$). Even without the truncation, the resulting P matrix is the same for this underdetermined system as it is for the Vandermonde formulation. This can be attributed to the fact that V is an [n x a] matrix and C is an [a x n] matrix, so when multiplied, P becomes an [n x n] matrix. Next, the concept of components of terrain is introduced.

$$R = R(v, \lambda) = \begin{bmatrix} <p_1, p_1> & <p_2, p_1> & \cdots & <p_n, p_1> \\ <p_1, p_2> & <p_2, p_2> & \cdots & <p_n, p_2> \\ \vdots & \vdots & \ddots & \vdots \\ <p_1, p_n> & <p_2, p_n> & \cdots & <p_n, p_n> \end{bmatrix} = [I] \qquad \text{Equation 70}$$

$$R = \langle P, P \rangle_W = P^T W P = C^T V^T W V C = I \qquad \text{Equation 71}$$

$$W \triangleq (V^T)^{-1}(C^T)^{-1}(C)^{-1}V^{-1} = P^{T-1}P^{-1} \qquad \text{Equation 72}$$

4.3.3   Defining Components of Terrain in the lth Principle Direction

The Hilbert space is developed such that the resulting projections onto the analytic basis vectors, $\sigma_{i,l}$, are insensitive to the selection of the transverse measurement locations composing vector $\boldsymbol{v}$ (and subsequently matrix $\boldsymbol{V}$) and with respect to the parameterization of the coefficient vectors, $\boldsymbol{c}_l$, and matrix $\boldsymbol{C}$. This insensitivity is demonstrated through a heuristic example in which the estimated true surface, $\hat{\boldsymbol{s}}_i$, is simply a linear combination of the analytic basis vectors, given in Equation 73. In this simplified case, the estimated true surface is constructed such that it lies completely in the global subspace. The θ_k coefficients are defined, for this example only, as the magnitude of the contribution of each analytic basis vector to this artificial estimated true surface. For this case, the choice of measurement locations (manifested in $\boldsymbol{V}$) and parameterization of the analytic basis vectors (manifested in $\boldsymbol{C}$) have no effect on the determination of the components of the terrain. This is due to the definition of the weighting matrix, $\boldsymbol{W}$, defined in Equation 65 and equivalently in Equation 72. Also by definition in Equation 64, the weighted inner product is only equal to one when k is equal to l and is zero otherwise, which is used in Equation 74. This is also consistent with the grammian formulation. Consider, for example, if the road surface is perfectly flat and level, but vertically offset by a specific value, then according to Equation 73, θ_1 must be proportional to this vertical offset and θ_k for k > 1 must be zero. In this case, the resulting first component of terrain $\sigma_{t,1}$ would be θ_1. Similarly, if the road were flat but banked, then θ_2 would be proportional to the surface's bank angle and all values for θ_k for $k \neq 2$ would be zero. If the estimated true surface is any linear combination of the analytic basis vectors, then the resulting projection onto each analytic basis vector, $\sigma_{i,l}$, given in Equation 74, is completely insensitive to the choice of transverse locations, $\boldsymbol{v}$, or the choice of the parameterization of the generating polynomial, λ.

$$\hat{\boldsymbol{s}}_i = \sum_{k=1}^{q} \theta_k \boldsymbol{p}_k \qquad \text{Equation 73}$$

$$\sigma_{i,l} = \langle \hat{\boldsymbol{s}}_i, \boldsymbol{p}_l \rangle_w = \langle \sum_{k=1}^{q} \theta_k \boldsymbol{p}_k, \boldsymbol{p}_l \rangle_w = \sum_{k=1}^{q} \theta_k \langle \boldsymbol{p}_k, \boldsymbol{p}_l \rangle_w = \theta_l$$

Equation 74

Using this definition of components of terrain, $\sigma_{i,l}$, a new representation is developed that captures the principle information about a terrain surface in a compact form, which is similar in structure to terrain profiles. Each terrain component $\sigma_{i,l}$, represents the contributions of the true surface in each principle direction. These terrain components are signals that are similar in form and simplicity to a terrain profile, except $\sigma_{i,l}$ represents the magnitude of each basis vector to the terrain surface. This concept is further developed in Chapter 5.

4.4 Chapter 4 Proof of Concept

4.4.1   Design of Experiment

An example using measured data is presented to demonstrate the effectiveness of using analytic basis vectors in capturing the characteristic shapes of the empirical basis vectors. Furthermore, the insensitivity to the discrete transverse locations at which the terrain heights are defined (either through measurement or spatial interpolation) is studied, showing the versatility of this approach in adapting to various profiler platforms. Four sections of U.S. Highway data (two concrete sections and two asphalt sections), each 151.1m long, were selected at the MnRoad testing facility in Albertville, Minnesota for this experiment. The sections were closed to traffic and each surface was measured ten times, in alternating directions, at a vehicle speed of approximately 10 m/s. The measurements were nominally corrected for body motion (Kern and Ferris 2007; Wagner, Kern et al. 2007) and mapped to a regularly spaced grid. The longitudinal spacing of 10mm resulted in 15110 longitudinal locations and a transverse spacing of 25.4mm at 51 transverse locations resulted in a 1.3 m wide terrain surface. The terrain heights at these uniform grid locations were captured in matrix form and a principle component analysis using an SVD algorithm was performed to identify the empirical basis vectors (refer to Figure 37). The empirical basis vectors used in this example represent a small portion of U.S. Highway data and, although they are not comprehensive, they provide a very

encouraging degree of support for representing terrain surfaces with analytic basis vectors.

4.4.2   Application of Theory

Before the analytic and empirical basis vectors are analyzed, the vector of transverse locations for the empirical data, v, is shifted horizontally 1.1375m to be centered on zero and scaled by a factor of 1.5 to exist on the same domain as the analytic basis vectors, [-1, 1].   The transverse locations defining v for the generation of the analytical basis vectors were evenly spaced among 51 points spanning [-1 ,1].   The magnitudes of the empirical basis vectors, b_l, were scaled by a factor of four to aid in visualizing the correlation. Clearly the correlation between the analytic basis vectors and the empirical basis vectors depends on the shifting and scaling of the measurement locations, v, but not on the scaling of the magnitude of the empirical basis vectors. The end of this Section will compare the root mean square error (RMSE) of each proposed analytic basis vector with the corresponding empirical basis vectors to determine which set of polynomials best represent this set of U.S. Highway data.

Gegenbauer Polynomials vs. Empirical Basis Vectors

For the Gegenbauer polynomials, the value of λ defines the relative scale of the tails of the polynomial; a value of $\lambda=0.1$ results in flat functions and a value of $\lambda=1$ results in polynomials with extreme tails. When $\lambda=0.5$, the Legendre polynomials are generated with the range bounded by [-1, 1]. The RMSE between the empirical and the analytic basis vectors corresponding to the Gegenbauer polynomials results when λ lies in the interval [0.45, 0.55], suggesting that the Legendre polynomials are most appropriate form of the Gegenbauer polynomials to compare to the empirical basis vectors. The Legendre polynomials will also be used to compare to the other two sets of polynomials to identify the most appropriate set of analytic polynomials to approximate the empirical data. This comparison will be discussed at the end of this Section. A comparison of the first four empirical basis vectors with the Legendre polynomials is shown in Figure 41. The first empirical basis vector and the first analytic basis vector provide a means to describe the overall change in elevation, however there is an obvious

offset between the empirical and analytic basis vectors; this is simply due to the scaling of the magnitude of the empirical basis vectors and, again, has no impact of the correlation. The second empirical basis vector is collinear with the second analytic basis vector, defining the bank angle in the road. The analytic and empirical basis vectors for the bank angle terrain characteristic correlate well. There was negligible crowning in the empirical data (corresponding to the third analytic basis vector), so no comparison is possible and no figure is provided. The third empirical basis vector representing the asymmetry terrain characteristic is compared with the fourth analytic basis vector (corresponding to the third order polynomial). The empirical data shows an asymmetry in the shape of the basis vector in that the local maximum is much narrower than the local minimum. The fourth empirical basis vector representing the road rutting terrain characteristic correlates well with the fifth analytic basis vector, however the empirical vectors exhibit deeper local minimums than the analytic basis vector.
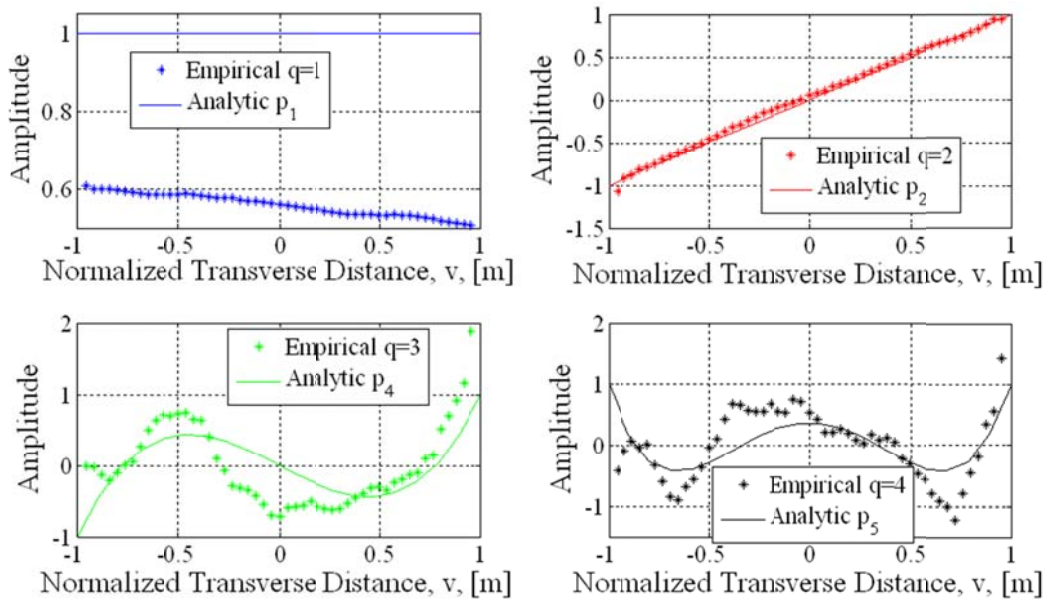
Figure 41. Empirical basis vectors compared with the Legendre polynomials.

Chebyshev Polynomials vs. Empirical Basis Vectors

A comparison of the first four empirical basis vectors with the Chebyshev polynomials is shown in Figure 42. The first empirical basis vector and the first analytic basis vector provide a means to describe the overall change in elevation, however there is

an obvious offset between the empirical and analytic basis vectors. The second empirical basis vector is collinear with the second analytic basis vector, defining the bank angle in the road. It is important to note that the analytic basis vectors pertaining to the first two Legendre and Chebyshev polynomials are identical. The main differences in appropriate curve fitting will become apparent in the latter basis vectors. There was negligible crowning in the empirical data (corresponding to the third analytic basis vector), so no comparison is possible and no figure is provided. The third empirical basis vector representing the asymmetry terrain characteristic is compared with the fourth analytic basis vector (corresponding to the third order polynomial). The analytic basis vector does not appear to correlate well with the empirical basis vector due to the property of the Chebyshev polynomials that the local minima and maxima are equal to +1 or -1. The fourth empirical basis vector representing the road rutting terrain characteristic correlates well with the fifth analytic basis vector.
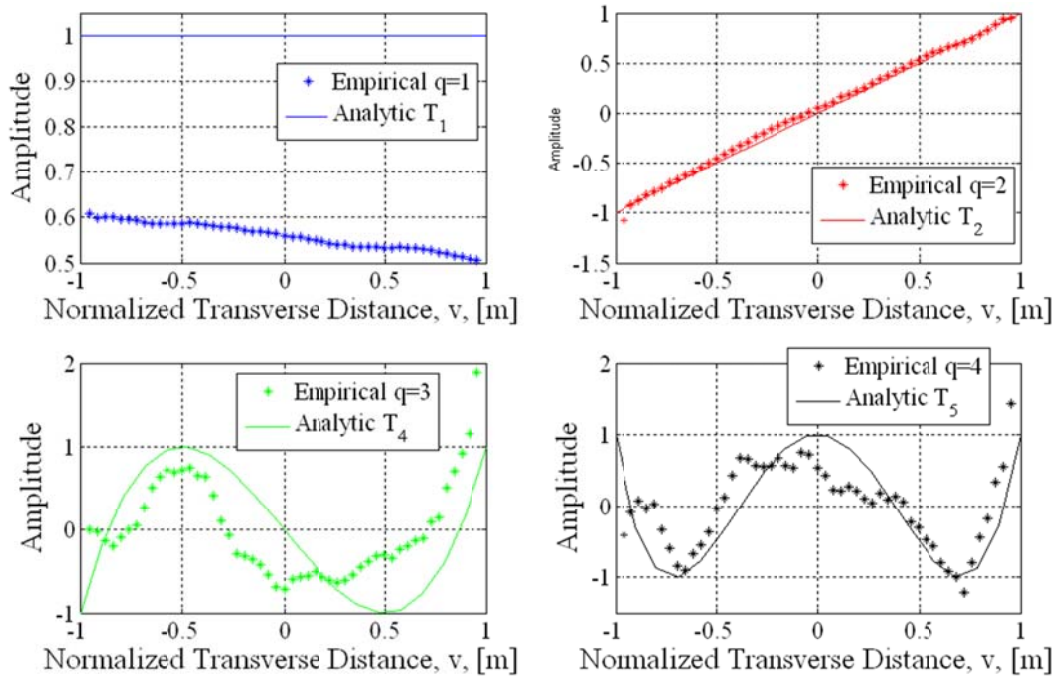
Figure 42. Empirical basis vectors compared with the Chebyshev polynomials.

Fourier Basis Functions vs. Empirical Basis Vectors

A comparison of the first four empirical basis vectors with the Fourier series basis functions is shown in Figure 43. The first empirical basis vector and the first analytic

116

basis vector provide a means to describe the overall change in elevation; again there is an obvious offset between the empirical and analytic basis vectors. It is important to note that the first analytic basis vectors are identical for the three sets of polynomials used in this analysis. More importantly, the second empirical basis vector is not represented in the basis functions of the Fourier series and thus does not show any correlation with the second analytic basis vector. The Fourier series could be summed to approximate the sloped line, but this negates the intention of this process and would require an infinite sum to obtain this. The fact that the Fourier basis functions do not capture the bank angle basis vector, and that this basis vector pertains to one of the two dominant features in the terrain surface indicates that the Fourier series are not appropriate to model the empirical data. Other differences will be seen in the third and fourth basis vectors. The third empirical basis vector representing the asymmetry terrain characteristic is compared with the third analytic basis vector. Similar to the third order Chebyshev polynomial, the local minimum and maximum are greater than that of the empirical data suggesting that the third order Fourier basis vector is not a good fit. The fourth empirical basis vector representing the road rutting terrain characteristic does not appear to correlate well with the fifth analytic basis vector. The local minimums do not line up with the minimums of the empirical data and the local maximum is too narrow as compared with the empirical data.
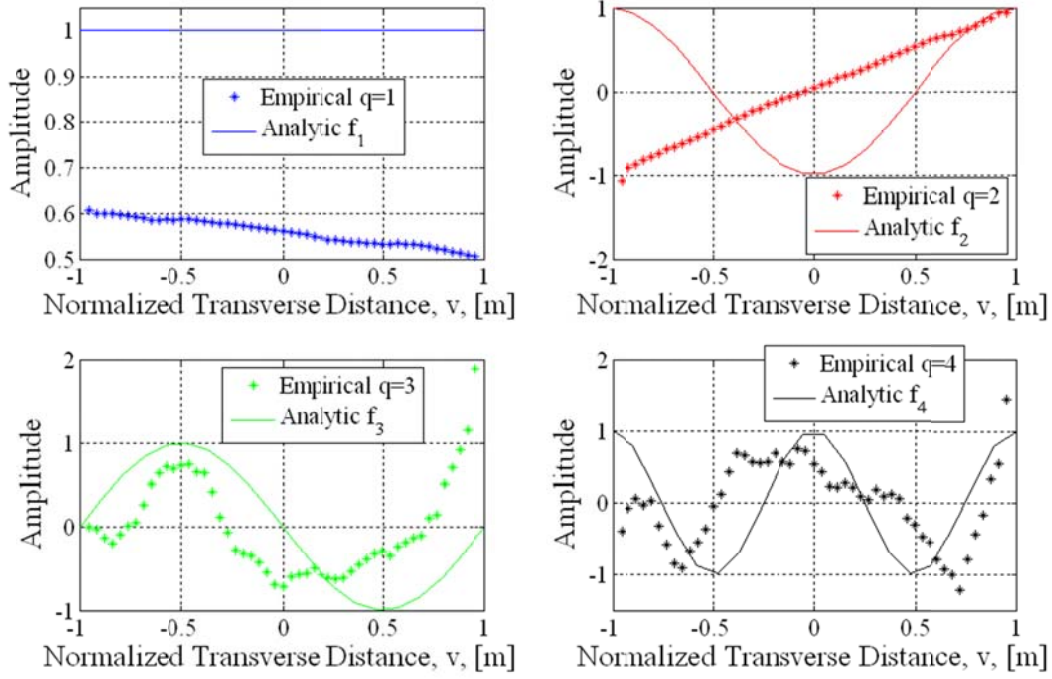
Figure 43. Empirical basis vectors compared with the Fourier series basis functions.

RMSE Study to Determine 'Best Fit'

The 'best fit' between the empirical basis vectors and the proposed sets of analytic polynomials is identified by the set of polynomials that presents the minimal RMSE between the first four empirical basis vectors and the corresponding analytic basis vectors (refer to Figure 41, Figure 42, and Figure 43 for a visual comparison). Plots comparing the RMSE of each polynomial set to the empirical basis vectors are described and shown next.

The RMSE between the Legendre, Chebyshev and Fourier polynomials and the first empirical basis vector was calculated and plotted in Figure 44. Referring to the previous section, it is important to note that each of these polynomial sets produced the same analytic basis vector, thus there is no differentiation between the polynomials. The trend in the RMSE plot is due to the fact that the empirical basis vector corresponding to the terrain characteristic of elevation is not perfectly level and exhibits a small downward slope. The error is then shown to increase when moving from the left to the right, with an average RMSE of 9.8%.
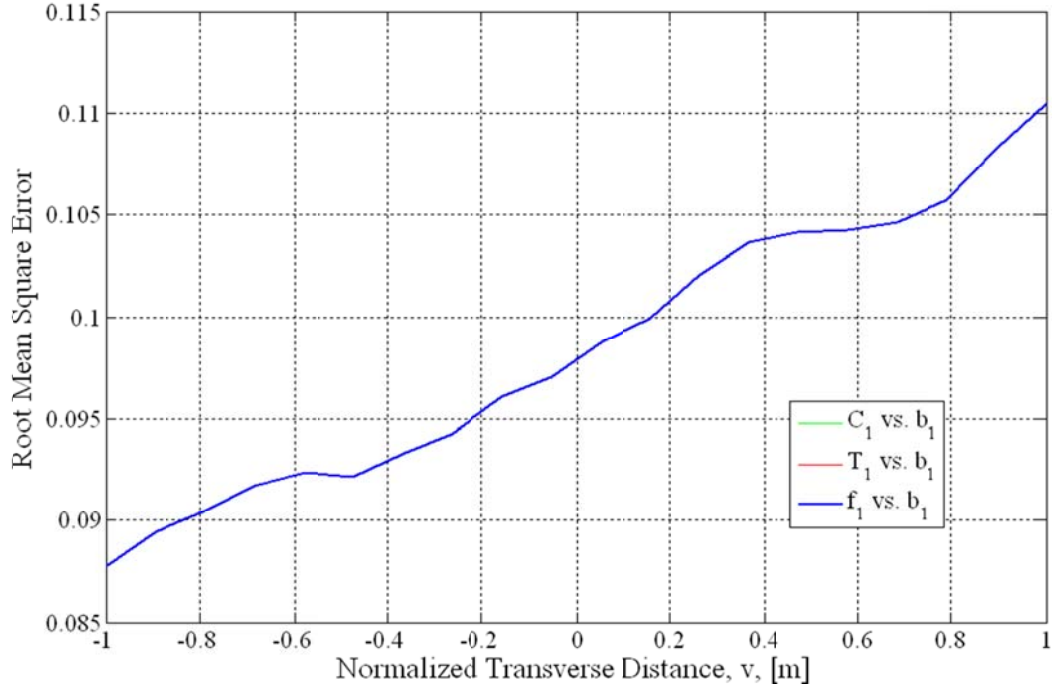
Figure 44. RMSE of analytic polynomials vs. empirical basis vectors for b_1.

The RMSE between the Legendre, Chebyshev and Fourier polynomials and the second empirical basis vector was calculated and plotted in Figure 45. Referring to the previous section, it is important to note that the Legendre and Chebyshev polynomials produced the same analytic basis vector, thus there is no difference between these two polynomials. The Fourier basis functions do not include a sloped line as one of its shapes and results in a large error when compared with the second empirical basis vector. The average RMSE of the Legendre and Chebyshev was calculated to be 0.83%, indicating that either polynomial is an appropriate fit for this basis vector.
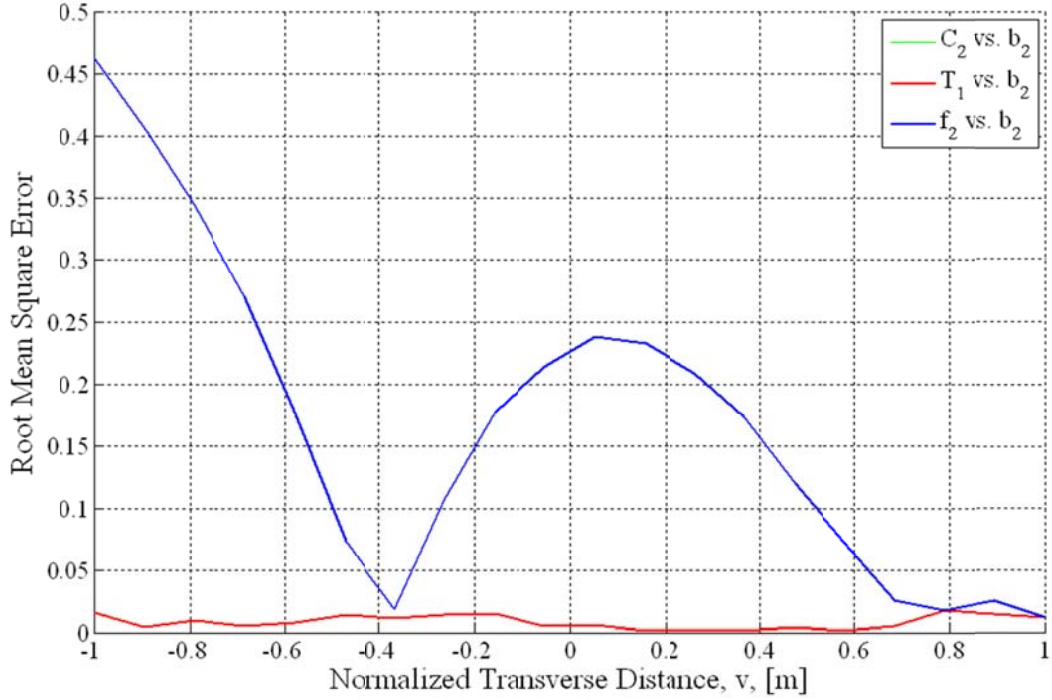
Figure 45. RMSE of analytic polynomials vs. empirical basis vectors for b_2.

The RMSE between the Legendre, Chebyshev and Fourier polynomials and the third empirical basis vector was calculated and plotted in Figure 46. All of the polynomials produced a significant amount of error, as high as 60%. Despite this, the Legendre basis vector produced the least amount of error of the three polynomials, with an average RMSE of 12.8%. The greatest amount of error occurred on the tails and when trying to match the asymmetric maximum of the empirical basis vector.
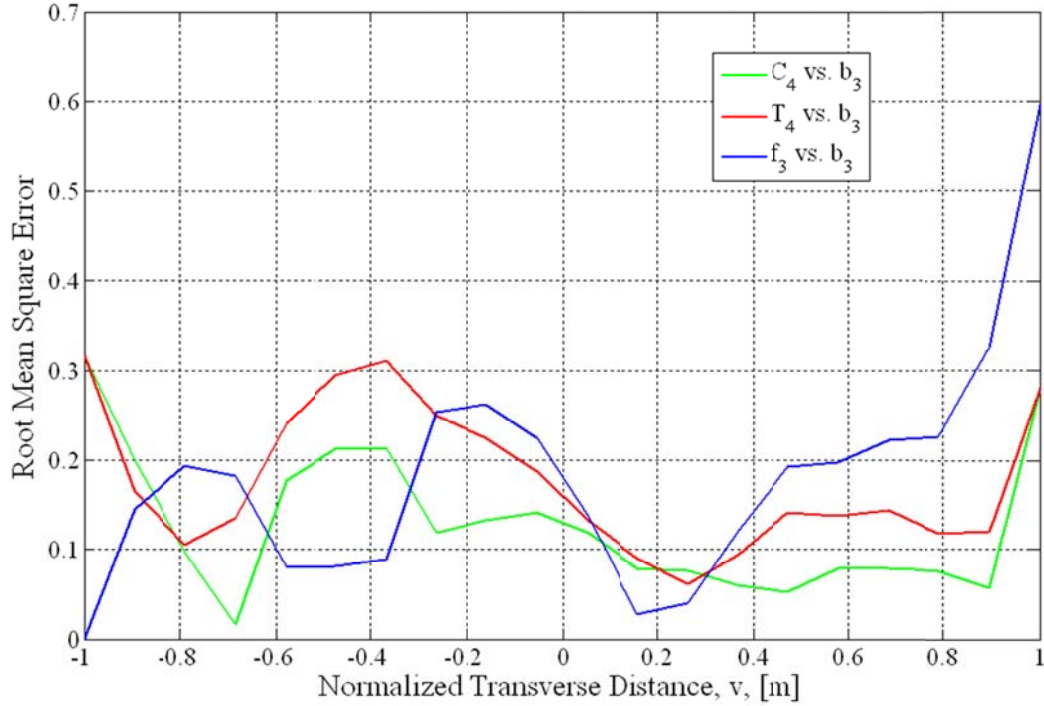
Figure 46. RMSE of analytic polynomials vs. empirical basis vectors for b_3.

The RMSE between the Legendre, Chebyshev and Fourier polynomials and the fourth empirical basis vector was calculated and plotted in Figure 47. The average RMSE of all of three polynomials were compared to identify which basis vector produced the minimum RMSE. The fifth Legendre basis vector produced an average RMSE of 8.1%. The fifth Chebyshev basis vector produced an average RMSE of 8.5%. The fourth Fourier basis vector produced an average RMSE of 15%. While the Legendre basis vector produced the lowest average RMSE, it also yielded the second highest peak value of error.
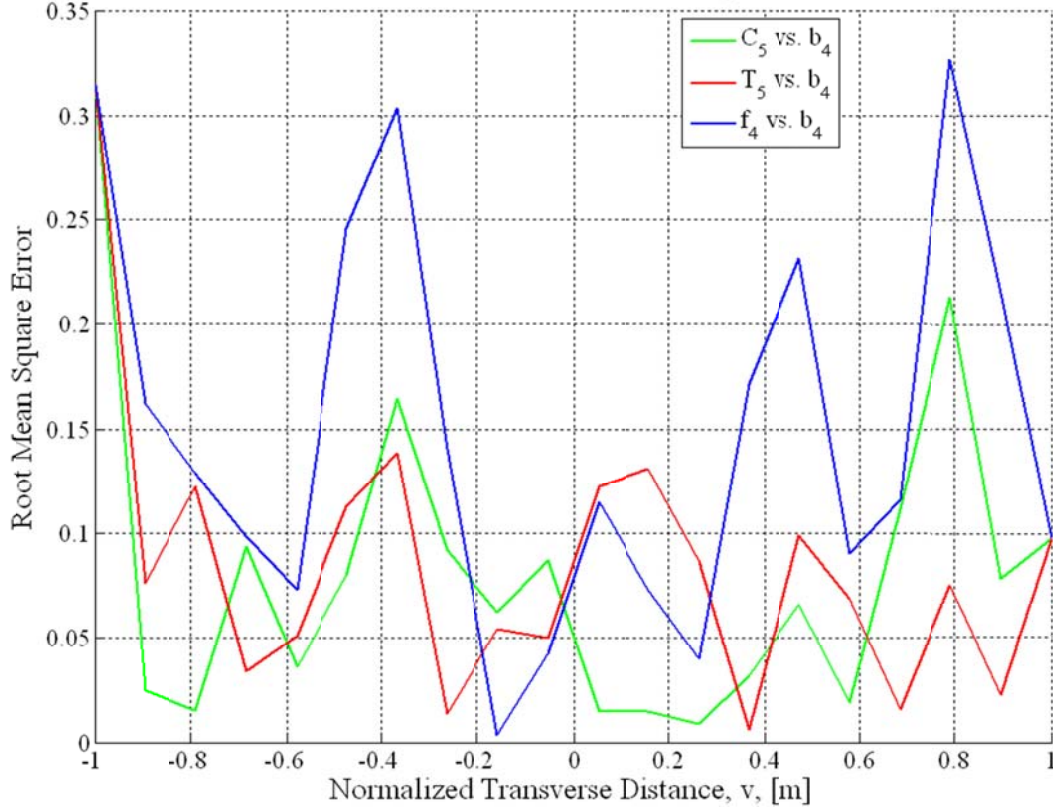
Figure 47. RMSE of analytic polynomials vs. empirical basis vectors for b_4.

The Fourier polynomials produced errors much greater than the Legendre and Chebyshev polynomials, as shown in Table 1. By calculating the average RMSE of each analytic basis vector against each empirical basis, it is concluded that the Legendre polynomials best define the analytic basis vectors for U.S Highways. It is expected that additional datasets will provide some spatial averaging to the empirical basis vectors, thus smoothing the empirical data and further supporting the correlation with the Legendre basis vectors.

Table 1. Average RMSE values of each analytic basis vector compared with each empirical basis vector.

	Basis 1	Basis 2	Basis 3	Basis 4
Legendre	9.8%	0.83%	12.9%	8.1%
Chebyshev	9.8%	0.83%	17.7%	8.5%
Fourier	9.8%	16.9%	17.8%	15%

4.4.3 Portability of Method to other Terrain Measurement Systems

Next, the insensitivity in the estimation of the components of terrain,   , with respect to the transverse grid locations is demonstrated. As previously demonstrated, the Legendre polynomials provide the best fit for U.S. Highway data and will be used as the representative set of basis vectors for the remainder of this work. Consider a simple terrain measurement system configuration with three optical sensors distributed along the width of the vehicle such that they measure the center of each wheel path and the center of the vehicle, with resulting analytic basis vectors shown in Figure 48A. This is compared to a system producing data at ten transverse sample locations with resulting analytic basis vectors shown in Figure 48B. Clearly a greater transverse sampling density (defined by v) provides more information about the surface being measured. That is, information about higher order contributions is unavailable from a system with few transverse measurement locations. To the extent possible, however, the method developed in this Chapter provides a means by which a consistent description of the surface can be generated by equipment with different sampling densities.
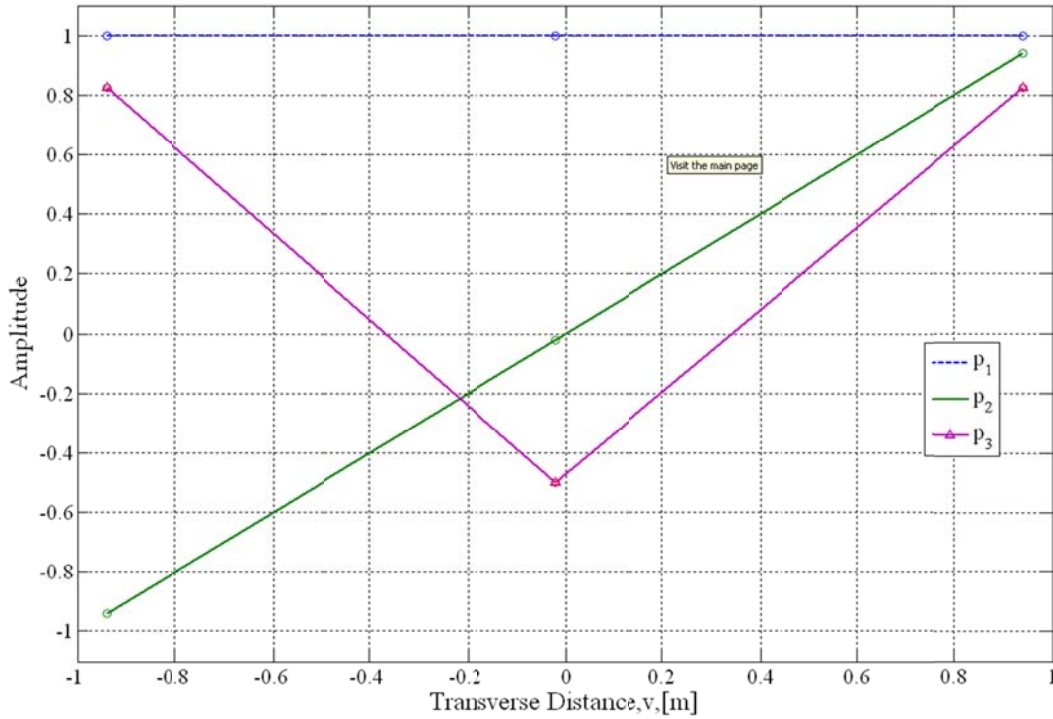
A)
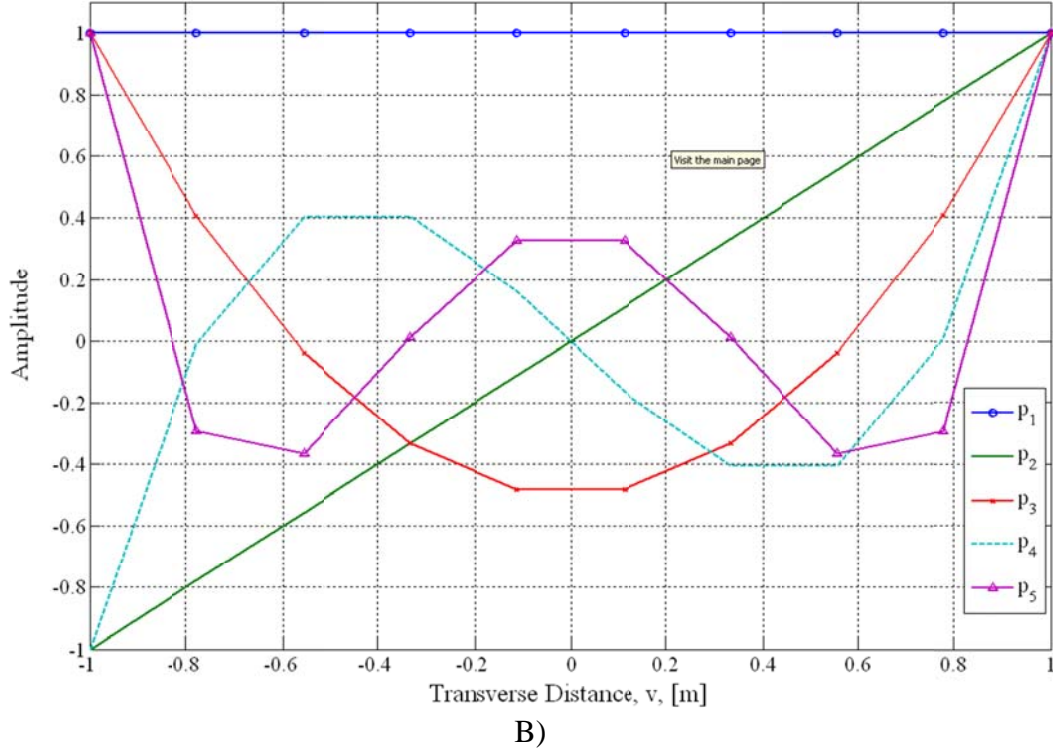
Figure 48. Legendre basis vectors for A) three samples comprising v and B) ten samples comprising v.

The components of terrain, $\sigma_{i,l}$, were studied for the first two principle directions for a system with three transverse locations (so that the length of v is three) and for a system with ten transverse locations. The first two components of terrain, $\sigma_{i,1}$ and $\sigma_{i,2}$, are shown in Figure 49 and Figure 50 respectively, along with the corresponding percent difference between the two components. The percent difference in the two signals represents the relative capability of a system with discrete sampling locations at three locations compared to a system sampling at ten locations to accurately capture the first two terrain surface characteristics. The first component of the terrain, $\sigma_{i,1}$, multiplied by the corresponding first basis vector, p_1, captures the elevation of the terrain surface. The mean percent difference in the first component of terrain for the two sampling vectors is 1.8E-5%. The second component of the terrain, $\sigma_{i,2}$, captures the banking in the terrain surface. Note that in Figure 50 the mean percent difference in the second component of terrain for the two sampling vectors (three locations vs. ten) is 0.0255%. It should be clear that the more basis vectors used to span the vector will reduce the amount of shared

124

content between basis vectors. In this example, the differences in the components of terrain arise due to the amount of information that is contained in the three basis vectors versus ten basis vectors. This estimation error is contributed by the definition of the vector space. When only three points are used to define the terrain surface, there is less confidence (more variation) in the estimate of the components of terrain as there would be with more transverse samples.
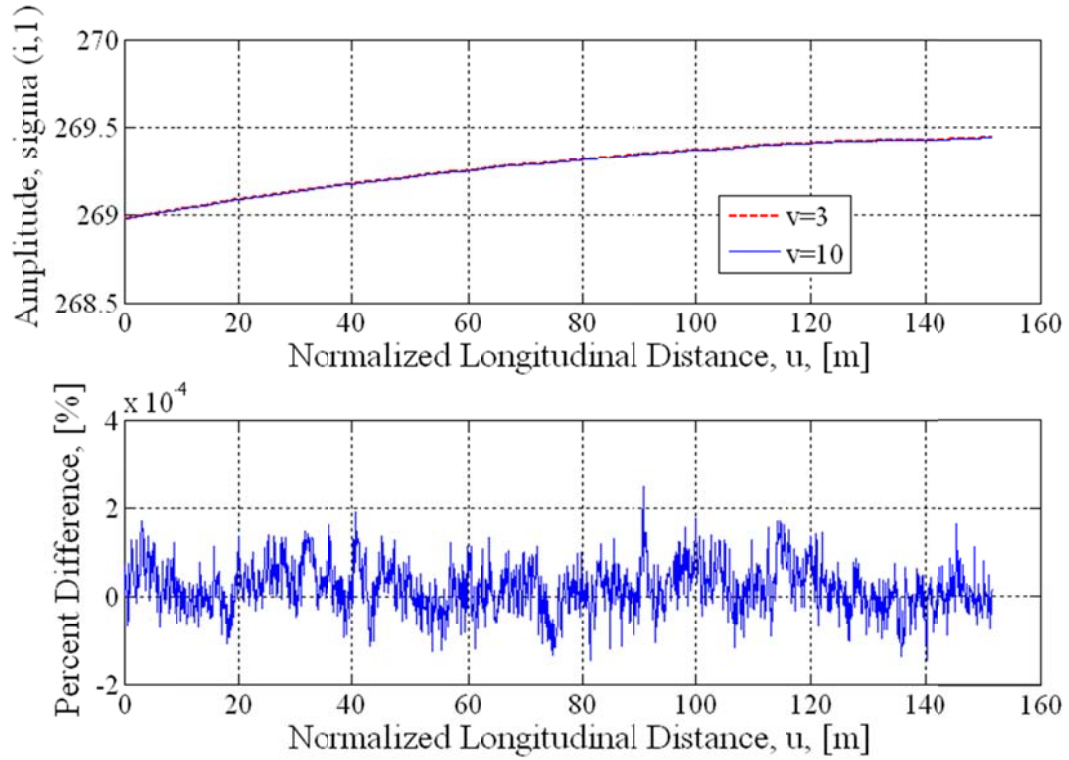
Figure 49. Comparing $\sigma_{i,1}$ for for two measurement systems.
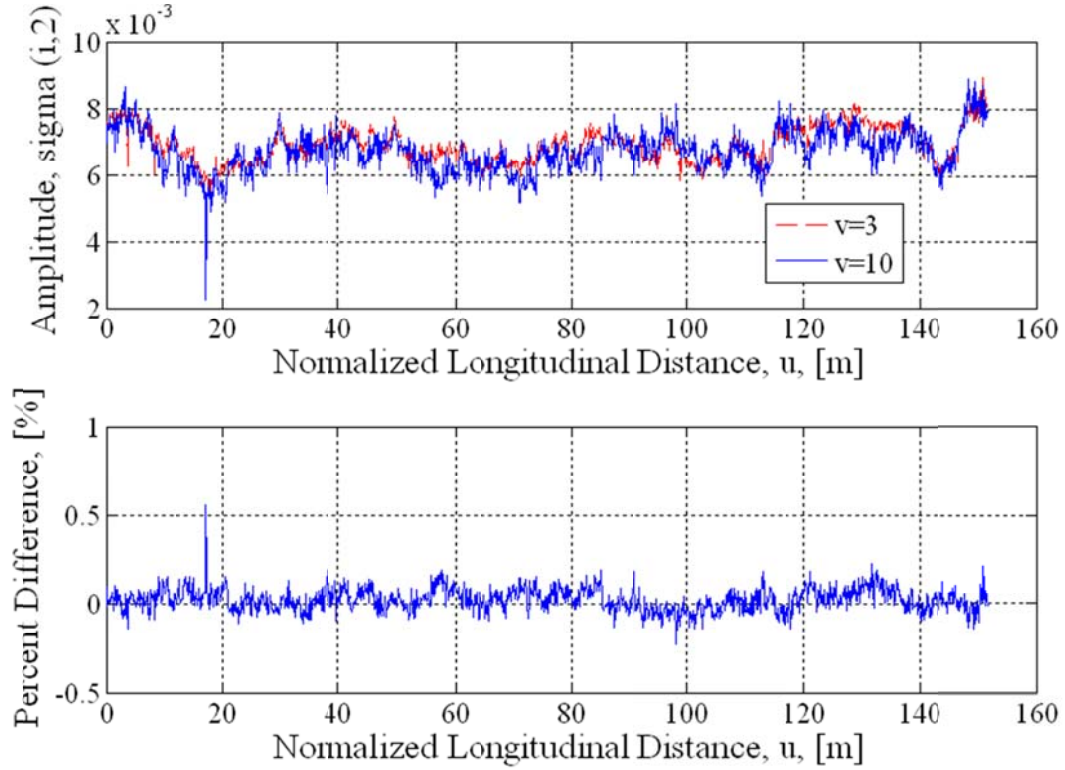
Figure 50. Comparing $\sigma_{i,2}$ for two measurement systems.

While this evidence cannot be regarded as conclusive without more extensive investigation of highway data, there is a strong indication that U.S. Highway data can be decomposed into projections onto analytic basis vectors that adequately capture the shapes of the empirical data. In this way, the method is portable in that the same set of analytic basis vectors can be used on all U.S. Highways. The portability of this method also extends to the ability to use measurements systems with different equipment configurations, as demonstrated by an example.

4.5 Chapter 4 Discussion

More empirical data representing various classes of terrain need to be studied to understand the limitations of this method. Terrain should be studied based on location, material composition, and paved vs. unpaved conditions and separated into appropriate types based on these characteristics. The appropriate selection of the polynomial sets would then be identified for each terrain type using the method developed in this Chapter.

126

It should be noted that the correlation of the analytic basis vectors to the empirical basis vectors is highly dependent on the scaling and shifting of the transverse locations, v, and this must be examined when determining the appropriate parameterization.

Expanding the database to include various terrain types may also suggest using a different set of generating functions to approximate the empirical basis vectors. Such techniques that may prove to be promising include, but are not limited to, the polynomials previous described here, Laguerre polynomials, non-uniform rational basis splines (NURBS), and Maclaurin polynomials (Taylor series polynomials when $a_0=0$). Once a more comprehensive set of data have been collected, each set of generating functions can be investigated using the method developed in this Chapter. For example, different sets of generating functions may be more appropriate for different types of terrain. Again, the appropriateness of each generating function would be determined by minimizing the RMSE between the resulting analytic basis vectors and the corresponding empirical basis vectors.

Although curve fitting the empirical data directly may appear to be the most efficient representation of the terrain, it does not exploit the specific anisotropicity of this path-specific method. Toward this end, a set of well-known generating polynomials describe a specific type of the terrain to maintain the portability of the method developed in this chapter. In this way, this method could be applied to any terrestrial terrain measurement system to formulate an accurate description of the terrain surface that contains path information. Additionally, these well-known sets of polynomials are commonly used in real-time data analysis. In this case, the empirical basis vectors are replaced by a recurrence relationship, thus eliminating the need to retain the information describing the empirical basis vectors and reducing the total information needed to describe the terrain surface. The next Chapter will further develop a method to represent terrestrial terrain measurement data in a more compact form.

4.5.1 Segmenting the Terrain

To reinforce the acceptability and appropriateness of the empirical basis vectors, the terrain sample was segmented into ten sections of approximately 1500 points in length. A SVD was performed on each section and the empirical basis vectors were

compared for all segments. This was completed in order to justify taking the SVD of the entire surface rather than using more of a FEA approach, where shape functions (basis vectors) are calculated for each grid node.

The first basis vector, corresponding to the principle direction of elevation, for the ten sections is shown in Figure 51A. It can be seen that each basis vector follows a similar trend with a maximum variation of approximately 0.001. Figure 51B identifies the maximum and minimum bounds of the first basis vector, as plotted in red dashed lines, by taking the maximum and minimum values across all ten sections. The mean of the ten sections is plotted as a dashed blue line. The elevation basis vector corresponding to the SVD of the entire surface is plotted in black. It can be seen that this basis vector falls within the calculated bounds. .

A)

B)
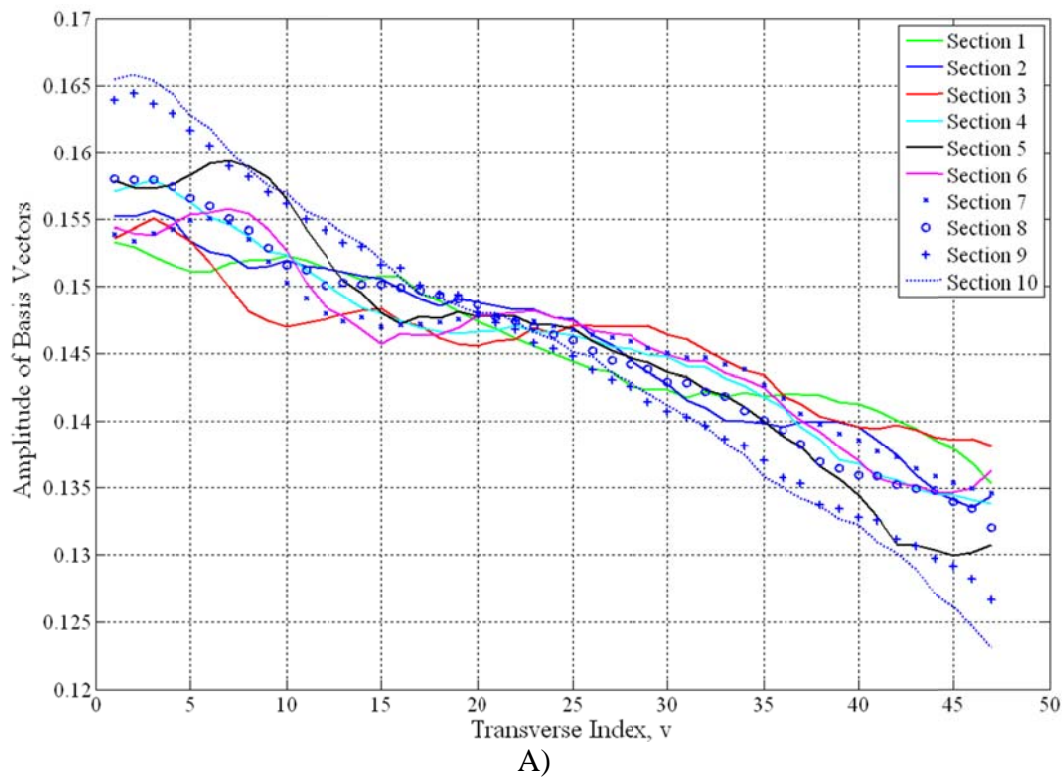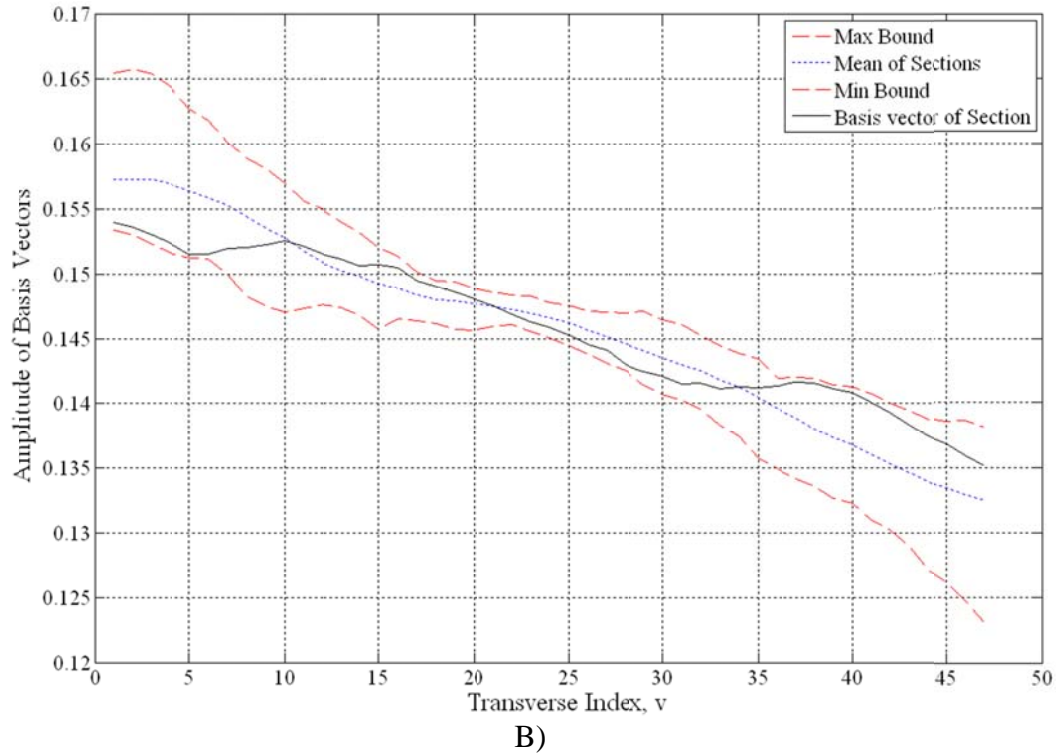
Figure 51. Identifying basis vector variation along the terrain section A) Comparing all ten segments B) Comparing the minimum bound, maximum bound, mean basis vector, and the basis vector of the entire surface.

The second basis vector, corresponding to the principle direction of bank angle, for the ten sections is shown in Figure 52A. It can be seen that each basis vector follows a similar trend with a maximum variation of approximately 0.0005. Figure 52B identifies the maximum and minimum bounds of the second basis vector, as plotted in red dashed lines, by taking the maximum and minimum values across all ten sections. The mean of the ten sections is plotted as a dashed blue line. The bank angle basis vector corresponding to the SVD of the entire surface is plotted in black. It can be seen that this basis vector falls within the calculated bounds.
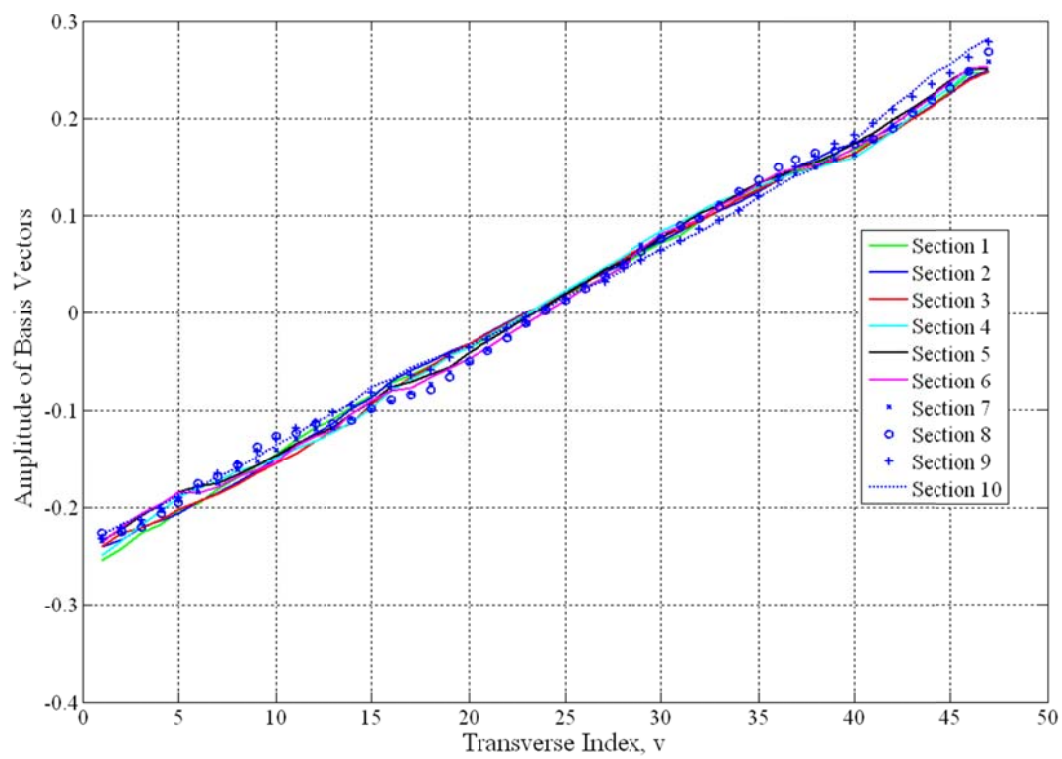
A)
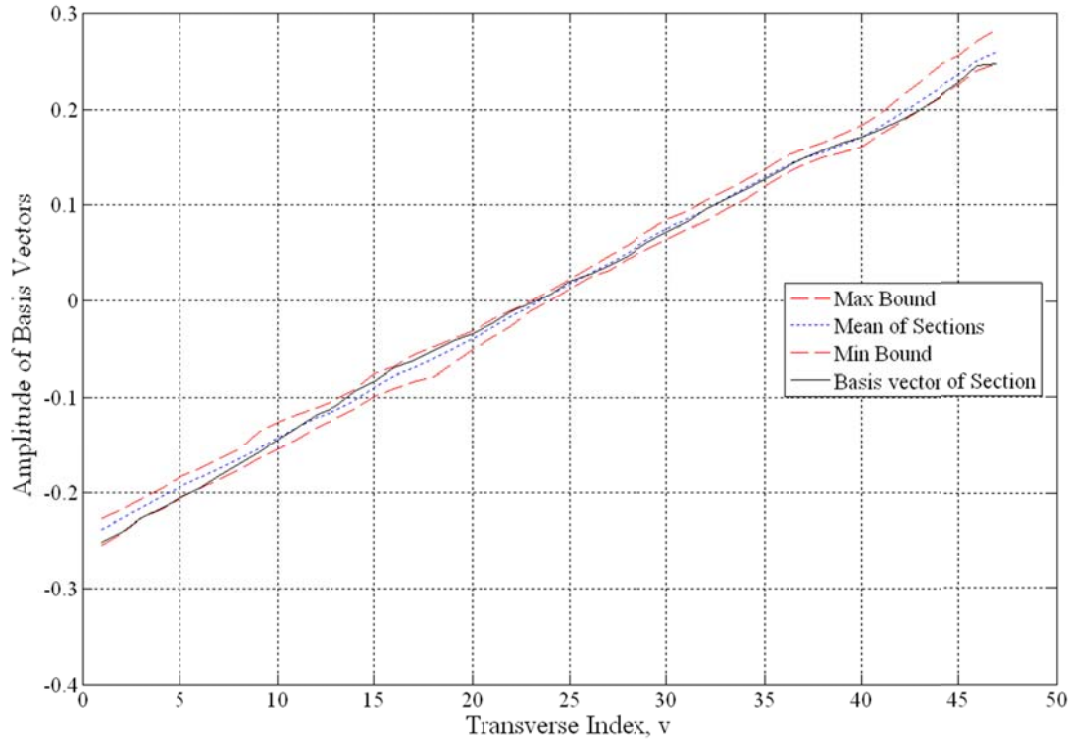
B)

Figure 52. Identifying basis vector variation along the terrain section A) Comparing all ten segments B) Comparing the minimum bound, maximum bound, mean basis vector, and the basis vector of the entire surface.

The third basis vector, corresponding to the principle direction of asymmetry, for the ten sections is shown in Figure 53A. It can be seen that the basis vectors follow a similar trend, but show a lot of variation in peak and valley location. This reduced correlation between basis vectors from different segments along the length of the road is partially explained by the fact that the third singular value is negligible compared to the first two. Additionally, it was shown in Chapter 3 that in order to obtain a zero-mean noise process, only two basis vectors need to be implemented for these examples of U.S. Highway data. Figure 53B identifies the maximum and minimum bounds of the third basis vector, plotted in red dashed lines, by taking the maximum and minimum values across all ten sections. The mean of the ten sections is plotted as a dashed blue line. The asymmetry basis vector corresponding to the SVD of the entire surface is plotted in black. It can be seen that this basis vector falls within the calculated bounds, however the bounds span from -0.3 to 0.2.

131
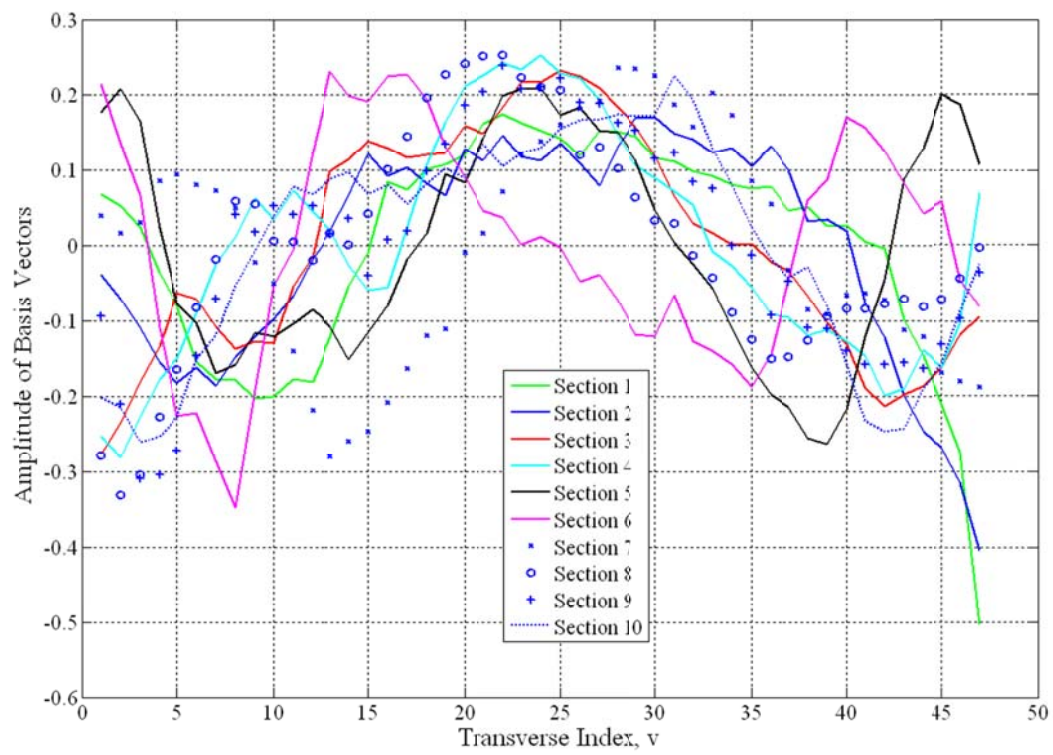
A)
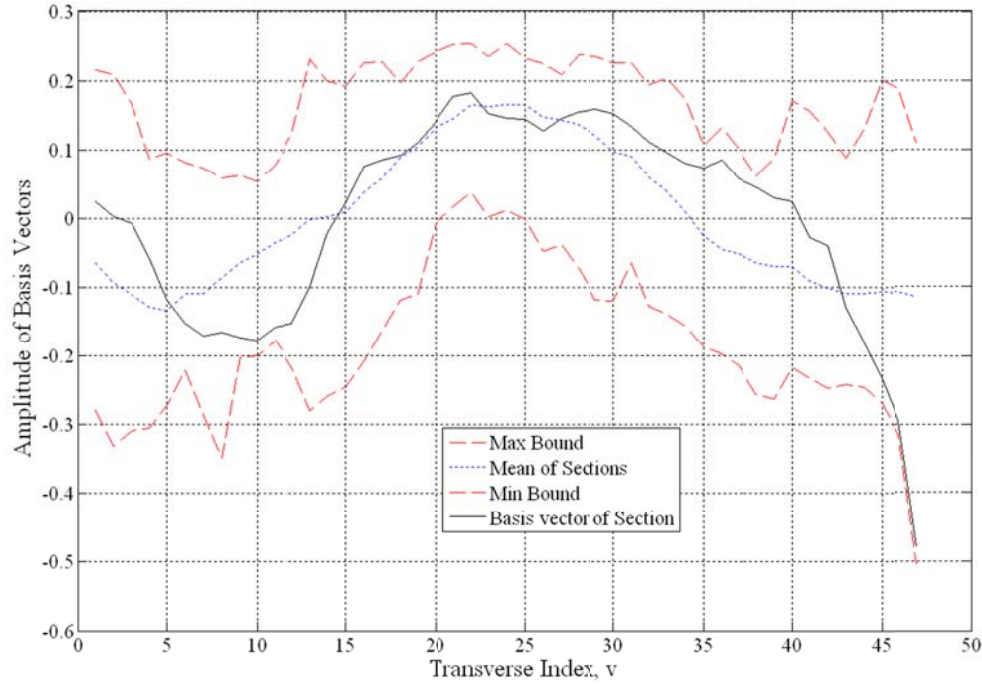
B)

Figure 53. Identifying basis vector variation along the terrain section A) Comparing all ten segments B) Comparing the minimum bound, maximum bound, mean basis vector, and the basis vector of the entire surface.

The fourth basis vector, corresponding to the principle direction of road rutting for the ten sections is shown in Figure 54A. It can be seen that the basis vectors do not appear to be correlated. Similar to the third basis vector, this reduced correlation between basis vectors from different segments along the length of the road is explained by the fact that the singular value is negligible. Again, it was shown in Chapter 3 that in order to obtain a zero-mean noise process for the proof of concept using U.S. Highway data, only two basis vectors need to be implemented. The lack of correlation between the basis vectors will deteriorate as the order of the basis vectors continues to increase due to the corresponding singular values approaching zero. Figure 54B identifies the maximum and minimum bounds of the fourth basis vector, as plotted in red dashed lines, by taking the maximum and minimum values across all ten sections. The mean of the ten sections is plotted as a dashed blue line. The road rutting basis vector corresponding to the SVD of the entire surface is plotted in black. It can be seen that this basis vector exceeds the calculated bounds at several locations, and the bounds span from -0.2 to 0.3.

133

A)

134

B)
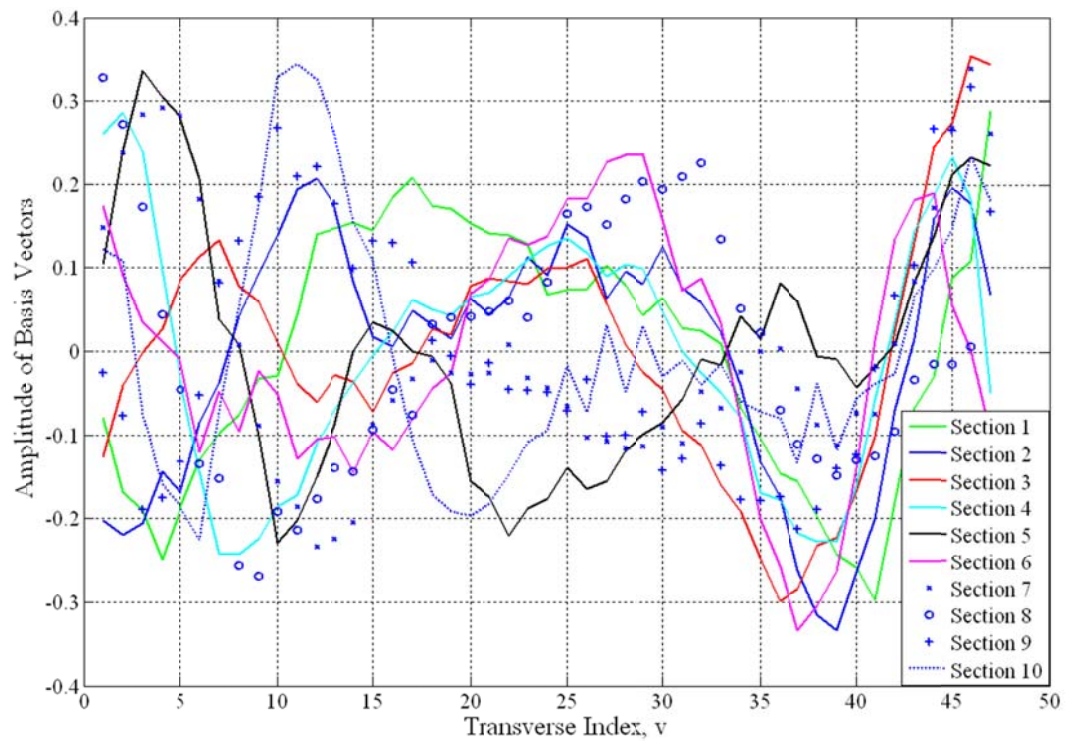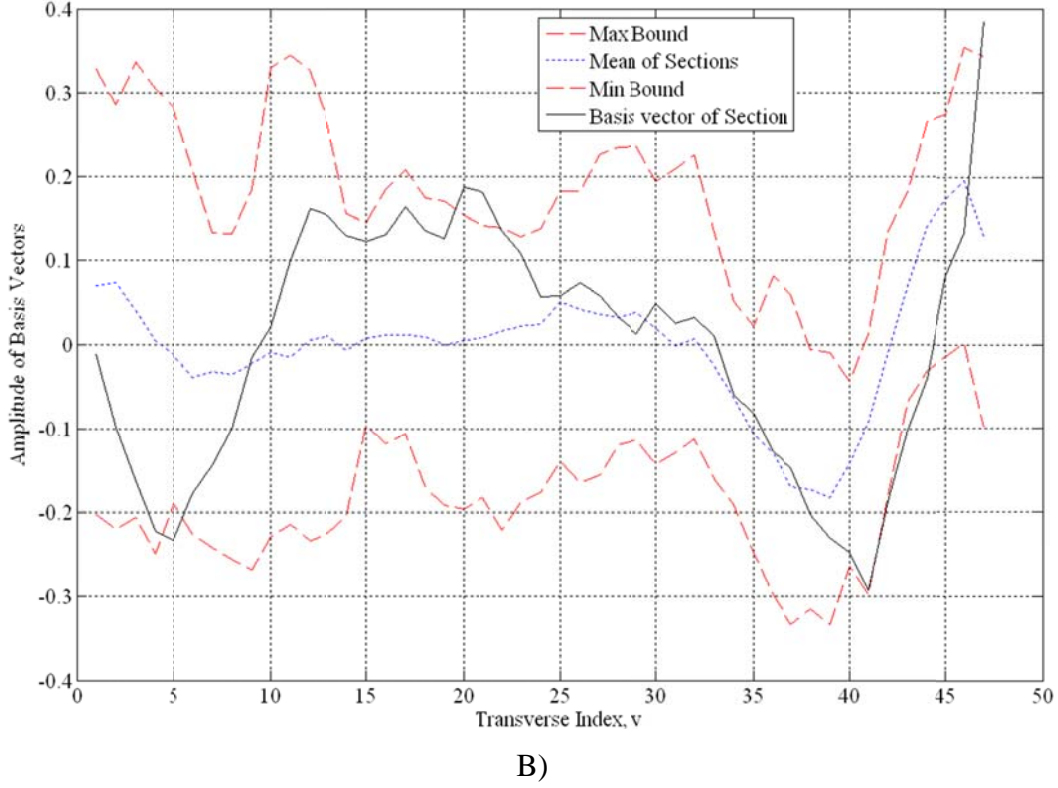
Figure 54. Identifying basis vector variation along the terrain section A) Comparing all ten segments B) Comparing the minimum bound, maximum bound, mean basis vector, and the basis vector of the entire surface.

4.5.2   Scaling and Shifting the Transverse Locations of the Analytic Basis Vectors

Another consideration for future work would be to scale and shift the transverse locations so that the transverse locations can be defined based on standard road and vehicle characteristics. For example, consider aligning the two local minima of the fifth analytic basis function (C_5 in Figure 38) with the track width of typical passenger vehicles, defined as t. The distance between the two local minima, defined as $\rho(\lambda)$, is a function of the parameterization, λ. The vector $\mathbf{v}$ would be normalized according to Equation 75, where $\mathbf{v}_{measured}$ are the original transverse locations and $\overline{\mathbf{v}_{measured}}$ is the mean of the original data. In this way, the transverse locations are normalized to correspond with the 'rutting' condition that arises from road wear due to tire interaction. The ability to quantitatively identify the presence of physical characteristics such as rutting in any pavement may lead to new insight into the mechanisms leading to pavement degradation.

135

$$v = \frac{v_{measured} - \overline{v_{measured}}}{\dfrac{t}{2}} \rho(\lambda)$$

<div align="right">Equation 75</div>

Once more data have been collected from a variety of terrain types, and they have been properly decomposed into their principle components, $\sigma_{i,l}$, investigations can be made into the stochastic or deterministic nature of the principle components. If components of the terrain surface, $\sigma_{i,l}$, can be considered realizations of an underlying stochastic process, then a stochastic model of these components can be developed. Terrain surfaces could then be represented by a combination of deterministic and stochastic processes. If all of the surface components are included and classified as deterministic, then the synthesized terrain surface would reproduce the originally measured terrain surface. As more components are classified as stochastic, more variation in the realizations will arise. This variation can be useful in chassis loading predictions where vehicle simulations conducted over several realizations of the same terrain type would identify more realistic maximum and minimum load cases. Reconstructing the terrain surface with these models would yield a compact, stochastic terrain surface representation. Since realizations of any length could be generated, vehicle responses could be simulated that are accurate for the given application, yet found quickly enough to make informed design decisions early in the design process. Synthesizing terrain surfaces in this manner will help shorten the development time to bring a new vehicle to market and provide a better fundamental understanding of how the vehicle is excited. This concept is further developed in Chapter 5 of this dissertation.

4.6 Chapter 4 Conclusions

Terrain surfaces capture a great deal more information about the terrain topology than simple terrain profiles. A new representation has been developed that captures the principle information about a terrain surface in a compact form, similar in structure to terrain profiles. Specifically, a Galerkin method was developed to define terrain surfaces as sequences of vectors in a Hilbert space composed of analytic basis vectors with a weighted inner-product and a weighted grammian. The analytic basis vectors that best fit the U.S. Highway example are generated from Gegenbauer polynomials that are parameterized to closely match the empirical basis vectors. Within the scope of U.S.

Highways, the proper parameterization of the generating polynomial results in a special case of the Gegenbauer polynomials: the Legendre polynomials. The resulting analytic basis vectors successfully capture the characteristics of the first four empirical basis vectors. A weighted inner-product was developed such that the resulting representation of the terrain surface is insensitive, within the limitations of the sampling density, to the measurement system used to acquire the terrain surface data. The basis vectors are then projected back onto the true surface to formulate components of terrain corresponding to each principle direction. These components of terrain describe the magnitude of each principle direction along the length of the road. Possible applications of this new method include modeling the changes in projections of the analytical basis vectors as a stochastic process and quantifying physical characteristics of the terrain such as elevation, bank angle, crowning, and road rutting.

5. Compact Models of Terrain Surfaces

5.1 Chapter 5 Introduction

Throughout a chassis development program it is necessary to possess load data representing severe customer usage to ensure that the chassis will perform as required. Unfortunately, actual loads are only available at the conclusion of the program. The design engineer is challenged with using predicted chassis loads early in the design process – when changes are relatively easy and inexpensive to make – and measured chassis loads late in the program – when changes to the design are extremely costly, if allowed to be implemented. It is clear that the terrain is the main excitation to the vehicle (Aurell and Edlund 1989). The non-deformable terrain imposes a unilateral geometric boundary constraint on rolling tires to which the chassis responds by generating loads, moments, motions, deformations, etc. The terrain surface remains a consistent excitation to the chassis, even as the chassis design changes. Knowledge of this excitation, when applied in conjunction with high fidelity tire and vehicle models, would allow chassis loads to be accurately predicted in vehicle simulations. Therefore, throughout the design process the system response to this consistent excitation can be calculated and compared for each chassis design considered. Accurate terrain models would then provide the chassis designer with a powerful tool to make informed design decisions early in the design process while changes are relatively inexpensive to implement. This will, in turn, shorten vehicle development time and reduce overall development costs.

High-fidelity terrain surfaces can be used as excitations to vehicle models in a virtual environment to accurately emulate the operation of a real vehicle. For example, specific events on a vehicle manufacturer's proving ground may be simulated and compared to measured data for model validation (Chaika, Gorsich et al. 2004). Current terrain measurement systems acquire approximately one million data points per second (Brandenburg and Rudd 1974; Karamihas 2005; Kern and Ferris 2007; Wagner, Kern et al. 2007). This data acquisition rate improves the available signal bandwidth and allows sharp disturbances to be detected in both the transverse and longitudinal directions. Capturing these disturbances is critical; the driver's perception of ride quality is largely

dictated by these events (Ferris 1999; Ahlin, Granlund et al.) and chassis durability is highly sensitive to transient loading (Lu and Lee 1996; Liu and Herman 1999; Stadterman, Connon et al. 2003; Bogsjo and Forsan 2004; Lei and Yang 2006; Duni, Toniato et al. 2008). Most terrain measurement systems are equipped with an Inertial Navigation System (INS) and accelerometers to remove body motion from the measured surface. INS systems are plagued with an inherent system drift, easily identified by taking multiple measurements of the same segment of road. Previous work by the authors produced a method, as presented in Chapter 3 of this dissertation, to compensate for the INS drift error of the terrain measurement system (Chemistruck, Binns et al. 2010).

The objective of this work is to develop a stochastic terrain surface that leverages the terrain representation implemented in the INS drift removal technique and previous work on Autoregressive (AR) terrain models. The remainder of this work is developed as follows. Background on terrain measurement is provided to further clarify the concept of terrain surfaces as a sequence of vectors and the compensation technique for INS drift is reviewed. A brief description of Autoregressive terrain models is also provided. The contribution of this work is then developed. The notion of global and local subspaces is generalized to a continuous scale in which the terms 'global' and local' are subjective locations along a well-defined continuum of possible scales. Specifically, the true surface is decomposed into a sequence of vectors; each vector represents the transverse heights of the terrain along a line perpendicular to the direction of vehicle travel. A Hilbert space is defined from this collection of vectors and principle directions are determined. These principle directions are cast as a set of orthonormal basis vectors, for both analytic and empirical basis vectors. The projection in each of the principle directions is a series that can be considered either a deterministic or stochastic component of the true surface where each series is termed 'a component of terrain in the lth principle direction'. The method is applied to experimental data to demonstrate the concept and to form a basis for discussion. Further application of this concept is discussed in the context of studying the partial autocorrelation of the stochastic components of the true surface to identify if it is appropriate to implement an AR model to characterize the process, followed by concluding remarks.

5.2 Chapter 5 Background

5.2.1   Terrestrial Terrain Measurement and INS Drift Compensation

Typically, terrestrial terrain surface measurement systems incorporate a scanning laser (Herr 1996) that is rigidly mounted to the body of a host vehicle. This vehicle traverses the terrain while simultaneously acquiring terrain measurements. When the vehicle encounters a disturbance, the laser translates and rotates with the body of the host vehicle. To obtain accurate terrain measurements, the motion of the vehicle must be accurately measured so that it can be removed from the laser measurement. Modern systems use INS to measure the vehicle movement (Kennedy, Hamilton et al. 2006). The accuracy of the INS depends on the alignment of the Inertial Measurement Unit (IMU) to the laser and satellite coverage of the Global Positioning System (GPS). The reduction of accuracy of the INS is contributed by many different error types (Lanyi and Roth 1988; Coco, Coker et al. 1991; Sardon, Rius et al. 1994; Sardon and Zarraoa 1997; Mannucci, Hajj et al. 2004; Sarma 2008). A detailed discussion of terrain measurement systems is presented in Section 2.1.1 and a thorough discussion of INS and GPS error is presented in Section 2.1.3 of this dissertation.

Recall the compensation method developed in Chapter 3 to correct INS drift in terrain measurements in which each terrain surface is considered to be a combination of a true surface and an error surface. The error is decomposed into drift (global error) and noise (local error). The global and local subspaces are constructed such that the drift is modeled as a random walk process while the noise is a zero-mean process. This theoretical development is coupled with careful experimental design to develop a method to identify the drift component of error and discriminate it from true terrain surface features, and correct for the INS drift (Gillespie, Sayers et al.). This Chapter classifies global and local terrain surface features into deterministic or stochastic processes. The number of basis vectors is incremented until the residual from the sum of the projections (i.e., the noise) becomes a zero-mean process. Experimentation has shown that two basis vectors are required to satisfy this criterion for a set of U.S. Highway data (Chemistruck, Binns et al. 2010). Experimental results have also shown that the first two basis vectors represent some vertical offset, defined as the principle characteristic of elevation and a

slope offset defined as the principle characteristic of bank angle. This work expands upon the work presented in Chapters 3 and 4 by examining the projection of the true surface onto orthonormal basis vectors that define the principle directions to formulate components of terrain and a truncated terrain surface.

5.2.2 Autoregressive Modeling

An Autoregressive (AR) model is comprised of a small set of coefficients and a residual process which drives the model. The AR model characterizes the current value of a profile as the linear combination of previous profile values and a residual process. If the profile is known, then the AR coefficients can be derived (e.g. via the Yule-Walker method), and the realization of the residual process is calculated. If the residual process is homogeneous, then the probability distribution can be represented as a Cumulative Probability Function (CPF). The set of AR model parameters for the stable AR model and the residual probability parameters compose a final set of model parameters that completely capture the physical characteristics of the terrain profile (Kern and Ferris 2006; Kern and Ferris 2006; Kern and Ferris 2007; Li and Sandu 2007; Wagner and Ferris 2007; Wagner and Ferris 2008; Wagner and Ferris 2010; Wagner and Ferris 2010). For this work, the AR model will be used to characterize suitable components of terrain and then synthesize a stochastic representation of the component for use in formulating a stochastic terrain surface.

5.2.3 Defining the Coordinate System

The terrain coordinate system used to develop the method presented in this chapter is described in Chapter 2, Section 2.1.4 of this dissertation.

5.2.4 Defining the Hilbert Space

A terrain surface measurement $\{z_{i,k}\}$, is considered a combination of the true surface, $\{s_i\}$, and a realization of a stochastic error surface process (refer to Section 3.3.2 for a detailed description). Each measured surface is then a realization of a stochastic process and is defined as a sequence of vectors that are elements of a Hilbert space, **H**. Since there is exactly one true surface, $\{s_i\}$, for any set of realizations, it is not indexed

141

by k. It is shown that the true surface must also span the Hilbert space **H** (Chemistruck, Binns et al. 2010).

The Hilbert space is decomposed into the principle directions of the terrain surface. A set of orthonormal basis vectors, $\mathbf{b}_l$, are constructed to span the space via a Singular Value Decomposition (SVD) of the set of measured vectors. SVD is used for principle component analysis (PCA) and was chosen for terrain surface analysis because paved roads share the same primary characteristics, such as elevation, bank angle, crowning, etc. By implementing SVD, these primary characteristics can be studied as individual terrain features. The order of the basis vectors is determined by the magnitude of their corresponding singular value, added in descending value. Basis vectors are added to the composition of the true surface until the residual noise vectors can be considered realizations of a zero-mean process; this number of basis vectors is defined as q. Recall that it was shown in Chapter 4 that the empirical basis vectors can be approximated by a set of analytic polynomials: the Legendre polynomials, in an effort to reduce the amount of information required to describe the terrain surfaces of similar types of terrain. Both formulations of components of terrain and truncated surfaces pertaining to empirical basis vectors and analytic basis vectors will be studied in this Chapter to develop a compact, high-fidelity, mixed deterministic-stochastic representation of terrain. The user will then be able to decide which formulation is most appropriate depending on the application.

Recall the following example from Chapter 4, Section 4.2.1, where the terrain surface is mapped to uniform grid spacing and the data have been acquired from a scanning laser. Let s_i be the vector representing the true surface of the section of terrain at the longitudinal location indexed by i. Let vector s_i contain elements indexed 1 to 80 corresponding to transverse points in the $v(j)$ direction, such that the terrain section has a width of 80 points, as shown in Figure 33. Furthermore, let the set of vectors $\{s_i\}$ span 10,000 locations in the $u(i)$ direction, such that the terrain section has a length of 10,000 points. Using this representation, $\{s_i\}$ contains 800,000 elevation points to describe the section of terrain. A method is developed in this Chapter that reduces the number of points required to describe the terrain surface to no more than the 20,000 points required to describe two wheel path profiles.
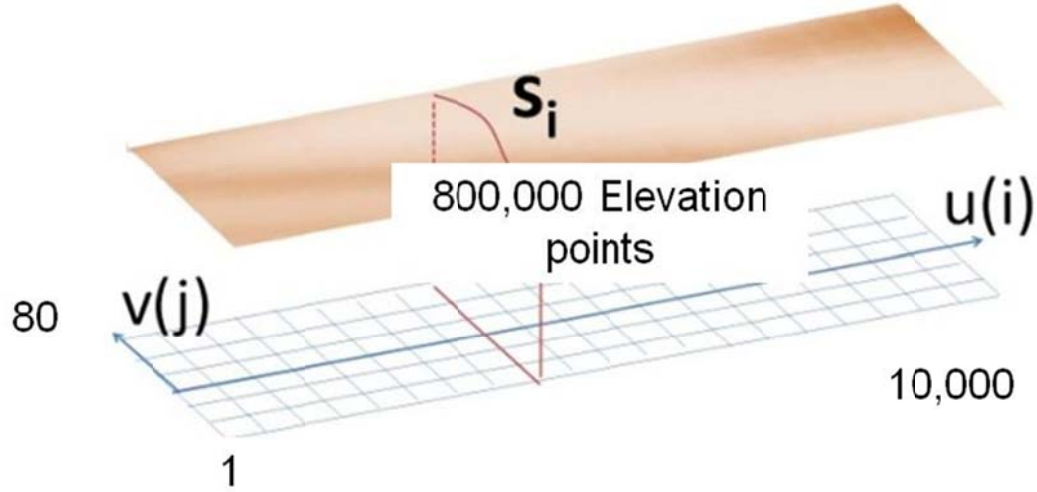
Figure 55. Example of discretized terrain surface in Curved Regular Grid Format.

The true surface is estimated by considering the estimated expectation of the sequence of measured vectors projected in the global vector space, **G**, and its complement separately as shown in Equation 76 (Refer to Section 3.3.5 for detail on how this equation was derived). Note that the second summation in Equation 76 would be identically zero, if the expectation were known exactly (or a linear estimator were used). The second term in the summation is retained for a more general derivation of results (when a nonlinear estimator is used). The projection of each measured surface vector, $z_{i,k}$, in the direction of each global direction for the empirical basis vectors is defined as $\beta_{i,k,l}$, as shown in Equation 77.

$$\hat{s}_i \triangleq \sum_{l=1}^{q} \widehat{E}[\beta_{i,k,l}]b_l + \sum_{l=1}^{q} \widehat{E}[z_{i,k} - \beta_{i,k,l}]b_l$$

Equation 76

$$\beta_{i,k,l} \triangleq \langle b_l, z_{i,k} \rangle$$

Equation 77

The true surface developed in terms of the analytic basis vectors is the same as Equation 76 except that the projections are onto the analytic basis vectors, p_l as shown in Equation 78. The projection of each measured surface vector, $z_{i,k}$, in the direction of each global direction for the analytic basis vectors is defined as $\beta_{i,k,l}$ as shown in Equation 79. Note that the same notation for $\beta_{i,k,l}$ is used since the developments in this Chapter are independent of the choice of basis vector (analytic or empirical).

$$\hat{s}_i \triangleq \sum_{l=1}^{q} \widehat{E}[\beta_{i,k,l}]\boldsymbol{p}_l + \sum_{l=1}^{q} \widehat{E}[\boldsymbol{z}_{i,k} - \beta_{i,k,l}]\boldsymbol{p}_l \qquad \text{Equation 78}$$

$$\beta_{i,k,l} \triangleq \langle \boldsymbol{p}_l, \boldsymbol{z}_{i,k} \rangle \qquad \text{Equation 79}$$

5.3 Classifying Terrain Surfaces

The true surface is decomposed into contributions in the principle directions defined by the empirical orthonormal basis vectors, as shown in Equation 76 or by the analytic orthonormal basis vectors in Equation 86. The projection of the estimated true surface is defined in each of the principle directions, $\sigma_{i,l}$, for the empirical basis vectors, as shown in Equation 88. Similarly, the projection of the estimated true surface is defined in each of the principle directions, $\sigma_{i,l}$, for the analytic basis vectors, as shown in Equation 89.

$$\sigma_{i,l} = \langle \boldsymbol{b}_l, \hat{s}_i \rangle = \sum_{l=1}^{q} \widehat{E}[\beta_{i,k,l}] + \sum_{l=1}^{q} \widehat{E}[\boldsymbol{z}_{i,k} - \beta_{i,k,l}] \qquad \text{Equation 80}$$

$$\sigma_{i,l} = \langle \boldsymbol{p}_l, \hat{s}_i \rangle_W \qquad \text{Equation 81}$$

Each component of terrain, $\sigma_{i,l}$, can be considered either a deterministic component of the true surface, or a stochastic component. This concept will be discussed further in the following section. A truncated surface can then be defined as the summation of the $\sigma_{i,l}$ multiplied by the truncated set of t empirical basis vectors ($l = 1, 2, \ldots, t$) and is defined as Equation 82. A truncated surface can also be defined in terms of the truncated set of t analytic basis vectors as defined in Equation 83. It is important to note that the truncated set of basis vectors can be incremented to include all of the basis vectors defining the space, n or it can be reduced to some value less than n such as q basis vectors used to identify contributions to the global subspace from Chapter 3.

$$\hat{s}_{trunc(i)} = \sum_{l=1}^{t \leq n} \sigma_{i,l} \, \boldsymbol{b}_l \qquad\qquad \text{Equation 82}$$

$$\hat{s}_{trunc(i)} = \sum_{l=1}^{t \leq n} \sigma_{i,l} \, \boldsymbol{p}_l \qquad\qquad \text{Equation 83}$$

5.3.1 Hybrid deterministic-stochastic representation of terrain surfaces

Features of the true surface can be recreated exactly if all of the components $\sigma_{i,l}$ are considered to be deterministic and are recorded for the entire longitudinal length of the surface. This approach does not allow the surface to be represented in a more compact form, nor give additional insight into the fundamental characteristics of the terrain. However, if some, or all, of the $\sigma_{i,l}$ are considered to be stochastic, then the sequence can be modeled as a stochastic process, and it may be possible to cast the process in a mathematical framework such as an AR model. In this case, the coefficients of the model would characterize the manner in which the process behaves in that principle direction.

The accuracy of the terrain surface representation is dependent on how each $\sigma_{i,l}$ is classified. If all of the terrain components, $\sigma_{i,l}$, are classified as deterministic, then the synthesized terrain surface would exactly resemble the true terrain surface. However, if only the first $\sigma_{i,l}$ is classified as deterministic, and the remaining $\sigma_{i,l}$ are classified as stochastic, then some variation in the realization will arise. The number of $\sigma_{i,l}$ that are modeled as stochastic is determined by the user and the application. This variation can be useful in chassis loading predictions based on tire and vehicle models. If it is desired to execute a vehicle simulation over several realizations of the same terrain type to identify maximum and minimum load cases, then the more $\sigma_{i,l}$ that are classified as stochastic will allow for more variation within that type of terrain. The effects of identifying the number of components that are deterministic or stochastic will be determined by studying the autocorrelation and the partial autocorrelation of each $\sigma_{i,l}$ to identify the correlation length. Examining the correlation length, will help to identify if that component should be modeled as a stochastic process and what model is most

appropriate. The longitudinal jointly tined concrete example is considered in the Proof of Concept section of this Chapter.

If some of the true surface components of the principle directions are modeled as an underlying stochastic process, then an Autoregressive (AR) technique can be employed. This characterization could be used to model the stochastic true surface components (Kern and Ferris 2006; Kern and Ferris 2006; Kern and Ferris 2006; Kern, Ferris et al. 2007; Wagner and Ferris 2007; Wagner and Ferris 2008). An example of the AR modeling capabilities is presented in the Proof of Concept section.

5.4 Chapter 5 Proof of Concept

An example is presented to demonstrate the decomposition of a terrain surface into components represented by the $\sigma_{i,l}$ and the implications of modeling these components as either deterministic or realizations of a stochastic process. A single lane of transversely tined jointed concrete, approximately 151.1m long, was selected at the MnRoad testing facility in Albertville, Minnesota for this experiment. A detailed description of the example of terrain and the data acquisition process are developed in Section 3.4.1. A profile taken from the true terrain surface is shown in Figure 56. The analysis is conducted in terms of longitudinal profiles to demonstrate the implications of this method. The empirical basis vectors are employed first and then the analytic basis vectors are employed for a comparison and to show the feasibility of approximating the empirical basis vectors with the analytic basis vectors.
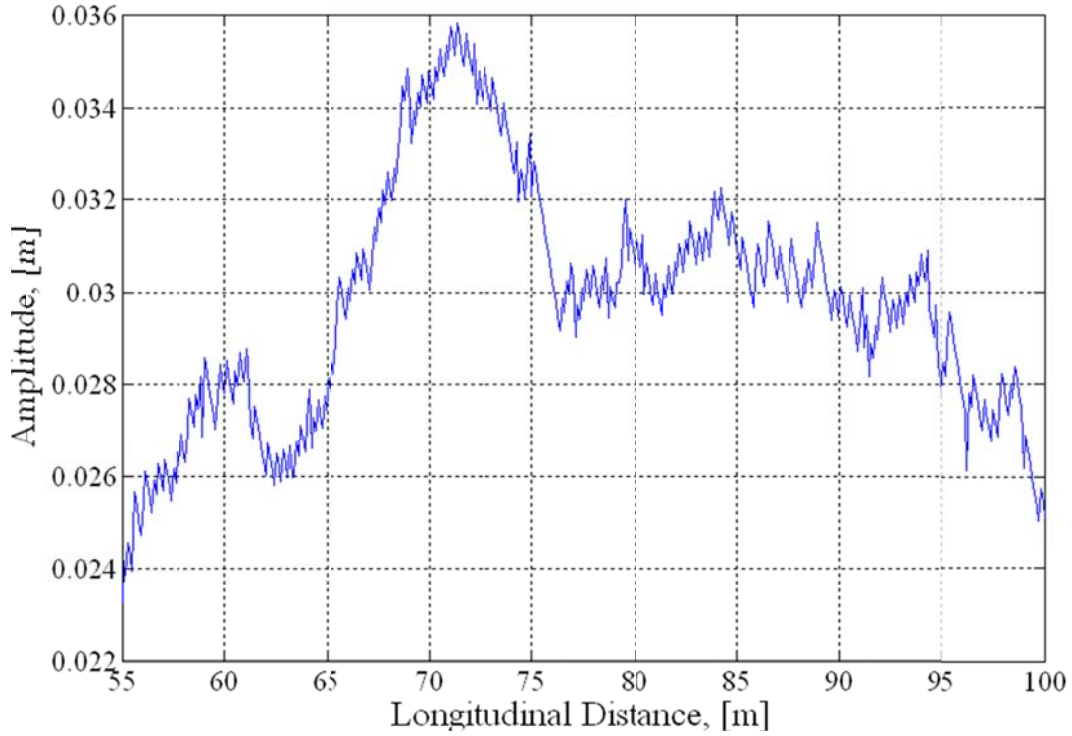
Figure 56. Original road profile.

5.4.1 Application of Theory

Basis Vectors- Empirical and Analytic

In practice, the number of basis vectors, q, sufficient to define the global subspace, **G**, must be incremented until the noise vectors, $n_{i,k}$, have a mean that is not statistically different than zero. For this set of terrain, two basis vectors were required to make the noise vectors a zero-mean process. The order in which these basis vectors are added is determined by the magnitude of their corresponding singular value in descending order. The first singular value has a magnitude of 6.41, and the second singular value has a magnitude of 1.59, and for this example the subsequent singular values are negligible. The shapes of the first two empirical basis vectors are plotted in Figure 57. These shapes correspond to the predominant changes in the road which typically include elevation change (uphill or downhill grades), bank angle (increased on turns, flatter on straights), crowning on major roads to allow for proper draining and road rutting due to tire wear. In general terms, the first basis vector accounts for a constant

147

offset in elevation and the second basis vector describes the 'bank angle' or 'road camber' in the terrain set.
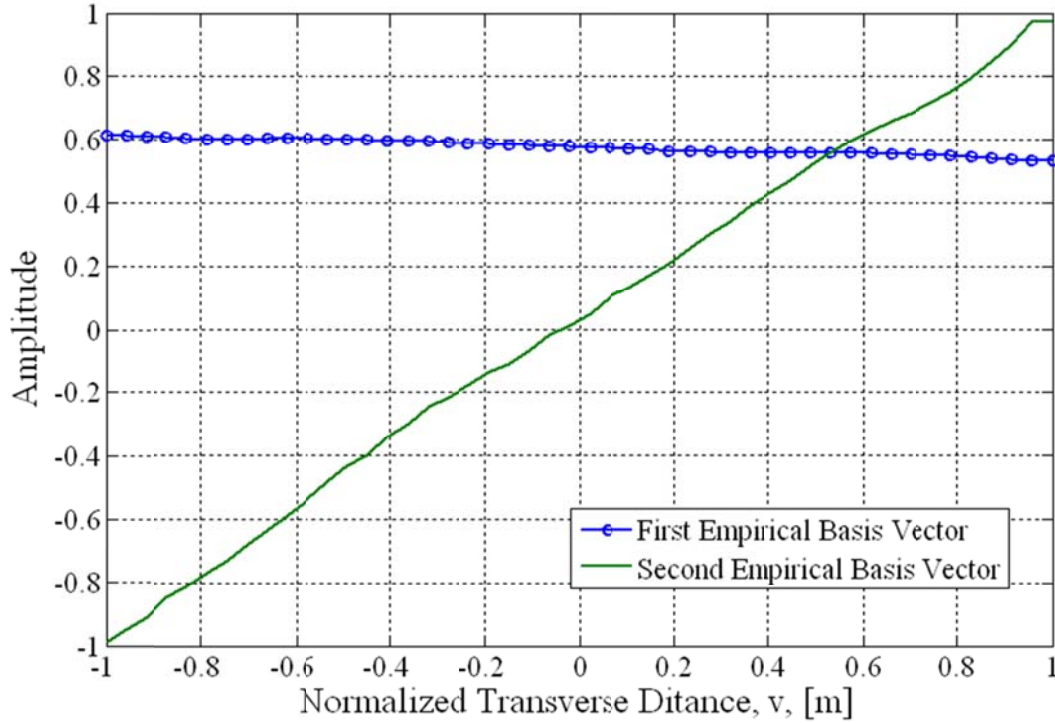
Figure 57. First two empirical basis vectors, b_l.

The shapes of the first two analytic basis vectors are plotted in Figure 58. The Legendre polynomials are used because they minimize the root mean square error between analytic and empirical basis vectors, as developed in Chapter 4. The analytic basis vectors correspond to the predominant changes in the road in the same way as the empirical vectors, which typically include elevation change (uphill or downhill grades), bank angle (increased on turns, flatter on straights), crowning on major roads to allow for proper draining and road rutting due to tire wear, except that the analytic basis vectors are smooth curves with no irregularity.
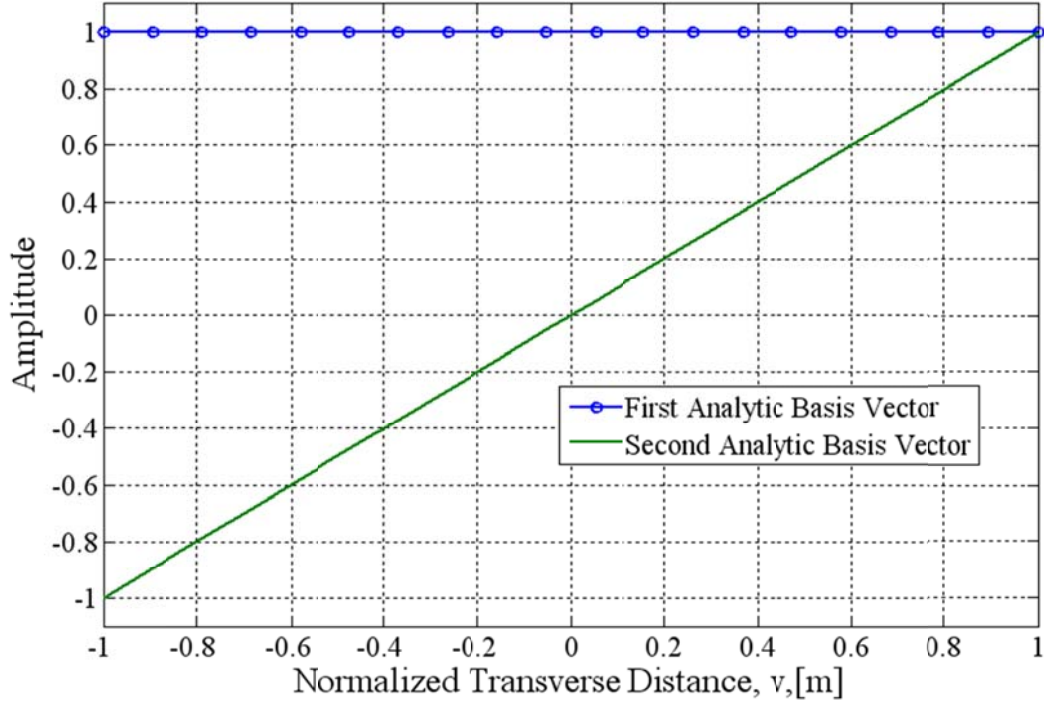
Figure 58. First two analytic basis vectors, $\boldsymbol{p}_l$.

Components of Terrain for the lth principle direction – Empircal and Analytic

The projection of the estimated true surface onto the first two empirical basis vectors is depicted in Figure 59. Executing this projection identifies the magnitude of each principle direction (basis vector) along the length of the road and is known as the component of terrain for the lth principle direction, $\sigma_{i,l}$ (refer to Equation 88). For example, if the terrain was perfectly flat, but vertically offset, then the corresponding component, $\sigma_{i,1}$, would be proportional to the magnitude of the offset. Likewise, if the terrain was perfectly flat and banked at a constant angle, then the corresponding component, $\sigma_{i,2}$, would be proportional to the bank angle of the terrain. There is minimal contribution from the second basis vector in this example as there does not appear to be an identifiable bank angle in the original profile. These observations agree with the magnitudes of the singular values; the first component, $\sigma_{i,1}$, has the greatest contribution in representing the terrain surface, and the remaining components, $\sigma_{i,1}$, have minimal contributions to representing the terrain surface for U.S. Highway data. In this example, consider the first component, $\sigma_{i,1}$, to be a deterministic process that captures the major

149

terrain characteristics of the terrain surface and subsequent $\sigma_{i,l}$ ($l > 1$) are identified as stochastic processes.
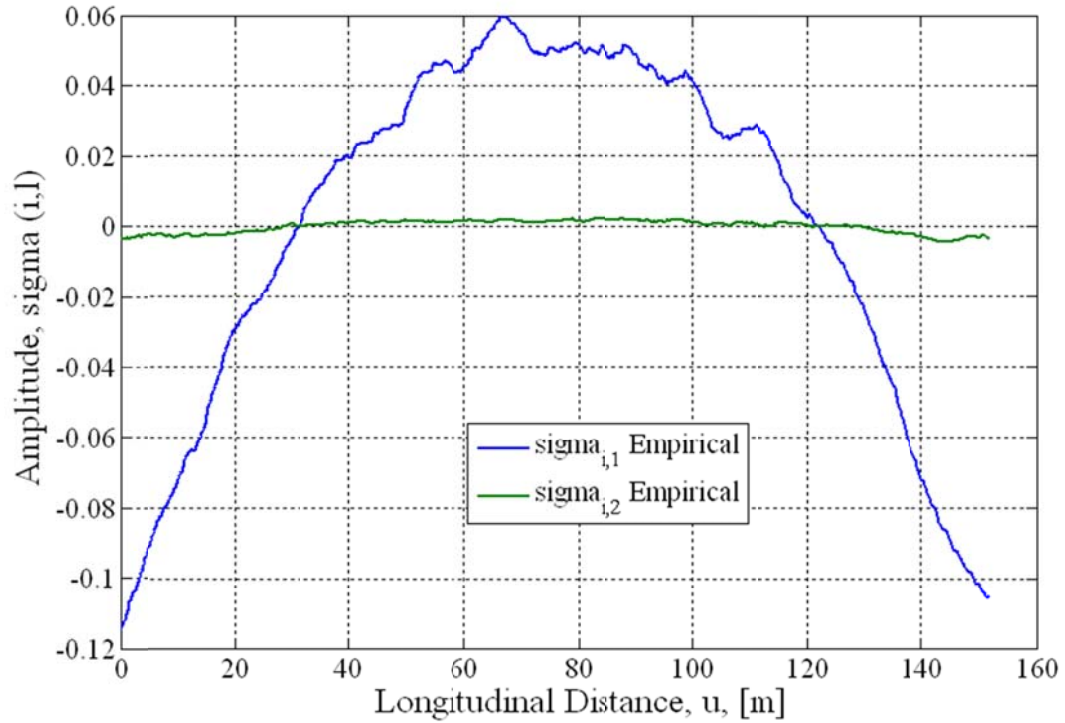
Figure 59. First two empirical $\sigma_{i,l}$.

The projection of the estimated true surface onto the first two analytic basis vectors is depicted in Figure 60. Executing this projection identifies the magnitude of each principle direction (basis vector) along the length of the road and is defined as the component of terrain for the lth principle direction, $\sigma_{i,l}$ (refer to Equation 89). It can be seen from comparing Figure 59 to Figure 60that there is more content in the analytic components of terrain. Referring to Section 4.4.2, this can be attributed to the fact that the empirical basis vectors contain the inherent variability in this particular terrain surface. The analytic basis vectors are developed to represent one *type* of terrain surface, rather than any one particular terrain surface of that type, and are therefore smoother than the empirical basis vectors. The differences in the magnitudes of the components of the terrain depend on the scaling of the corresponding basis vectors, therefore, the magnitudes of the components of terrain differ between the empirical and analytical basis

150

vectors. For example, consider the first empirical basis vector; the magnitude is approximately 0.6 as compared with the analytic basis vector that is exactly one.
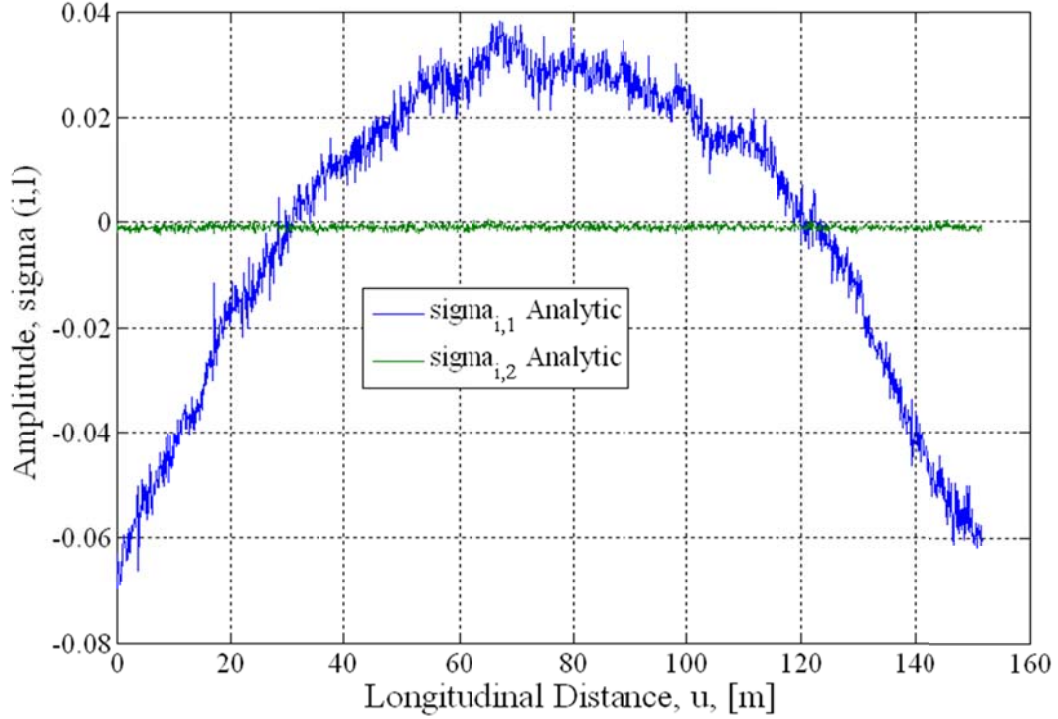
Figure 60. First two analytic $\sigma_{i,l}$.

Truncated Surface Formulation- Empirical and Analytic

The truncated surfaces generated from the empirical basis vectors and the analytic basis vectors both agree well with the true terrain surface. The truncated surfaces are based on using all of the available information on the terrain, i.e. all basis vectors and all components of terrain- a deterministic approach (48 basis vectors and 48 components of terrain for this example). The average difference between the analytically generated terrain surface and the true terrain surface is on the order of 10^{-5}m. This disagreement may be attributed to a numerical round-off issue in the calculation of the weighting matrix. The average difference between the empirically generated terrain surface and the true terrain surface is on the order of 10^{-16}m and is negligible. This is expected since the truncated surface based on the empirical basis vectors and components of terrain are produced from the same data sets that formulated the true terrain surface. None-the-less, the approximation of the empirical basis vectors by Legendre polynomials proves to be a good fit, and a difference of 10^{-5}m is two orders of magnitude smaller than the resolution

151

of the original terrain surface measurement and deviations of 10 μm will not affect the handling, ride, or durability of a vehicle.

Consider the first two components of the terrain, $\sigma_{i,l}$, corresponding to the first two empirical basis vectors $\boldsymbol{b_l}$. Understanding how each $\sigma_{i,l}$ and $\boldsymbol{b_l}$ contributes to the truncated surface compared with the true surface is identified in Figure 61. It can be seen that as the contributions from each $\sigma_{i,l}$ are added, the truncated surface converges to the true surface. The true terrain surface is represented by the blue line, the truncated surface pertaining to the first empirical basis vector and component of terrain is represented by the green line, and the first plus the second empirical basis vectors and components of terrain are represented by the green line. The peaks between the second truncated terrain surface and the true terrain surface differ by 0.4mm.
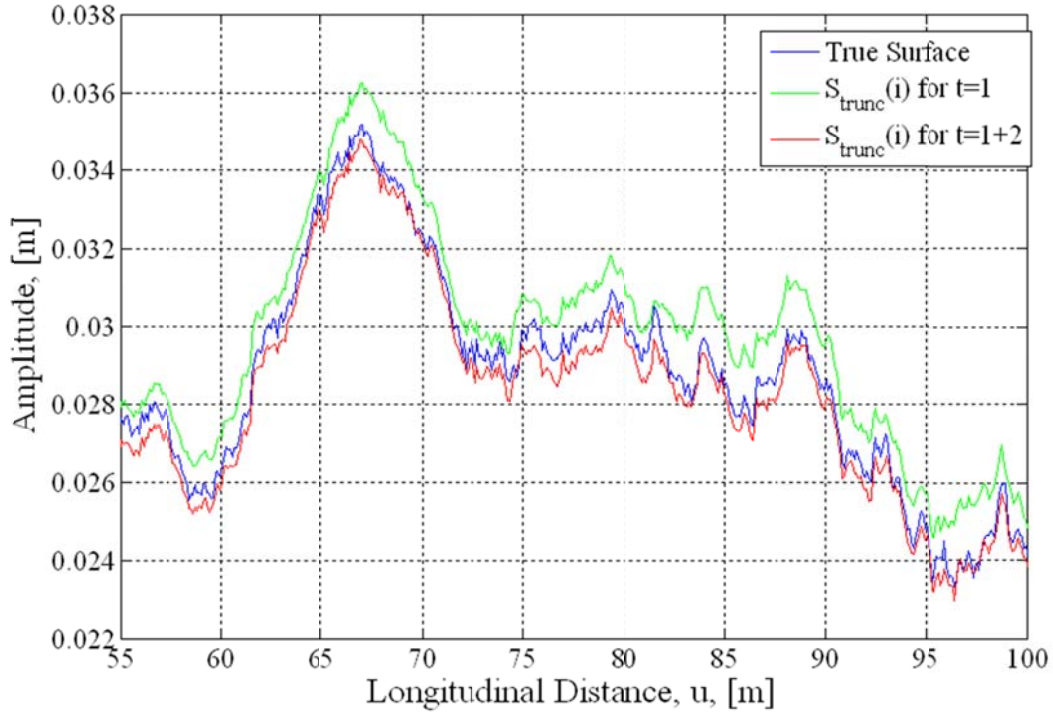
Figure 61. Comparing true surface to subsequent empirical truncated surfaces.

5.4.2 Hybrid Deterministic-Stochastic Representation of Terrain Surfaces

Consider the transversely tined jointed concrete studied in this Chapter and the normalized partial autocorrelation of the first component in Figure 62. The partial autocorrelation of $\sigma_{i,l}$ does not decay until lag 22. This could indicate that $\sigma_{i,l}$ should be

modeled as a deterministic component of the true surface. However, the partial autocorrelation of $\sigma_{i,2}$ as shown in Figure 63 and $\sigma_{i,3}$ (not shown) decay within 6 lags and are more appropriate candidates to be represented as a stochastic process. Since the partial autocorrelation decays quickly, an AR model would be appropriate to implement. This brief study suggests that the first $\sigma_{i,1}$ could be represented as a deterministic process and the remaining components can be modeled as a stochastic process. If this were the case, then elevation changes would be exactly preserved in the first component, $\sigma_{i,1}$, and stochastic variation would be induced in the remaining components, $\sigma_{i,l}$ ($l > 1$) in directions such as roll and crowning. Thus the stochastic nature would produce different variations of the original road surface.
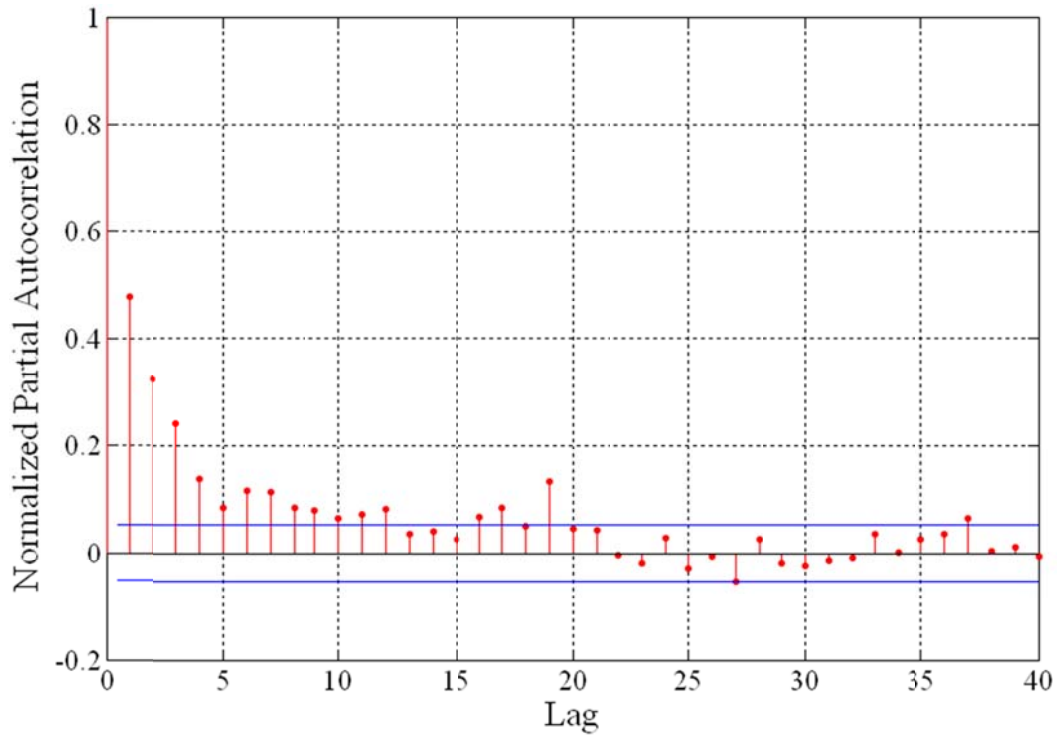
Figure 62. Partial autocorrelation of $\sigma_{i,1}$.
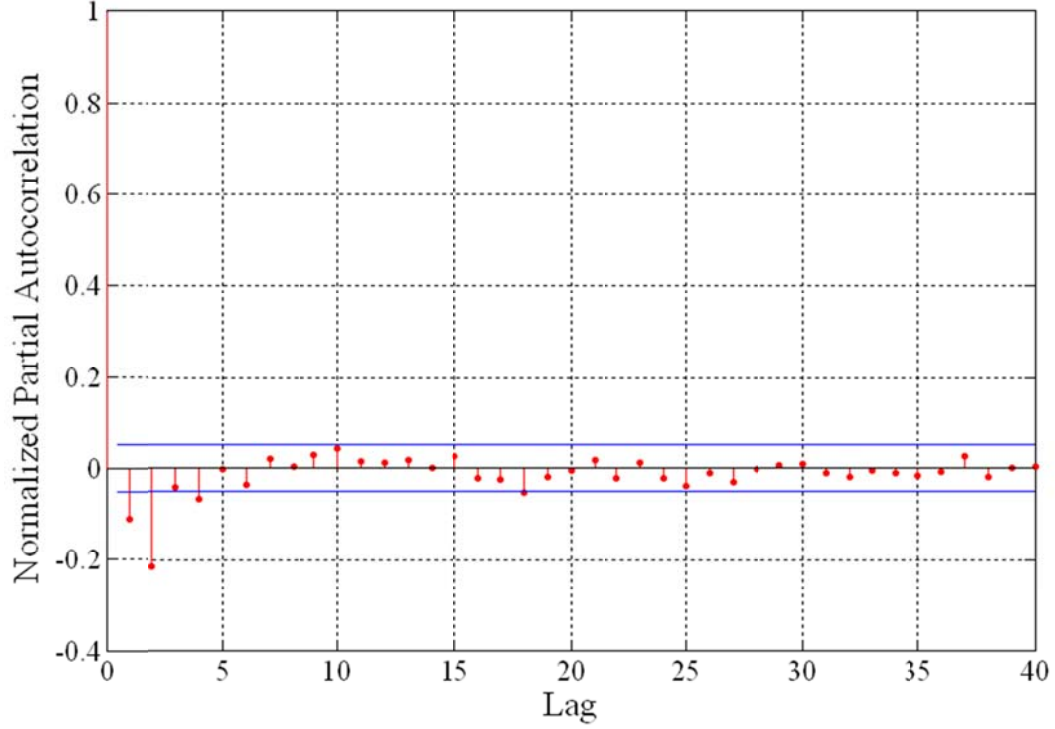
Figure 63. Partial autocorrelation of $\sigma_{i,2}$.

Consider an example of the AR modeling capabilities for the bank angle component of terrain in Figure 64. Since the correlation length is short (approximately a lag of 6 from Figure 63), modeling this component seems appropriate. The empirically calculated bank angle component of terrain is shown in blue, and the AR synthesized component of terrain is shown as a red dashed line. It can be seen that the synthesized curve maintains the major characteristics of the original component of terrain but due to the AR residual process, the synthesized component exhibits different trends than the empirical component. In terms of the truncated surface, the differences in the synthesized bank angle component of terrain can be seen in Figure 65. Again, the synthesized truncated surface is shown as a red dashed line and the true surface is shown as a blue solid line. Since the variation exhibited by the two components of terrain is of small magnitude, profiles taken from the surface were used to illustrate the effects of a stochastically represented bank angle component of terrain. It would be expected that more drastic changes would result from stochastically modeling the elevation component of terrain as that component has the greatest effect in depicting the terrain surface.
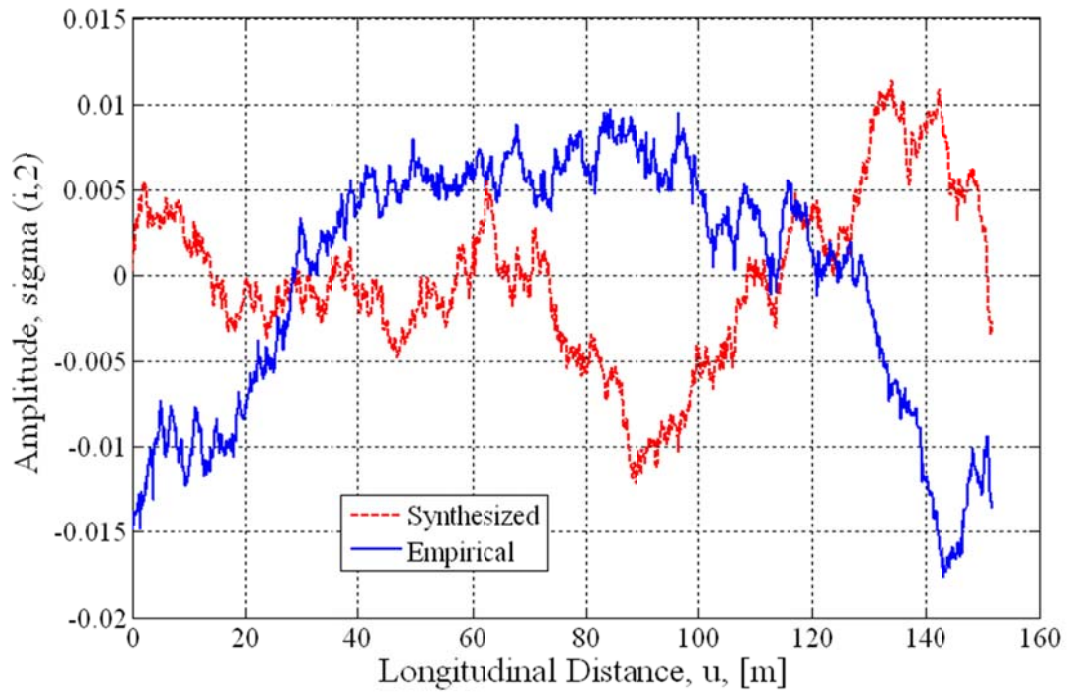
Figure 64. Empirical bank angle component of terrain and AR synthesis of bank angle component of terrain.
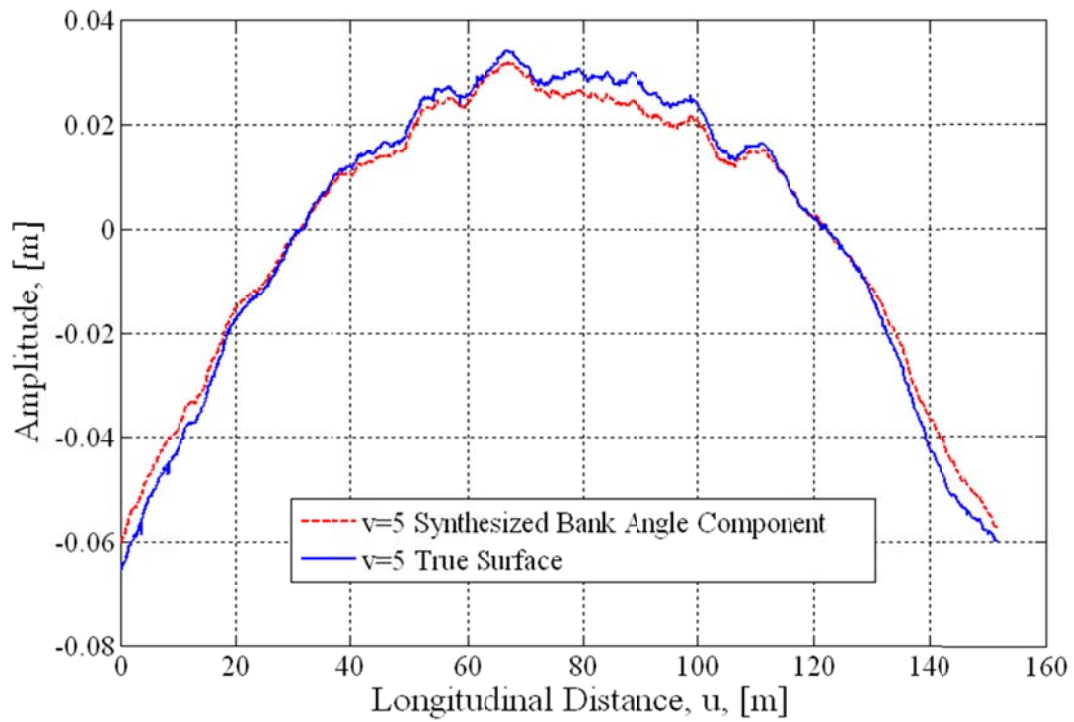
Figure 65. Comparing true surface to truncated surface with AR synthesized bank angle component of terrain.

5.4.3 Implication of Surface Representation on Statistical Analyses

This new formulation of a compact model improves the ability to analyze the roughness of the terrain surface. Currently, all roughness indices analyze longitudinal profiles. A better representation of the terrain roughness is realized by analyzing each component of terrain. For example, consider the first component of terrain, the elevation. The International Roughness Index (IRI) as described in Section 2.1.3.1 is the standard practice for calculating road roughness. The calculations of the IRI are also described in Section 2.1.3.1. The IRI of the elevation component of terrain, left and right wheel paths of the true surface and mean of the left and right wheel paths were calculated and presented in Table 2. There are obvious differences between the IRI calculated for the left and right so the mean IRI between the wheel paths was calculated. Comparing the IRI of the elevation component of terrain to the mean IRI of the left and right wheel paths produces a 0.138% difference. Since the elevation component of terrain identifies the magnitude of the first basis vector along the length of the terrain section, it seems more appropriate than arbitrarily using the left and right wheel path longitudinal profiles as an analysis.

Table 2. IRI comparison between components of terrain and longitudinal profiles.

	International Roughness Index
Elevation Component of Terrain	0.5756
Left Wheel Path Longitudinal Profile	0.5680
Right Wheel Path Longitudinal Profile	0.5848
Mean of Wheel Paths	0.5764

In addition to response surface modeling (discussed in Section 3.5), previous research with Empirical Mode Decomposition (EMD) showed that combining certain features from each level of decomposition produced an identifiable terrain feature (Attoh-Okine, Ayenu-Prah Jr et al. 2006). For example, if contributions from the first and third basis vectors are combined, it may be possible to identify a new surface characteristic, other than elevation, bank angle, crowning or rutting. By construction, the basis vectors are orthonormal, however the projection of the terrain surface vectors onto these basis

vectors may not be independent. It is likely that there is a relationship between the components of terrain, $\sigma_{i,l}$, that can be used to further simplify the representation.

5.5 Chapter 5 Discussion

A detailed review of the effectiveness of this method via a specific example provides more insight into the usefulness of this method in developing a compact representation of terrain. It was shown in the INS drift compensation example from Chapter 3 and discussed earlier in this Chapter (see Figure 55) that two basis vectors are needed to establish a zero-mean noise process, suggesting that only the first two components of terrain are required to form the truncated surface. Consider the first two empirical basis vectors of size [48 x 1] and two corresponding components of terrain of the size [15,180 x 1]. The amount of data needed to represent this truncated terrain surface is 30,360 points (plus the 96 points needed to define the two empirical basis vectors). The same number of points is required with using the analytic basis vectors, however the data required to define the analytic basis vectors is simplified by using known polynomial expressions. In either case, these representations are drastic reductions from the original terrain surface, a matrix of the size [48 x 15,180] or a total of 728,640 elevation points. Now consider the use of a Hybrid deterministic-stochastic representation of the terrain surface in which the first component is deterministic, requiring 15,180 points, and the second component is stochastic, requiring perhaps 11 AR model coefficients and 3 coefficients to define the probability distribution of the residual process (Wagner, et al.). The total reduction in complexity is reduced by a factor of 50. If both of the components were represented as stochastic processes, then this factor improves dramatically to nearly 6,000 – nearly four orders of magnitude.

The demonstration of the method developed in this work is limited by the amount of available measured data. Future work will study the capabilities of this method beyond the application to U.S. Highway data and investigate its compatibility with off-road terrain and harsher on-road terrain surfaces. The effect that the stochastic components will have in comparison with the original terrain should also be investigated. The original terrain and stochastic terrain will differ when compared feature by feature, but the overall physical properties will be the same. This overall effect will be better

understood when terrain surfaces are used in tire and vehicle simulations. Future work will investigate the overall effects of the stochastic surface on the loads generated by the tire and vehicle based on changing different surface components.

5.6 Chapter 5 Conclusions

Terrain profiles capture an incomplete representation of terrain topology data whereas terrain surfaces capture a great deal more information about the terrain. Extracting two longitudinal wheel path profiles using a collocation approach from the example presented in Section 5.2.4 results in a 20,000 point realization. Alternatively, the discrete terrain surface can be represented exactly, requiring 800,000 elevation points for this example. A computational problem arises when trying to work with and manipulate measured terrain surface data due to the large file sizes. A new representation has been developed that captures the principle information about a terrain surface in a compact form, similar in structure to terrain profiles. If two basis vectors representing elevation and bank angle are used in place of the two longitudinal wheel path profiles and two components of terrain corresponding to these basis vectors were used instead, then the principle information about the terrain surface (in the Galerkin sense) is captured using the same amount of data, 20,000 points total. This new development increases computational efficiency while maintaining sufficient terrain surface knowledge. It was demonstrated through an example that this new process successfully captures the characteristics of a terrain surface with as few as two principle directions. Once this decomposition is complete, the projection of the true surface onto the principle directions can be modeled as either deterministic or a stochastic process. If the component is identified as stochastic, then a mathematical model, such as an Autoregressive model, can be implemented. Through the example shown in the Discussion Section of this Chapter, the terrain surface could be further reduced to two basis vectors, one deterministic component of terrain and one stochastically modeled component terrain, significantly reducing by a factor of 50 the data necessary to represent the surface. It was also shown that using the elevation component of terrain may provide a better way to analyze the road roughness of the surface. These components of terrain identify the magnitude of elevation along the length of the road and therefore serve as a better representation of the

surface as opposed to identifying two arbitrary profiles to represent the left and right wheel paths, and then formulate roughness for the two profiles.

6. Conclusions

6.1 Summary of Research

The concept of simulation-based engineering has been embraced by virtually every research and industry sector (Sinha, Liang et al. 2001; Mocko and Fenves 2003). Engineering and science communities have become increasingly aware that computer simulation is an indispensable tool for resolving a multitude of scientific and technological problems. Design and analysis engineers are simulating increasingly complex mechanical systems. It is clearly desirable to gain a reliable perspective on the behaviour of a system early in the design stage, long before building costly prototypes (Chul and Ro 2002; Letherwood, Gunter et al. 2004; Makarand Datar 2007; Ersal, Fathy et al. 2008; Mueller, Ferris et al. 2009). Simulation tools have become a critical part of the automotive industry due to their ability to reduce the time and money spent in the development process.

Terrain is clearly the principle source of vertical excitation to the vehicle and must be accurately represented in order to correctly predict the vehicle response in simulation. Ideally, an efficient terrain surface definition could be developed that maintains the high-fidelity information required to accurately excite vehicle models. Modern terrain measurement systems use an Inertial Navigation System (INS) to measure and remove vehicle movement from laser measurements of the terrain surface. Instrumental and environmental errors inherent in the INS produce noise and drift errors in the resulting estimates of vehicle position and orientation. The evolution and implications of terrain surface measurement techniques and existing methods for correcting INS drift have been reviewed as a framework for a new compensation method for INS drift in terrain surface measurements. Each measurement is considered a combination of the true surface and the error surface, defined on a Hilbert vector space, in which the error is decomposed into drift (global error) and noise (local error). The global and local subspaces are constructed such that the drift is modeled as a random walk process and the noise is a zero-mean process. This theoretical development is

coupled with careful experimental design to identify the drift component of error and discriminate it from true terrain surface features, thereby correcting for the INS drift. It is shown through an example that this new compensation method dramatically reduces the error variation in the measured surfaces to within the resolution of the measurement system itself. It was also shown through the example using U.S. Highway data that classifying the first two basis vectors as spanning the global subspace results in a zero-mean noise process with a standard deviation of 0.35mm. This is approximately three standard deviations within the 1mm bound defined by the resolution of the measurement equipment.

It is also desirable to develop a compact, path-specific, terrain surface representation that exploits the inherent anisotropicity in terrain traversed by vehicles. This representation should also minimize the effect of the choice of measurement system used to sample the terrain surface. Non-deformable terrain surfaces are defined as a sequence of vectors, where each vector comprises terrain heights at locations oriented perpendicular to the direction of travel. A vector space is formed by the span of these vectors and a corresponding set of empirical basis vectors is developed. A set of analytic basis vectors is formed from Gegenbauer polynomials, parameterized to approximate the empirical basis vectors. Other polynomials were studied, but the Gegenbauer polynomials with λ equal to ½ resulted in the smallest Root Mean Square Error. A weighted inner-product is defined to form a Hilbert space and the terrain surface vectors are projected onto the set of analytic basis vectors. The weighting matrix is developed such that these projections are insensitive to the number and placement of the discrete transverse locations at which the terrain heights are defined. Furthermore, components of terrain are developed from a projection of each analytic basis vector onto the terrain surface. These components of terrain capture the magnitude of specific terrain characteristics along the length of the road, such as elevation, bank angle, or road crowning, for statistical analysis and modeling techniques similar to that of a terrain profile. This method is successfully demonstrated on sets of paved road surfaces to show that a high-fidelity but compact definition of terrain surfaces is developed.

It is also desirable to evaluate vehicle models and tire models over a wide range of terrain types, but it is computationally impractical to store long distances of every

terrain surface variation. This dissertation examines the terrain surface, rather than the terrain profile, to maximize the information available to the tire model (i.e. wheel path data). A method to decompose the terrain surface as a combination of deterministic and stochastic components is developed. If any of the components of the terrain surface are considered to be stochastic, then the sequence can be modeled as a stochastic process. As a result, a stochastic terrain surface that is more computationally efficient to implement in simulation can be synthesized for any desired length. Such a surface will include variation in the synthesized data, whereas measured data is deterministic and the simulation results for measured terrain will always result in the same responses. These stochastic representations of terrain surfaces can then be implemented in tire and vehicle models to estimate chassis loads. It was also shown that the elevation component of terrain may provide a better method to analyze the road roughness of the surface rather than single-path road profiles.

6.2 Future Work

Future work that may leverage the work presented here could include but is not limited to an in-depth study of rough, non-deformable terrain surfaces and an optimal selection for sampling the terrain surface.

6.2.1 Rough, Non-deformable Terrain Surfaces

All the examples presented in this work were based on data acquired from U.S. Highways. It would be vital to include rough, non-deformable terrain surfaces in subsequent studies for classifying terrain surfaces and develop a compact representation of rough terrain. Then, durability roads could be synthesized using the methods presented in this dissertation as a concatenation of various terrain features that are important for properly assessing vehicle durability, reliability, and ride and handling characteristics.

6.2.2 Optimal Selection of {v} for Sampling the Terrain Surface

Chapter 4 of this dissertation identified the need to develop a method to define the terrain surface that can be universally applied to various types of terrain measurement systems, from point lasers to scanning lasers. The solution presented in Chapter 4, was a

best fit analytic polynomial approximation to empirical basis vectors for a specific terrain type. The analytic polynomials can be discreteized to match the capabilities of any existing terrain measurement system. While this is a numerical approach, an empirically based standard currently exists and defines what types of lasers, how many and at what locations along the width of a host vehicle the lasers should be mounted to adequately sample the terrain to identify road ruts (ASTM 2009). Other empirical testing has been conducted to identify the flaws in this standard (Mehta, Roque et al. 2001; Ihs 2004; Offrell and Sjogren 2004; Sjogren and Lundberg 2004), and thus a need to extend the work presented in Chapter 4 to develop an optimal selection of discrete locations to identify specific terrain features, such as crowning and rutting, needs to be developed. Future work leveraging the findings in Chapter 4 could focus on generating an optimal selection of the discrete sampling locations along the width of the host vehicle, also known as $\{v\}$.

6.3 Conclusions

Recall the thesis statement from Chapter 1.

A high-fidelity yet compact representation of terrain surface types can be developed that are insensitive to the particular measurement system being used and allows for the study of principle terrain characteristics.

The primary goal of this research is to develop a technique to develop a compact representation of a terrain surface while improving the fidelity of the estimated terrain surface. The methodology in this work is currently limited to:

a) Non-deformable terrain- terrain whose surface deformation due to a single vehicle traversing the surface is negligible, such as paved roads (both asphalt and concrete), gravel roads, and typical off-road trails; deformable terrain such as sand and snow are beyond the scope of this work.

b) Anisotropic terrain- terrain having an inherent path defined over which vehicles travel and where that travel, over time, has defined obvious paths or lanes.

This dissertation focuses on the development of a high-fidelity and compact representation of terrain surfaces that are insensitive to the terrain measurement system used to acquire the data. The main contributions of this research are:

1) The first representation of terrain surfaces as elements of a Hilbert Space, $\mathbf{H}$
    a) Terrain surfaces are explicitly defined as a sequence of vectors
    b) Each vector comprises terrain heights at locations oriented perpendicular to the direction of travel.
    c) A vector space is defined by the span of these vectors
    d) A weighted inner-product is defined and the norm is defined as the induced norm

2) A novel method to remove INS drift:
    a) The Hilbert space is decomposed into a Global subspace, $\mathbf{G}$, and the complementary local subspace, $\mathbf{G^c}$.
    b) A sequence of error vectors are defined such that each error vector is the difference between the measured terrain surface vector and the true surface vector.
    c) The three vectors comprising the measurement, error, and true surface at each longitudinal location have components in both the Global and Local subspaces.
    d) The Global subspace is defined such that the component of the error vectors in the Global subspace can be modelled as a random-walk process and the elements of the error vectors in the Local subspace are elements of a zero-mean uncorrelated noise process.
    e) The Global error, arising from Inertial Navigation System (INS) drift, is identified and removed from the measurements such that the standard deviation of the residual noise process (Local error) is within the resolution of the measurement system (+/-1mm)
    f) The true surface is computed from the drift-free measurements using a non-linear expectation estimator.

3) The first rigorous definition of principle terrain characteristics:

a) Although principle terrain characteristics, such as elevation, bank angle, crowning, and rutting have been subjectively described and techniques for *ad hoc* measurement of these characteristics have been defined (AASHTO 2009), this research develops sets of empirical basis vectors for U.S. Highways that are consistent between asphalt and concrete roads

b) The principle terrain characteristics are defined with analytic basis vectors that closely approximate these consistent empirical basis vectors by minimizing the root mean square error between the empirical and analytic basis vectors.

4) It is the first method that has been developed that can define terrain surfaces without sensitivity to the measurement system used to acquire the terrain data

   a) Analytic basis vectors can be discretized based on terrain measurement system capabilities to sample the terrain

5) A novel method to stochastically represent a terrain surface

   a) Components of terrain (known as $\sigma_{i,l}$) identify the magnitude of the principle terrain characteristics along the length of the terrain surface

   b) Components of terrain are classified as deterministic or stochastic

   c) Stochastic components of terrain can be modelled in the same manner as terrain profiles

   d) Truncated terrain surface is synthesized based on stochastically modelled components of terrain

The results of this study are intended for application in the accurate prediction of:

   a) Tire loads- based on full-knowledge of the tire patch: lateral, longitudinal, and radial tire forces and the overturning, rolling resistance and self-aligning tire moments acting on the contact patch.

   b) Chassis loading scenarios- better ground vehicle reliability predictions

   c) Pavement Life- crack propagation, rutting, roughness

Better modeling of the contact patch will yield more representative tire loading, which will lead to better estimates of chassis loading scenarios and will inevitably result in better ground vehicle reliability predictions. Current application of this work is limited to the analysis of U.S. Highways, comprised of both asphalt and concrete, but the

methods developed in this dissertation are applicable to any non-deformable path-specific terrain.

References

(1973). Mays Ride Meter Booklet. Austin, TX, Rainhart Co.: 12.

AASHTO (2009). Standard Practice for: Determining Pavement Deformation Parameters and Cross-slope from Collected Transverse Profiles.

Ahlin, K., J. Granlund, et al. (2004). "Comparing road profiles with vehicle perceived roughness." International Journal of Vehicle Design 35(2-3): 270-86.

Andren, P. (2006). "Power spectral density approximations of longitudinal road profiles." International Journal of Vehicle Design 40(1-3): 2-14.

Andren, P. (2006). "Power spectral density approximations of longitudinal road profiles." International Journal of Vehicle Design 40(1/2/3): 2-14.

Arfken, G. (1985). Mathematical Methods for Physicist, 3rd Ed. Orlando, Academic Press.

ASTM (1989). Standard Terminology Relating to Traveled Surface Characteristics, American Society for Testing and Materials.

ASTM (1996). Standard Test Method for Measuring Road Roughness by Static Level Method. West Conshohocken, PA, ASTM International.

ASTM (2006). E867-06A. Standard Terminology Rating to Vehicle-Pavement System, American Society for Testing and Materials.

ASTM (2009). ASTM E950-Standard Test Method for Measuring the Longitudinal Profile of Traveled Surfaces with an Accelerometer Established Inertial Profiling Reference.

Attoh-Okine, N. O., A. Y. Ayenu-Prah Jr, et al. (2006). Application of the empirical mode decomposition to pavement profile analysis, Atlanta, GA, United states, American Society of Civil Engineers.

Aurell, J. and S. Edlund (1989). Operating Severity Distribution - A Base for Vehicle Optimization. 11th International Association of Vehicle System Dynamics Symposium (IAVSD 1989). Kingston, Ontario, Canada. 18: 42-56.

Ayenu-Prah, A. Y. and N. O. Attoh-Okine (2009). "Comparative study of hilbert-huang transform, fourier transform and wavelet transform in pavement profile analysis." Vehicle System Dynamics 47(4): 437-456.

Bendat, J. S. and A. G. Piersol (2000). Random Data Analysis and Measurement Procedures, John Wiley & Sons inc.

Binns, R. (2010). Improved Terrain Measurement System for Estimation of Global Terrain Features, Surface Roughness, and Texture. Mechanical Engineering. Blacksburg, Virginia Tech. Masters of Science: 63.

Bogsjo, K. and A. Forsan (2004). "Fatigue relevant road surface statistics." Vehicle System Dynamics 41: 724-733.

Brandenburg, E. L. and T. J. Rudd (1974). Development of an Inertial Profilometer. F. R. Administration. United States. FRA-ORD&D-75-15: 47p.

Briese, C. and N. Pfeifer (2001). "Airborne Laser Scanning and Derivation of Digital Terrain Models." Optical 3D Measurement Techniques: 80-87.

Bruscella, B. and V. Rouillard (1999). "Analysis of Road Surface Profiles." Journal of Transportation Engineering 125(1): 55.

Capuruço, R. A. C., T. Hegazy, et al. (2005). "Full-Car Roughness Index as Summary Roughness Statistic." Transportation Research Record: Journal of the Transportation Research Board 1905(1): 9.

Chaika, M., D. Gorsich, et al. (2004). "Some statistical tests in the study of terrain modelling." International Journal of Vehicle Design 36(2-3): 132-148.

Chemistruck, H., R. Binns, et al. (2010). "Correcting INS Drift in Terrain Measurements." Journal of Dynamic Systems, Measurement, and Control(DS-10-1034).

Chul, K. and P. I. Ro (2002). "An accurate full car ride model using model reducing techniques." Transactions of the ASME. Journal of Mechanical Design 124(4): 697-705.

Coco, D. S., C. Coker, et al. (1991). "Variability of GPS satellite differential group delay biases." IEEE Transactions on Aerospace and Electronic Systems 27(6): 931-938.

Cook, R. D., D. S. Malkus, et al. (2002). Concepts and Applications of Finite Element Analysis. Madison, WI, Jon Wiley & Sons, Inc.

Dembski, N., G. Rizzoni, et al. (2006). Development of a terrain severity measurement system utilizing optical lasers, Bellingham WA, WA 98227-0010, United States, International Society for Optical Engineering.

Dembski, N., G. Rizzoni, et al. (2006). Development of aTerrain Severity Measurement System Utilizing Optical Lasers. SPIE: Modeling and Simulation for Military Applications, Orlando, FL, USA, SPIE.

Dembski, N., G. Rizzoni, et al. (2006). The development of a terrain severity measurement system, New York, NY 10016-5990, United States, American Society of Mechanical Engineers.

Detweiler, Z. R. and J. B. Ferris (2008). Uniform Grid Spacing for Three Dimensional High Fidelity Terrain Profiles. 16th International Conference of the International Society for Terrain-Vehicle Systems, Turin, Italy.

Detweiler, Z. R. and J. B. Ferris (2009). "Interpolation methods for High-fidelity three dimensional terrain surfaces." Journal of Terramechanics.

Donoho, D., H. Kakavand, et al. (2007). The Simplest Solution to an Underdetermined System of Linear Equations. IEEE International Symposium on Information Theory, Seattle, WA.

DOT, U. (2006). Number of Vehicles and Vehicle Classification. Washington, D.C., Bureau of Transportation.

DOT, U. (2008). Table1-1: System Mileage within the United States. Washington D.C., National Trasportation Statistics & Beureau of Transportation Statistics.

Doyle, F. J. (1978). "Digital Terrain Models: An Overview." Photogrammetric Engineering and Remote Sensing 44(12): 1481-1485.

Duni, E., G. Toniato, et al. (2008). Vehicle Fatigue Load Prediction based on Finite Element Tire/Road Interaction implemented in an Integrated Implicit-Explicit Approach. 2008 Abaqus Users' Conference: 16.

Elmqvist, M. (2001). Ground Estimation of Laser Radar data with Active Shape Models. Stockholm, Sweden, OEEPE workshop on Airborne Laserscanning and Interferometric SAR for Detailed Digital Elevation Models.

Elmqvist, M., E. Jungert, et al. (2001). "Terrain Modelling and Analysis Using Laser Scanner Data." International Archives of Photogammetry Remote Sensing and Spatial Information Sciences XXXIV(3/W4): 219-226.

Ersal, T., H. K. Fathy, et al. (2008). "A Review of Proper Modeling Techniques." Journal of Dynamic Systems, Measurement, and Control 130(6): 061008.

Favey, E., M. Cerniar, et al. (1999). Sensor Attitude Determination Using GPS Antenna Array and INS. ISPRS WG III/1 Workshop: "Direct Versus Indirect Methods of Sensor Orientation", Barcelona, Spain.

Ferris, J. B. (1999). Factors affecting perceptions of ride quality in automobiles. Dynamic Systems and Controls. Nashville, TN, USA, ASME IMECE. 65: 649-654.

Ferris, J. B. (2004). "Characterising road profiles as Markov Chains." International Journal of Vehicle Design 36(2-3): 103-115.

Ferris, J. B. and J. L. Larsen (2002). "Establishing Chassis Reliability Testing Targets Based on Road Roughness." International Journal of Materials and Product Technology 17: 453-461.

Frederick, D. and T. S. Chang (1965). Continuum Mechanics. Boston, Allyn and Bacon, Inc.

Fritsch, D. and J. Kilian (1994). "Filtering and Calibration of Laser Scanner Measurements." International Archives of Photogammetry Remote Sensing and Spatial Information Sciences XXX(Part 3/1): 227-234.

Gillespie, T. D. (1992). Everything You Always Wanted to Know about the IRI, But Were Afraid to Ask! Road Profile Users Group Meeting, Lincoln, Nebraska.

Gillespie, T. D., M. Sayers, et al. (1980). Calibration of Response-Type Road Roughness Measuring Systems, International Journal of Materials and Product Technology: 453-461.

Gillespie, T. D., M. W. Sayers, et al. (1980). "Calibration of Response-Type Road Roughness Measuring Systems." National Cooperative Highway Research Program Report(228): 453-461.

Gimmler, H., D. Ammon, et al. (2005). "Road profiles: Mobile measurement, data processing for efficient simulation and assessment of road properties." VDI Berichte(1912): 335-352.

Giremus, A., A. Doucet, et al. (2004). A Rao-Blackwellized particle filter for INS/GPS integration. Acoustics, Speech, and Signal Processing, 2004. Proceedings. (ICASSP '04). IEEE International Conference on.

Gorsich, D. J., D. Gunter, et al. (2003). Terrain Roughness Standards for Mobility and Ultra- Reliability Prediction. SAE World Congress & Exhibition, Detroit, MI.

Hajek, J. J., T. J. Kazmierowski, et al. (1998). "Switching to International Roughness Index." Transportation Research Record(1643): 116-124.

Haykin, S. (2003). Adaptive Filter Theory. Singapore, Pearson Education Pte. Ltd, Inidan Branch.

Herr, W. J. (1996). Highway Profile Measuring System. United States of America.

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

Ihs, A. (2004). The influence of road surface condition on traffic safety. 2nd European Pavement and Asset Management Conference. Berlin, Germany, VTI: 25-32.

ISO8608 (1995). Mechanical Vibration- Road Surface Profiles- Reporting of Measured Data. Geneva, International Organization of Standardization.

ISO (2004). 2631:2004. <u>Mechanical vibration and shock -- Evaluation of human exposure to whole-body vibration</u> International Organization for Standardization.

Karamihas, S. M. (2005). ACPA Profiler Repeatability Tests. T. R. Institute. Ann Arbor, MI, University of Michigan: 32p.

Karamihas, S. M. (2005). Critical Profiler Accuracy Requirements. United States: 130p.

Kelly, R. E., P. R. H. McConnell, et al. (1977). "The Gestalt Photomapping System." <u>Photogrammetric Engineering and Remote Sensing</u> 43: 1407-1417.

Kennedy, S., J. Hamilton, et al. (2006). <u>Architecture and system performance of SPAN-NovAtel's GPS/INS solution</u>, Piscataway, NJ 08855-1331, United States, Institute of Electrical and Electronics Engineers Inc.

Kern, J. and J. B. Ferris (2007). Developing ARIMA Road Profile Models with a Stationary Residual Process. <u>IFAC</u>.

Kern, J. V. and J. B. Ferris (2006). <u>Characterizing 2-D topographic mappings of roads</u>, Chicago, IL, United States, American Society of Mechanical Engineers, New York, NY 10016-5990, United States.

Kern, J. V. and J. B. Ferris (2006). <u>Preliminary results for model identification in characterizing 2-D topographic road profiles</u>. Modeling and Simulation for Military Applications, Kissimmee, FL, United States, SPIE - The International Society for Optical Engineering.

Kern, J. V. and J. B. Ferris (2006). <u>Preliminary results for model identification in characterizing 2-D topographic road profiles</u>, Kissimmee, FL, United states, SPIE.

Kern, J. V. and J. B. Ferris (2007). Development of a 3-D Vehicle-Terrain Measurement System Part I: Equipment Setup. <u>ISTVS: Innovations in Terrain and Vehicle Systems</u>. Fairbanks, AK, ISTVS.

Kern, J. V. and J. B. Ferris (2007). <u>Development of a 3D Vehicle-Terrain Measurement System Part I: Equipment Setup</u>. Proceedings of the Joint North America, Asia-Pacific ISTVS Conference, Fairbanks.

Kern, J. V., J. B. Ferris, et al. (2007). <u>Characterizing 2D road profiles using ARIMA modeling techniques</u>, Orlando, FL, United States, SPIE, Bellingham WA, WA 98227-0010, United States.

Khashei, M., M. Bijari, et al. (2008). "Improvment of Auto-Regressive Integrated Moving Average Models using Fuzzy Logic and Artifical Neural Networks (ANNs)." <u>Neurocomputing</u> 72: 956-967.

Kim, J. and S. Sukkarieh (2004). "Autonomous airborne navigation in unknown terrain environments." <u>Aerospace and Electronic Systems, IEEE Transactions on</u> 40(3): 1031-1045.

Kraus, K. (1997). <u>Photogrammetry</u>. Dummler, Bonn.

Kraus, K., C. Briese, et al. (2004). "Quality Measures for Digital Terrain Models." <u>International Archives of Photogammetry Remote Sensing and Spatial Information Sciences</u> 35 (Part B2): 113-118.

Kraus, K. and N. Pfeifer (1998). "Determination of terrain models in wooded areas with airborne laser scanner data." <u>Journal of Photogrammetry and Remote Sensing</u> 53: 193-203.

Kropáč, O. and P. Múčka (2004). "Non-standard longitudinal profiles of roads and indicators for their characterisation." International Journal of Vehicle Design 35(2-3): 149-72.

Kropáč, O. and P. Múčka (2005). "Be careful when using the International Roughness Index as an indicator of road unevenness." Journal of Sound and Vibration 287(4-5): 989-1003.

Kropáč, O. a. M., P. (2008). "Indicators of longitudinal unevenness of roads in the USA." International Journal of Vehicle Design 46(4): 393-415.

Kuchar, J. K. (2001). "Markov model of terrain for evaluation of ground proximity warning system thresholds." Journal of Guidance, Control, and Dynamics 24(3): 428-435.

Kutay, M. E., E. Weaver, et al. (2007). Development of Transfer Functions for Spectral Compatibility of Inertial Profilers Used in the LTPP Program. Transportation Research Board. Washington D.C., USA: 1-12.

Lanyi, G. E. and T. Roth (1988). "A comparison of mapped and measured total ionospheric electron content using Global Positioning System and beacon satellites." Radio Science 23(4): 483-492.

Lei, G. and R. J. Yang (2006). "On reliability-based optimisation methods for automotive structures." International Journal of Materials & Product Technology 25(1-3): 3-26.

Lenzen, A. and H. Waller (1997). "Identification using the algorithm of Singular Value Decomposition- An application to the realisation of dynamic systems and to fault detection and localisation." Mechanical Systems and Signal Properites 11(3): 441-457.

Letherwood, M. D., D. D. Gunter, et al. (2004). Spatial multibody modeling and vehicle dynamics analysis of advanced vehicle technologies, SPIE.

Li, L. and C. Sandu (2007). Modeling and Simulation of 2D ARMA Terrain Models for Vehicle Dynamics Applications. SAE International. Detroit, MI, Transactions Journal of Commercial Vehicles. SAE 2007-01-0138: 1-9.

Li, Z. (1994). "A comparative study of the accuracy of digital terrain models (DTMs) based on various data models." ISPRS Journal of Photogrammetry and Remote Sensing 49(1): 2-11.

Liu, C. and R. Herman (1999). "Road Profile, Vehicle Dynamics, and Ride Quality Rating." Journal of Transportation Engineering 125(2): 123-128.

Liu, F., N. Dembski, et al. (2008). A Kalman-filter-based multi-sensor terrain profile measurement system: principle, implementation and validation. Modeling and Simulation for Military Operations III, Orlando, FL, USA, SPIE.

Liu, F., N. Dembski, et al. (2008). A Kalman-filter based multi-sensor terrain profile measurement system: Principle, implementation and validation, Bellingham WA, WA 98227-0010, United States, SPIE.

Lopez, C. (1997). "Locating some types of random errors in Digital Terrain Models." International Journal of Geographical Information Science 11(7): 677-698.

Loucks, R. B. (1996). Chebyshev Polynomial Fit for Terrain Elevation. United States: 11p.

Lu, M.-W. and Y.-L. Lee (1996). Reliability based strength/fatigue design criteria. Proceedings of the Annual Reliability and Maintainability Symposium, Piscataway, NJ, USA, IEEE.

Makarand Datar, D. N. (2007). Technical Report on Virtual Protoyping of Ground Vehicles.

Mannucci, A. J., G. A. Hajj, et al. (2004). GPS-based remote sensing of the geospace environment: Horizontal and vertical structure of the ionosphere and plasmasphere. Proceedings of SPIE - The International Society for Optical Engineering, Honolulu, HI, United states, SPIE.

Maxwell, D. (1970). "Mathematical Surface Approximation of the Terrain." Transportation Research Record(319): 16-29.

Mehta, Y., R. Roque, et al. (2001). "Evaluation of road surface profiler and transverse profilograph for determination of rut depths." Transportation Research Record(1764): 157-163.

Miller, C. L. and R. A. Laflamme (1958). "The Digital Terrain Model-Theory and Application." Photogrammetric Engineering XXIV(3): 433.

Mocko, G. M. and S. J. Fenves (2003). A Survey of Design- Analysis Integration Issues, National Insitute of Standards and Technology.

Moore, I. D., R. B. Grayson, et al. (1991). "Digital Terrain Modelling: A Review of hydrological geomorphological and biological applications." Hydrological Processes: Digital Terrain Modelling in Hydrology 5(1): 3-30.

Mueller, T., J. B. Ferris, et al. (2009). Identifying Vehicle Model Parameters using Measured Terrain Data, Detroit, MI, Submitted to the SAE World Congress.

Offrell, P. and L. Sjogren (2004). Crack Measures and Reference Systems for a Harmonised Crack Data Collection Using Automatic Systems. 2nd European Pavement and Asset Management Conference. Berlin, Germany, VTI: 14-24.

Papagiannakis, A. T. and B. Raveendran (1998). "International Standards Organization-Compatible Index for Pavement Roughness." Transportation Research Record: Journal of the Transportation Research Board 1643(1): 6.

Perera, R. W. and S. D. Kohn (2002). Issues in Pavement Smoothness: A Summary Report. N. Project: 1-200.

Petrie, G. and T. J. M. Kennie (1987). "Terrain modelling in surveying and civil engineering." Computer-Aided Design 19(4): 171-187.

Pfeifer, N., A. Kostli, et al. (1998). "Interpolation and filtering of laser scanner data - implementation and first results." International Archives of Photogammetry Remote Sensing and Spatial Information Sciences XXXIII(Part 3/1): 153.

Pfeifer, N., P. Stadler, et al. (2001). Derivation of digital terrain models in the SCOP++ Environment. OEEPE Workshop on Airborne Laserscanning and Interferometric SAR for Digital Elevation Models. Stockholm: 1-13.

Prem, H. (1987). A Laser-Based Highway-Speed Road Profile Measuring System. 10th International Association for Vehicle System Dynamics (IAVSD 1987), Symposium on The Dynamics of Vehicles on Roads and on Tracks Prague, Czechoslovakia.

Quinn, P., K. Beven, et al. (1991). "The prediction of hillslope flow paths for distributed hydrological modelling using digital terrain models." Hydrological Processes: Digital Terrain Modelling in Hydrology 5(1): 59-79.

Rauh, J. and M. Mossner-Beigel (2008). "Tyre simulation challenges." Vehicle System Dynamics 46(SUPPL.1): 49-62.

Rose, S. M. (2001). The Effect of Digital Elevation Model Resolution on Wave Propagation Predictions at 24Ghz. Geography. Blacksburg, VA, Virginia Tech. Master of Science: 53.

Rouillard, V. (2004). "Using predicted ride quality to characterise pavement roughness." International Journal of Vehicle Design 36(2-3): 16.

Rouillard, V., B. Bruscella, et al. (2000). "Classification of Road Surface Profiles." Journal of Transportation Engineering 126(1): 41-45.

Rouillard, V., M. A. Sek, et al. (2001). "Simulation of Road Surface Profiles." Journal of Transportation Engineering 127(3): 247-253.

Sardon, E., A. Rius, et al. (1994). "Estimation of the transmitter and receiver differential biases and the ionospheric total electron content from Global Positioning System observations." Radio Science 29(3): 577-586.

Sardon, E. and N. Zarraoa (1997). "Estimation of total electron content using GPS data: How stable are differential satellite and receiver instrumental biases?" Radio Science 32(5): 1899-1910.

Sarma, A. D. S. R., G.; Somasekhar Rao, P.V.D.; Ramalingam, K. (2008). "GPS Satellite and Receiver Instrumental Biases Estimation using SVD Algorithm." IEEE Transactions on Aerospace and Electronic Systems 44(4): 1560-1566.

Sayers, M. and T. G. W. Paterson (1986). Guidelines for Conducting and Calibrating Road Roughness Measurements. T. W. Bank. Washington D.C., UMTRI: 1-96.

Sayers, M. W. (1989). "Two quarter-car models for defining road roughness. IRI and HRI." Transportation Research Record(n 1215): 165-172.

Sayers, M. W. (1995). "On the calculation of international roughness index from longitudinal road profile." Transportation Research Record(1501): 1-12.

Sayers, M. W. and M. S. Karamihas (1998). The Little Book of Profiling. T. R. Institute. Ann Arbor, MI, The Regent of the University of Michigan.

Sayers, M. W. and S. M. Karamihas (1996). "Estimation of rideability by analyzing longitudinal road profile." Transportation Research Record(1536): 110-116.

Segu, W. P. (1985). "Terrain approximation by fixed grid polynomial." Photogrammetric Record 11(65): 581-91.

Shoop, S. A. (2001). Finite Element Modeling of Tire-Terrain Interaction. United States: 69p.

Sinha, R., V.-C. Liang, et al. (2001). "Modeling and Simulation Methods for Design of Engineering Systems." Journal of Computing and Information Science in Engineering 1: 84-91.

Sjogren, L. and T. Lundberg (2004). Designing an up to date rut depth monitoring profilometer, requirements and limitations. 2nd European Pavement and Asset Management Conference. Berlin, Germany, VTI: 1-13.

Smith, H. (2009). "Incorporating tire envelopment and bridging algorithms with 2D scanning laser technology to improve the validity of 2D terrain." IJVSMT.

Smith, H. and J. B. Ferris (2008). Post Processing Techniques to Correct Inertial Errors in High Fidelity Terrain Topology Measurements. International Society for Terrain-Vehicle Systems. Torino, Italy.

Smith, H. and J. B. Ferris (2008). Post Processing Techniques to Correct Inertial Measurements in High Fidelity Terrain Topology Measurements, Turin, Italy, International Conference of the ISTVS.

Smith, H. and J. B. Ferris (2009). Excitation Event Design and Accuracy Verification Procedure for High Fidelity Terrain Measurement Systems. Submitted to the SPIE, Orlando, FL.

Smith, H. F., John B. (2009). Calibration Surface Design and Accuracy Verification Procedure for Terrain Measurement Systems. ASTM.

Smith, H. F., John B. (2009). "Techniques for averting and correction inertial errors in 2D terrain topology measurements." Journal of Terramechanics.

Spangler, E. B. (1962). Servo-Seismic Method of Measuring Road Profiles, Highway Research Board Bulletin 328: 33-51.

Spangler, E. B. and W. J. Kelly (1966). GMR Road Profilometer-A Method for Measuring Road Profile. Washington D.C., Highway Research Board. 121: 27-54.

Spangler, E. B. and W. J. Kelly (1966). GMR Road Profilometer - A Method for Measuring Road Profile. Highway Research Record, Highway Research Board: 27-54.

Stadterman, T. J., W. Connon, et al. (2003). "Dynamic modeling and durability analysis from the ground up." Journal of the IEST 46: 128-34.

Stein, A., F. v. d. Meer, et al., Eds. (2002). Spatial Statistics for Remote Sensing. Dordrecht, Kluwer Academic Publishers.

Stein, M. L. (1999). Interpolation of Spatial Data: Some Theory for Kriging. New York, Springer.

Sun, H., J. Fu, et al. (2008). "Analysis of the Kalman Filter with Different INS Error Models for GPS/INS Integration in Aerial Remot Sensing Applications." The International Archives of the Phogrammetry, Remote Sensting and Spatial Information Sciences XXXVII(Part B5): 883-890.

Sylvester, J. J. On the reduction of a bilinear quantic of the nth order to the form of a sum of n products by a double orthogonal subsitution. Messenger of Mathematics: 1-5.

VIRES. (October 2008, November 12, 2009). "OpenCRG." Retrieved November 18, 2009, from http://www.opencrg.org/.

Wackernagel, H. (2010). Multivariate Geostatistics. Berlin, Springer.

Wagner, S. and J. B. Ferris (2010). "Developing Compact Characterizations of Terrain Topology Profiles." Submitted to Journal of Dynamic Systems Measurement and Control (DS-10-1070).

Wagner, S. and J. B. Ferris (2010). "Residual Analysis of Autoregressive Models of Terrain Topology." Submitted to Journal of Dynamic Systems Measurement and Control.

Wagner, S. M. and J. B. Ferris (2007). A polynomial chaos approach to ARIMA modeling and terrain characterization, Bellingham, WA SPIE.

Wagner, S. M. and J. B. Ferris (2008). Reduced order ARIMA models of 2-D terrain profiles using singular value decomposition, New York, NY 10016-5990, United States, American Society of Mechanical Engineers.

Wagner, S. M., J. V. Kern, et al. (2007). Development of a 3-D Vehicle-Terrain Measurement System Part II: Signal Processing and Validation. ISTVS: Innovations in Terrain and Vehicle Systems. Fairbanks, AK, ISTVS.

Wagner, S. M., J. V. Kern, et al. (2007). Development of a 3D Vehicle-Terrain Measurement System Part II: Signal Processing and Validation, Fairbanks,

Alaska, Proceedings of the Joint North America, Asia-Pacific ISTVS Conference and Annual Meeting of the Japanese Society for Terramechanics.

Walker, R. S. and E. Becker (2006). Collecting Stop and Go Inertial Profile Measurements. United States: 56p.

Wang, H. (2006). Road Profiler Performance Evaluation and Accuracy Criteria Analysis. Civil and Environmental Engineering. Blacksburg, VA, Virginia Tech. Masters of Science: 73.

Wilson, J. P. and J. C. Gallant (2000). Terrain Analysis: Principles and Applications, John Wiley and Sons.

Yoon, S. (2003). "A study on terrain-surface modeling and searching algorithms for real-time simulation of off-road vehicles." Vehicle System Dynamics 39(5): 353-63.

Nomenclature

H	Hilbert Space spanned by sets of vectors $\{z_{i,k}\}$, $\{s_i\}$, and $\{e_{i,k}\}$
G	Subspace of **H**, in which global events are defined
G^c	Complement of subspace **G**, in which local events are defined
$z_{i,k}$	i^{th} vector component of the kth realization of the terrain surface height
s_i	i^{th} vector component of the true surface height
$\hat{s}_i$	i^{th} vector component of the estimated true surface height
$\hat{s}_{trunc(i)}$	i^{th} vector component of the truncated true surface
$e_{i,k}$	Error in realization $z_{i,k}$ with respect to s_i
$n_{i,k}$	Noise component of the ith error vector of the kth realization
b_l	Orthonormal empirical basis vector in the principle direction indexed by l
p_l	Orthonormal analytical basis vector in the principle direction indexed by l
$\sigma_{i,l}$	Estimated true surface projected onto basis vector p_l or b_l
$\delta_{i,k,l}$	Projection of $e_{i,k}$ onto b_l representing the contribution of the measurement error in the lth global direction
$\alpha_{i,k,l}$	Difference in the ith transverse profile, k^{th} realization in the l direction from γ_l
γ_l	Ensemble average drift in the l direction across all realizations
$\beta_{i,k,l}$	Projection of $z_{i,k}$ onto the basis vector b_l
i	Index for discrete longitudinal locations along the path coordinate u, where $i \in \{1,2,\ldots,m\}$
j	Index for transverse locations along the path coordinate v, where $j \in \{1,2,\ldots,n\}$
k, h	Indices for the realization (measurement) $[1, 2, \ldots r]$
l	Index for the principle directions of **H**, where $l \in \{1,2, \ldots, q\}$
q	Number of principle directions of **H** (i.e., the dimension of **G**)
$\langle \bullet, \bullet \rangle$	Inner-product operator on two vectors
$\langle \bullet, \bullet \rangle_w$	Weighted inner-product operator on two vectors

$E[\bullet]$	Expectation operator
$\hat{E}[\bullet]$	Estimated expectation operator (some ensemble averaging technique, such as mean, median, 50% trimmed mean, etc.)
m	The number of discrete longitudinal locations along the terrain surface
r	the number of realizations (measurements) used to estimate the true terrain surface
t	truncated set of basis vectors
a	Number of polynomial terms (polynomial order is a-1)
(x, y)	Global coordinate system
(u, v)	Path-centered coordinate system
$\boldsymbol{u}$	Vector of discrete longitudinal locations along path coordinate u
$\boldsymbol{v}$	Vector of discrete transverse locations at which the terrain heights are defined either through measurement or spatial interpolation
RMS_H	Root Mean Square of terrain height
D_H	Single measure of the combined variance of the random component of the road profile with the harmonic component of the road profile
Z_n	Stochastic process
n	Process index
S	State space of the stochastic process
z_n	Elements of S
$G_d(n)$	The roughness level of the displacement PSD
C	The general roughness parameter (PSD)
w	Dimensionless parameter called waviness (PSD)
λ	Wavelength of the road (PSD)
$P(\ )$	Probability
$s_1,\ s_2,\ ...$	States of State Space of stochastic process
$P^n(\ )$	n-step transition function
p	Order of the autoregressive model
d	The number of times the data were differenced
q	The order of the moving average model
φ	The autoregressive coefficient

177

a_i	The distribution of the uncorrelated residuals
∇^d	Backward difference operator
$R_{PRS}(n)$	Parametric Road Spectrum
$R_{SSR}(n)$	Shifted Spatial Range Spectrum
μ	Mean
σ^2	Variance
m	Total vehicle mass (kg)
m_s	Sprung mass (kg)
m_{us}	Unsprung mass of each wheel assembly (kg)
$k_{s,f}$	Front suspension stiffness (N/m)
$k_{s,r}$	Rear suspension stiffness (N/m)
k_t	Tire stiffness (N/m)
c_f	Damping coefficient front suspension N/(m/s)
c_r	Damping coefficient rear suspension N/(m/s)
a	Distance front axle to center of mass (m)
b	Distance rear axle to center of mass (m)
t_f	Track width front axle (m)
t_r	Track width rear axle (m)
$z(t)$	Measured suspension travel over time
L	Length of the profile [m]
$\dot{z}_s$	Motion of sprung mass [m/s]
$\dot{z}_u$	Motion of unsprung mass [m/s]
V	Velocity of at which the simulation was executed
IRI	International Roughness Index
RN	Ride Number
$RIDE$	Roughness Index for Driving Expenditure
SVD	Singular Value Decomposition
AR	Autoregressive Model
f	Sprung mass acceleration of reference vehicle
$y_i(t)$	Road input to full-car model, i indicates the corner of the vehicle

FRI	Full-car roughness index
PSD	Power Spectral Density
RCR	Ride Condition Rating
p_i	Weight function for robust estimation (linear prediction)
a, b	Steepness of weight function (linear prediction)
g	Shift value (linear prediction)
w	Ceiling elevation value that is user defined (linear prediction)
v_i	Surface residuals (linear prediction)
P(z)	Gaussian distribution
$\overline{n_{i,j}}$	Mean of the terrain surface for each grid location
$\overline{n}$	Scalar representation of the mean of the terrain surface
$\sigma_{i,j}$	Standard deviation of the terrain surface for each grid location
σ	Scalar representation of the standard deviation of the terrain surface
(ξ, η)	Node based coordinate system
$N_i(\xi)$	Shape function in (ξ)
$N_i(\eta)$	Shape function in (η)
$N_i(\xi, \eta)$	Shape function in (ξ, η)
$u(\xi, \eta)$	Displacement field
λ	Parameter to tune Gegenbauer polynomials
$C_n^{(\lambda)}(x)$	Gegenbauer polynomials
$T_k(x)$	Chebyshev polynomials
$f(x)$	Fourier series basis functions
a_0	Fourier series coefficient
a_n	Fourier series coefficient
b_n	Fourier series coefficient
$P(v)$	Continuous polynomial
v	Continuous coordinate on which the continuous polynomial is defined
$\boldsymbol{p}(\boldsymbol{v}, \lambda)$	Discretized polynomial in vector form as a function of λ and $\boldsymbol{v}$
$\boldsymbol{V}(\boldsymbol{v})$	Vandermonde matrix, *n*x*a* for which *n=a*
$\boldsymbol{c}(\lambda)$	Gegenbauer coefficients in vector form

C	Matrix of coefficients
W	Weighting matrix
$P(v, \lambda)$	Discretized polynomial in matrix form as a function of λ and v
$C(\lambda)$	Gegenbauer coefficients in matrix form
R	Weighted grammian
θ_l	Coefficients pertaining to the magnitude of the contribution of each analytic basis vector to the true terrain surface
$RMSE$	Root mean square error
$v_{measured}$	Original transverse measurement locations
$\overline{v_{measured}}$	Mean of the original transverse measurement locations
t	Track width of typical passenger vehicle
$\rho(\lambda)$	Distance between the two local minima for the fifth analytic basis vector

Appendix A - VTMS Detailed Description

The VTMS is equipped with a Phoenix Scientific PPS scanning laser and a Novatel SPAN INS. The scanning laser acquires data at 1kHz with an accuracy of +/- 0.5mm in the horizontal plane and +/- 0.5mm in elevation. The INS combines an inertial measurement unit (IMU) which measures three translational accelerations and three rotational velocities at 100 Hz with a differential GPS acquiring data at 5 Hz. The differential GPS and IMU signals are then combined through Kalman filtering.

The horizontal precision of the complete system is defined in terms of the standard deviation in the height measurement vs. the horizontal distance the vehicle has traveled between repeated measurements, as shown in Figure 66. This can be thought of as the precision as a function of wavelength. The INS is capable of a horizontal precision of 1 mm for short distances traveled; the limiting factor is the laser spot size. Precision levels of 10 mm are possible for distances of up to 10 m; this is limited by the drift in the uncorrected IMU data (1deg/hr). For distances greater than 10 m, the precision is limited by the DGPS, which has a precision of 10mm + 1ppm (e.g., 20 mm at 10 km traveled); the DGPS is affected by the number of satellites and atmospheric interference.
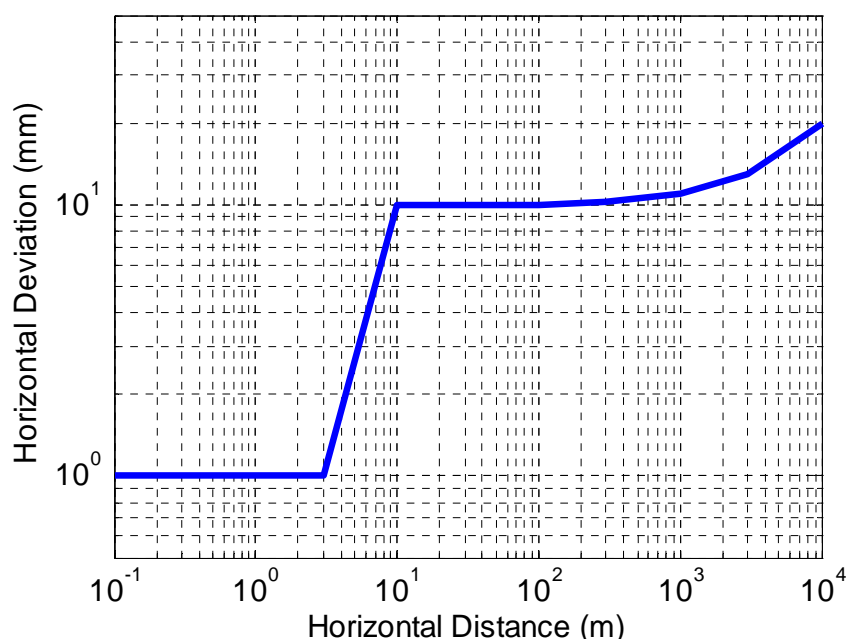
Figure 66. Horizontal Precision of VTMS.

The vertical accuracy of the system was assessed by exciting the vehicle over large 70 mm wooden cleats; the vehicle was driven at a speed such that its resonant frequencies were excited as described by Smith (Smith 2009). A steel calibration surface, as shown in Figure 67 (located between the wooden excitation event) was measured during this severe excitation, resulting in an RMS error of 2 mm.

Figure 67. Sample calibration surface for inertial profilers.

Appendix B - Parameter Values for Calculating IRI

Table 3: Golden Car Parameters

Parameter	Description
C	C_s/M_s=6.0 : Normalized damping coefficient
K_1	K_t/M_s=653 : Normalized tire stiffness
K_2	K_s/M_s=63.3 : Normalized spring stiffness
μ	M_u/M_s=0.15 : Normalized mass ratio
L_B	250mm, Moving average base length
V	V=80 km/hr : Forward speed

Table 4: Ride Number Quarter Car Parameters

Parameter	Description
C	C_s/M_s=17 : Normalized damping coefficient
K_1	K_t/M_s=5120 : Normalized tire stiffness
K_2	K_s/M_s=390 : Normalized spring stiffness
μ	M_u/M_s=0.036 : Normalized mass ratio
V	V=80 km/hr : Forward speed

Appendix C - Gegenbauer Polynomials

Gegenbauer polynomials are solutions to the Gegenbauer differential equation for integer n. They are proportional, or depending on the normalization, equal to the ultraspherical polynomials $P_n^\lambda(x)$. The Gegenbauer polynomials are extensions of Legendre Polynomials. The generating function for these polynomials are then given by

$$\frac{1}{(1 - 2xt + t^2)^\lambda} = \sum_{n=0}^{\infty} C_n^{(\lambda)}(x)t^n$$

The first few Gegenbauer Polynomials are then

$$C_0^{(\lambda)}(x) = 1$$

$$C_1^{(\lambda)}(x) = 2\lambda x$$

$$C_2^{(\lambda)}(x) = -\lambda + 2\lambda(1 + \lambda)x^2$$

$$C_3^{(\lambda)}(x) = -2\lambda(1 + \lambda)x + 4/3\lambda(1 + \lambda)(2 + \lambda)x^3$$

A recurrence relation is presented as

$$nC_n^{(\lambda)}(x) = 2(n + \lambda - 1)xC_{n-1}^{(\lambda)}(x) - (n + 2\lambda - 2)C_{n-2}^{(\lambda)}(x), \ \forall n$$
$$= 2,3 \ ...$$

Appendix D - Legendre Polynomials

Legendre polynomials are solutions to Legendre differential equations. The Legendre polynomial can be defined by the contour integral over the range [-1,1], where n is an integer and the contour integral encloses the origin and is traversed in a counterclockwise direction (Arfken 1985).

$$P_n(z) = \frac{1}{2\pi i} \oint (1 - 2tz + t^2)^{-1/2} t^{-n-1} dt$$

Equation 90

The first few Legendre Polynomials are

$$P_0(x) = 1$$
Equation 91

$$P_1(x) = x$$
Equation 92

$$P_2(x) = \frac{1}{2}(3x^2 - 1)$$
Equation 93

$$P_3(x) = \frac{1}{2}(5x^3 - 3x)$$
Equation 94

$$P_4(x) = \frac{1}{8}(35x^4 - 30x^2 + 3)$$
Equation 95

$$P_5(x) = \frac{1}{8}(63x^5 - 70x^3 + 15x)$$
Equation 96

$$P_6(x) = \frac{1}{16}(231x^6 - 315x^4 + 105x^2 - 5)$$
Equation 97

The Legendre Polynomials can also be generated using Gram-Schmidt Orthonormalization on the open interval (-1, 1) with the weighting function 1. Normalizing so that $P_n(1)=1$ yields the expected Legendre Polynomials.

$$P_0(x) = 1$$
Equation 98

$$P_1(x) = \left[x - \frac{\int_{-1}^{1} x\, dx}{\int_{-1}^{1} dx} \right] * 1 = x$$
Equation 99

$$P_2(x) = x \left[x - \frac{\int_{-1}^{1} x^3 dx}{\int_{-1}^{1} x^2 dx} \right] - \left[\frac{\int_{-1}^{1} x^2 dx}{\int_{-1}^{1} dx} \right] * 1 = x^2 - \frac{1}{3}$$
Equation 100

$$P_3(x) = \left[x - \frac{\int_{-1}^{1} x \left(x^2 - \frac{1}{3} \right)^2 dx}{\int_{-1}^{1} \left(x^2 - \frac{1}{3} \right)^2 dx} \right] \left(x^2 - \frac{1}{3} \right) - \left[\frac{\int_{-1}^{1} (x^2 - \frac{1}{3})^2 dx}{\int_{-1}^{1} x^2 dx} \right]$$

$$* x = x^3 - \frac{3}{5}$$

The Legendre Polynomials are a special case of ultraspherical functions with α=1/2, a special case of the Jacobi Polynomials $P_n^{(\alpha,\beta)}$ with α=β=0. A generating function for the Legendre Polynomial is written as

$$g(t, x) = (1 - 2xt + t^2) = \sum_{n=0}^{\infty} P_n(x) t^n$$

The recurrence form of the Legendre Polynomial is written as

$$(l + 1)P_{l+1}(x) - (2l + 1)xP_l(x) + lP_{l-1}(x) = 0$$

Appendix E - Chebyshev Polynomials

Chebyshev polynomials are solutions to the Chebyshev differential equation. They are typically used as an approximation to a least squares fit and are a special case of the Gegenbauer polynomial when α=0. Chebyshev polynomials of the first kind are defined by the contour integral over the range [-1, 1].

$$T_n(z) = \frac{1}{4\pi i} \oint \frac{(1 - t^2)t^{-n-1}}{(1 - 2tz + t^2)} dt$$

<div align="right">Equation 104</div>

The generating function for these polynomials exists for |x|≤1 and |t|<1 are then given by

$$g_1(t,x) = \frac{1 - t^2}{1 - 2xt + t^2} = T_0(x) + 2\sum_{n=1}^{\infty} T_n(x)t^n$$

$$g_2(t,x) = \frac{1 - xt}{1 - 2xt + t^2} = \sum_{n=0}^{\infty} T_n(x)t^n$$

<div align="right">Equation 105</div>

A direct representation of the generating function is given by

$$T_n(x) = \frac{1}{2}x^2 \left[\left(\sqrt{1 - \frac{1}{z^2}} + 1 \right)^n + \left(\sqrt{1 - \frac{1}{z^2}} \right)^n \right]$$

<div align="right">Equation 106</div>

The first few Chebyshev Polynomials are then

$$T_0(x) = 1$$

<div align="right">Equation 107</div>

$$T_1(x) = x$$

<div align="right">Equation 108</div>

$$T_2(x) = 2x^2 - 1$$

<div align="right">Equation 109</div>

$$T_3(x) = 4x^3 - 3x$$

<div align="right">Equation 110</div>

$$T_4(x) = 8x^4 - 8x^2 + 1$$

<div align="right">Equation 111</div>

A recurrence relation is presented that is true after setting T0(x)=1 and T1(x)=x

$$T_k(x) = 2xT_{k-1}(x) - T_{k-2}(x) \quad \forall k = 2,3, \dots$$

<div align="right">Equation 112</div>

Appendix F - Fourier Series Polynomials

Fourier series polynomials take on a similar form to that of power series or Taylor series, except instead of an infinite series of powers, Fourier series are based on an infinite series of sines and cosines. Fourier series can be expressed in the following form

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos \frac{n\pi}{l} x + \sum_{n=1}^{\infty} b_n \sin \frac{n\pi}{l} x$$

Equation 113

Let f(x) and f'(x) be periodic functions having period 2*l*, where f(x) and f'(x) are piecewise continuous on -∞ < x < ∞. Let the Fourier coefficients be defined by

$$a_n = \frac{1}{l} \int_{-l}^{l} f(x) \cos \frac{n\pi}{l} dx, \quad n = 0,1 \dots$$

Equation 114

$$b_n = \frac{1}{l} \int_{-l}^{l} f(x) \sin \frac{n\pi}{l} dx, \quad n = 1,2 \dots$$

Equation 115

Appendix G - Inner Product Properties

Inner product properties

- Conjugate Symmetry

$$\langle x, y \rangle = \langle \overline{y, x} \rangle = \langle y, x \rangle \qquad \text{Equation 116}$$

- Bilinearity

$$\langle \alpha x, y \rangle = \alpha \langle y, x \rangle \qquad \text{Equation 117}$$

- Bilinearity

$$\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle \qquad \text{Equation 118}$$

- Non-negativity (Positive Definiteness)

$$\langle x, x \rangle > 0 \; if \; x \neq 0 \qquad \text{Equation 119}$$

$$\langle x, x \rangle = 0 \; if \; x = 0$$

Appendix H - Weighted Inner Product

Weighted Inner Product

- For a finite dimensional vector space, a weighted inner product can be obtained by inserting a Hermitian weighting matrix W between the elements. The concept of orthogonality is defined with respect to the particular inner product used: changing the inner product may change the orthogonality relationship between vectors.

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle_W = \boldsymbol{y}^H W \boldsymbol{x} \qquad\qquad \text{Equation 120}$$

- In order for the weighted inner product to be used to define a norm, as in

$$\|\boldsymbol{x}\|_W^2 = \langle \boldsymbol{x}, \boldsymbol{x} \rangle_W = \boldsymbol{x}^H W \boldsymbol{x} \qquad\qquad \text{Equation 121}$$

- It is necessary that $\boldsymbol{x}^H W \boldsymbol{x} > 0$ for all x not equal to 0. A matrix W with this property is said to be positive definite.

- A Hermitian matrix is a square matrix which is equal to its own conjugate transpose. A matrix that has only real entries is Hermitian if and only if it is a symmetric matrix with respect to the main diagonal. A real and symmetric matrix is simply a special case of a Hermitian matrix.

- Weighting can also be applied to integral inner products. If there is some function w(t)>=0 over [a,b], then an inner product can be defined as

$$\langle f, g \rangle_w = \int_b^a w(t) f(t) g(t) dt \qquad\qquad \text{Equation 122}$$

- The weighting can be used to place more emphasis on certain parts of the function

Appendix I – Hilbert Space

Hilbert Space: A finite dimensional vector space where the norm is the induced norm and an inner-product exists.

- Norm is the induced norm for a Hilbert space. The length of vector **x**.

$$\|x\| = \langle x, x \rangle^{\frac{1}{2}}$$
Equation 123

- Form basis vectors so that

$$\langle x_i, x_j \rangle = \delta(i - j) = \begin{cases} 1 \; for \; i = j \\ 0 \; for \; i \neq j \end{cases}$$
Equation 124

- Angle between vectors **x** and **y** is consistent with dot product

$$\langle x, y \rangle = \|x\| \|y\| cos\theta_{xy}$$
Equation 125

To be a basis vector, the vectors must be orthogonal. Normalize for mathematical convenience.

Appendix J – INS Drift Sample Matlab Code

This Appendix provides a sample of Matlab code used to develop the INS drift compensation method.

```matlab
function arrayData = remove_ins_drift_hmc_2010_09_07(directory)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% REMOVE INS DRIFT from crg structs
%
% Title: remove_ins_drift
% Author:JBF, RMB, HMC
% Date: 2010 09 07
%
% DESCRIPTION
%
% Assumptions
% 1) directory contains subdirectories that in turn contain .mat files
% which are stuctures, D, with a D.z matrix that represents the best
% estimate of the surface for that run
%
% Inputs
% 1) directory containing a set of subdirectories, each subdirectory
containing an .mat file from
% an individial measurements/runs/realization of a particular terrain.
% That is, the directory only contains information about a single
terrain
% surface, not different terrain surfaces
%
% Variables
% dirName: the name of the diirectory in which the subdirectories
exist,
%          without the superdirectories appended
% Outputs
%
%
% Tasks
% 1) Define bigZ from directory
% 2) Define a 'large' number of basis vectors
% 3) Calculate the projection of the measured vectors onto the basis
% vectors (beta (i,k,l))
% 4) Sum projections of measured vectors onto basis vectors
% 5) Calculate the estimated True Surface
% 6) Calculate the error surfaces
% 7) Calculate the drift surfaces
% 8) Evaluate noise surface to be sure it is zero-mean
% 9) Calculate the drift-free surfaces and over-write arrayData(k).z
% 10) create additional structure in the arrayData array with True
Surface
%
% Validation sets
% 1) directory='\\172.16.56.51\datashare\Unsecured\Data\MnRoad 2009 10
20\Mnroad_20091019_Section1'
```

```matlab
% 2) directory='\\172.16.56.51\datashare\Unsecured\Data\MnRoad 2009 10
20\Mnroad_Section5'
% 3) directory='\\172.16.56.51\datashare\Unsecured\Data\Clean Data - DO
NOT MODIFY\Mnroad_20091019_Section1'
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%

% maxnBasis is the maximum number of basis vectors and is being
% defined as a parameter. If the length of the vectors (n) is less than
% maxnBasis, then maxnBasis is reduced later in the program
maxnBasis=20;

% Get the name of the main directory
remainder = directory;
while true
    [dirName, remainder] = strtok(remainder, '\');
    if isempty(remainder),  break;  end
end

% 1) Define bigZ from the directory
% define a 3D matrix with each surface being added as a layer to the 3D
% matrix bigZ.  This is done to simplify further processing of data
% (without the meta info carried along)
% bigZ is the concatenation of the crg surfaces
%   bigZ(i,j,k) then
%   i: index of transverse profile, length of longitudinal profile
("row"), i=[1,2,..., m]
%   j: index of longitudinal profile, length of transverse profile
("col"), j=[1,2,..., n]
%   k: index of realization (measurement), k=[1,2,..., r]

arrayMatFilename = get_matFilename(directory);  % this is an array of
structures containing a field called .matFilename, this is not an array
of strings (or all the filenames would have to be the same length?)
r=length(arrayMatFilename);


% 1b) load first mat file
load(arrayMatFilename(1).matFilename);    % There should now be a
struct called 'D' in the workspace

% Account for different versons of the mat file...
if exist('Surface','var'), D=Surface; clear Surface; end

% Determine 'best' reported value of z (median best, mean next best,
'z',
% third best...
D.z = D.weightCdf(:,:,5);

[m,n]=size(D.z);

if ~D.suppressOutput,
    disp(' ');
    disp(['Processing ' dirName]);
    disp(' ');
```

193

```matlab
    disp(['Number of longitudinal points (length of profiles, m): ',
num2str(m)]);
    disp(['Number of transverse points (size of vector space, n): ',
num2str(n)]);
    disp(['Number of realizations (measurements runs taken, r):   ',
num2str(r)]);
    figure(1); clf(1); surf(D.z); title(['Heights for ',D.filename]);
    figure(2); clf(2); contour(D.z); title(['Heights for
',D.filename]);
end


bigZ=zeros(m,n,r);


bigZ(:,:,1) = D.z;        % take best reported value of z


%arrayData(1:r)=struct;
arrayData(1:r)=D;        % pre-allocate memory so that all r structures
start off looking like the first one
clear D;


% 1b) load subsequent mat files
for k = 2:r,
    load(arrayMatFilename(k).matFilename);     % There sould now be a
struct called 'D' in the workspace

    % Account for different versons of the mat file...
    if exist('Surface','var'), D=Surface; clear Surface; end

    % Determine 'best' reported value of z
    D.z = D.weightCdf(:,:,5);

    arrayData(k)=D;
    bigZ(:,:,k) = D.z;        % take best z value
    %bigZ(:,:,k)=detrend(bigZ(:,:,k));
    clear D;
end




% 2) Define the basis vectors
% calculate the first maxnBasis basis vectors, they will not all be
used in the
% final analysis, but they are available.  There can not be more basis
% vectors that the size of the vector space itself, which is n, the
length
% of each vector
maxnBasis = min(maxnBasis, n);
[singularValues, basis] = get_basis_vectors(bigZ, maxnBasis);
singularValues = diag(singularValues);
% if ~arrayData(1).suppressOutput,
%     disp(['Singular Values: ', num2str(singularValues')]);
%     figure(3); clf(3);
%     stem(singularValues);
%     xlabel('Singular Value index (l)');
%     ylabel('Singular Value');
```

```matlab
%       title(['Singular Values for ', dirName]);
%       figure(4); clf(4);
%       plot(basis(:,1:min(5,maxnBasis)),'-');
%       hold on;
%       plot(basis(:,1:min(5,maxnBasis)),'o');
%       xlabel('Transverse Location Index (j)');
%       ylabel('Basis Vector Height');
%       title(['Basis Vectors for ', dirName]);
% end

% 3) Calculate the projection of the measured vectors onto all the
basis
% vectors (beta (i,k,l))
beta = zeros(m,r,maxnBasis);
for l=1:maxnBasis,
    for k=1:r,
        for i=1:m,
            beta(i,k,l) = bigZ(i,:,k)*basis(:,l);
        end
    end
end

% 4) Loop on increasing number of basis vectors being USED, nBasis,
until
% the noise is is zero-mean with a standard deviation less than 1/2 a
% milimeter (therefore the standard deviation is 2 standard deviations
away from the resolution of the laser).

%Initialize variables for loop to run
StdTotal=1;
MeanTotal=1;
nBasis = -1;           % initialize nBasis to zero, it is immediately
incremented inside the while loop
maxNoise = 1;
minNoise = -1;
%while ( (StdTotal>0.0005 || MeanTotal > 0.0001 || maxNoise > 0.010 ||
minNoise < -0.010) && nBasis<maxnBasis ),
 while ( (StdTotal>0.0005 || MeanTotal > 0.0001 ) && nBasis<maxnBasis
),
    nBasis = nBasis + 1;

    % 4a) Sum projections of measured vectors onto basis vectors
    sumProjection = zeros(m,n,r);
    for i=1:m,
        for k=1:r,
            for l=1:nBasis,
                sumProjection(i,:,k) = sumProjection(i,:,k) +
beta(i,k,l)*basis(:,l)';
            end
        end
    end

    bigZMinusSumProjection = bigZ-sumProjection;
```

```matlab
    % 4b) Calculate the estimated True Surface
    % [expectedS] = get_expectation_20090515(s, dim, method)
    trueSurface = get_expectation_20090515(bigZMinusSumProjection, 3, ...
2) + ...
        get_expectation_20090515(sumProjection, 3, 2);

    % 4c) Calculate the error surfaces
    error = zeros(m,n,r);
    for k=1:r,
        error(:,:,k)=bigZ(:,:,k)-trueSurface;
    end

    % 4d) Calculate the drift surfaces
    drift = zeros(m,n,r);
    for i=1:m,
        for k=1:r,
            for l=1:nBasis,
                drift(i,:,k) = drift(i,:,k) +
(error(i,:,k)*basis(:,l))*basis(:,l)';
            end
        end
    end

    %Calculate the projection of the True Surface onto the principle
    %directions
  sigma = zeros(m,nBasis);
    for i=1:m,
        for l=1:nBasis,
                sigma(i,l)=trueSurface(i,:)*basis(:,l);
        end
    end

    %Calculate the truncated surface
    sTrunc=zeros(m,n);
    for i=1:m,
        for l=1:nBasis,
            sTrunc(i,:)=sigma(i,l)*basis(:,l)';
        end
    end


    % 4e) The noise surface is then the error surface - the drift
surface
    noise = error - drift;

    % 4f) Evaluating the noise surface to see if it is a zero-mean
process
    % with a standard deviation less than 0.5mm

    [myMean, myStd, StdTotal, MeanTotal, VarTotal, maxNoise,
minNoise]=CheckNoise(noise);

    %xtrans=linspace(0.5,1.8,n);
    xtrans=arrayData(1).v;
    %xlong=linspace(1,151.2,m);
```

```matlab
    xlong=linspace(0,arrayData(1).nu*arrayData(1).uinc,m)';

    % 4g) Plotting every tenth longitudinal profiles
    for i=1:1:min(9,floor(m/300)),          % loop on drift-free
longitudinal profile

        figure(10+i); clf(10+i); hold all;
        title(['Error Vectors at u = ', num2str(xlong(i*300)), ' m, .
', num2str(nBasis), ' basis vectors']);
        xlabel('Transverse Distance, [m]'); ylabel('Amplitude, [mm]');
grid on;
        for k=1:r,
            plot(xtrans,1000*error(i*300,:,k));
            %plot(tranProf(:,k));          % drift-free
        end
        %plot(trueSurface(i*200,:),'k:');
        figure(20+i); (20+i); hold all;
        title(['Noise Vectors at u = ', num2str(xlong(i*300)), ' m, .
', num2str(nBasis), ' basis vectors']);
        xlabel('Transverse Distance, [m]'); ylabel('Amplitude, [mm]');
grid on;
        for k=1:r,
            plot(xtrans,1000*noise(i*300,:,k))
        end
        figure(30+i); (30+i); hold all;
        title(['Drift Vectors at u = ', num2str(xlong(i*300)), ' m, .
', num2str(nBasis), ' basis vectors']);
        xlabel('Transverse Distance, [m]'); ylabel('Amplitude, [mm]');
grid on;
        for k=1:r,
            plot(xtrans,1000*drift(i*300,:,k))
        end

    end

    for j=1:1:min(9,floor(n/10)),           % loop on drift-free
longitudinal profile

        figure(40+j); (40+i); hold all;
        title(['Drift-Free and True Surface at v = ',
num2str(xtrans(j*10)), ' m, .  ', num2str(nBasis), ' basis vectors']);
        xlabel('Longitudinal Distance, [m]'); ylabel('Amplitude, [m]');
grid on;
        longProf=zeros(m,r);
        for k=1:r,
            longProf(:,k)=bigZ(:,j*10,k)-drift(:,j*10,k);       % these
are the samples of drift-free long profiles
            plot(xlong,longProf(:,k));          % drift-free
        end
        plot(xlong,trueSurface(:,j*10),'k:');

        disp(['cross correlation for long prof. ', num2str(j*10)]);
        c=corr(bigZ(:,j*10,:)),
```

```matlab
        %disp(['standard deviation of noise for long prof. ',
num2str(j*10)]);
        %stand = std(noise(:,j*10,:),0,2),

        % look at drift
        figure(50+j); (50+i); hold all;
        title(['Drift at v = ', num2str(xtrans(j*10)), ' m, .   ',
num2str(nBasis), ' basis vectors']);
        xlabel('Longitudinal Distance, [m]'); ylabel('Amplitude,
[mm]'); grid
on;legend('k=1','k=2','k=3','k=4','k=5','k=6','k=7','k=8','k=9','k=10')
        for k=1:r,
            driftProf(:,k)=drift(:,j*10,k);       % these are the
samples of long drift profiles
            plot(xlong,driftProf(:,k)*1000);
        end
        % look at noise
        figure(60+j); (60+i); hold all;
        title(['Noise at v = ', num2str(xtrans(j*10)), ' m, .   ',
num2str(nBasis), ' basis vectors']);
        xlabel('Longitudinal Distance, [m]'); ylabel('Amplitude,
[mm]'); grid on;          for k=1:r,
            %noiseProf(:,k)=longProf(:,k)-trueSurface(:,j*10);       %
these are the samples of longitudinal noise profiles
            noiseProf(:,k)=noise(:,j*10,k);
            plot(xlong,noiseProf(:,k)*1000);
        end
    end    % for j=1:1:8,       % loop on drift-free longitudinal
profile

end    % while (CheckNoise),

disp(StdTotal);
disp(MeanTotal);
disp(nBasis);          % initialize nBasis to zero, it is immediately
incremented inside the while loop
disp(maxNoise);
disp(minNoise);

% 9) Calculate the drift-free surfaces etc.
for k=1:r,
    arrayData(k).driftFreeZ = bigZ(:,:,k)-drift(:,:,k);
    arrayData(k).noise = noise(:,:,k);
    arrayData(k).drift = drift(:,:,k);
    arrayData(k).trueSurface = trueSurface;
end

% % 10) create additional structure in the arrayData array with True
Surface
% arrayData(r+1).z = trueSurface;

save test.mat;

for k=1:r,
    save_D(arrayData(k));
```

```matlab
    end


    return    % remove_ins_drift



function [sigma, b] = get_basis_vectors(bigZ, nBasis)
% This function performs a SVD on all the differences in measured
vectors
% to determine a set of basis vectors that spans the vector space of
the
% error.
% bigZ is the concatenation of the measured surfaces
%   bigZ(i,j,k) then
%   i: index of transverse profile ("row"), i=[1,2,..., m]
%   j: index of longitudinal profile ("col"), j=[1,2,..., n]
%   k: index of realization (measurement), k=[1,2,..., r]
[m,n,r]=size(bigZ);

% Determine the number of combinations of differences in measured
vectors
nDiffMeas = m*(r-1)*r/2;
maxCounterDiffMeas = 1e4;    % only consider the first
maxCounterDiffMeas vectors...
maxCounterDiffMeas = min(nDiffMeas, maxCounterDiffMeas);

diffMeas = zeros(maxCounterDiffMeas,n);
counterDiffMeas = 0;          % counter the number of vectors for SVD

% TODO randomly select the pairs of measured vectors to use, using
% [index]=generate_random_indices(m,num)
for i=1:m,
    for k=1:(r-1),
        for h=k+1:r,
            counterDiffMeas = counterDiffMeas + 1;
            diffMeas(counterDiffMeas,:) = bigZ(i,:,k)-bigZ(i,:,h);
            if counterDiffMeas == maxCounterDiffMeas, break; end
        end
        if counterDiffMeas == maxCounterDiffMeas, break; end
    end
    if counterDiffMeas == maxCounterDiffMeas, break; end
end

[~,sigma,b] = svds(diffMeas, nBasis);


    return    % get_basis_vectors



function [expectedS] = get_expectation_20090515(s, dim, method)
%The expectation of variable (in any number of dimensions(, s,
% is returned after s has ben 'averaged' over
% the 'dim' dimension using the 'method' method.

[nRows,nCols,nRuns] = size(s);
```

199

```matlab
switch method
    case 1
        expectedS = mean(s, dim);
    case 2
        expectedS = trimmean(s, 50, dim);
    case 3
        expectedS = median(s, dim);
    otherwise
end

return    % get_expectation


function [myMean, myStd, StdTotal, MeanTotal, VarTotal, maxNoise, minNoise]=CheckNoise(noise)
%CheckNoise: This function loads the noise surface(s) and evaluates for a
%zero-mean process in terms of the mean and standard deviation
%
% Input:  noise: noise vectors from INS drift code
%
% Output: myMean: mean noise surface combining mean of each kth value
%         myStd:  standard deviaiton noise surface combining std of each
%                 kth value
%         StdTotal: Single number representation of std deviation of
%                   surface
%         MeanTotal: Single number representation of mean of surface
%         VarTotal: Single number representation of variance of surface

[m,n,r] = size(noise);
%Establish size of matrices
myMean = zeros(m,n);
myStd = zeros(m,n);

% Calculate mean noise surface and standard deviation noise surface for
% each kth measurement
for i=1:m
    for j=1:n
        myMean(i,j)=mean(noise(i,j,:));
        myStd(i,j)=std(noise(i,j,:));
    end
end

%Std Deviation and variance of surface as represented by a single number
VarTotal=0;
for i=1:m
    for j=1:n
        VarTotal=VarTotal+var(noise(i,j,:));
    end
end
VarTotal=VarTotal/(m*n);
StdTotal=sqrt(VarTotal)
```

```matlab
%Mean of surface as represented by a single number
NoiseTotal=0;
maxNoise = 0;
minNoise = 0;
for i=1:m
    for j=1:n
        NoiseTotal=NoiseTotal+(noise(i,j,:));
        maxNoise=max( maxNoise, max(noise(i,j,:)) );
        minNoise=min( minNoise, min(noise(i,j,:)) );
    end
end
MeanTotal=mean(NoiseTotal/(m*n))

return
```

Appendix K – Analytic Basis Vectors Sample Matlab Code

This Appendix provides a sample of Matlab code used to calculate the RMSE between the empirical and analytic basis vectors.

```matlab
function [sigma_Gegen, sTrunc_GegenSum, sigma, sTruncSum, basis,
basis_trunc, Gegen_basis, RMS_gegen, Fourier_basis, RMS_fourier,
Chebyshev_basis,
RMS_chebyshev,v]=AnalyticvsEmpirical_2010_09_29(v_Indexa,lambda,shift_v
_horiz, scale_v, scale_basis,q)
%% Description
%This function compares the Analytic basis vectors with the Emperical
Basis
%Vectors for Gegenbauer tuned on lambda, Chebyshev and Fourier
%
%Inputs:  (v_Indexa,lambda,shift_v_horiz, scale_v, scale_basis,q)
%
%   v_Indexa: vector of whole number indices i.e. [1:1:50] creates.
Will
%   map the space from [-1, 1] by 50 points.
%   lambda: parameter to tune Gegenbauer polynomials, must exist on
(0,1]
%   shift_v_horiz: shift the empirical basis vectors to exist and be
%   centered on [-, 1]
%   scale_v:  scales the width of the span of the empirical basis
vectors
%   scale_basis: magnifies empirical basis vectors
%   q: number of analytic basis vectors to be generated
%
%
%% Adjust Empirical basis vectors and formulate v vector
%Load Empirical basis vectors
load('C:\Documents and Settings\IALR\Desktop\Candidate
Release\TestData\EmpiricalBasis.mat');
load('C:\Documents and Settings\IALR\Desktop\Candidate
Release\TestData\TrueSurface_Section1.mat');

v_Empir_Orig=linspace(0.5,1.8,48);  %Original data spacing in meters

%Need to scale original data to be centered on 0.  In this case, the
data
%spans 1.3 meters.

v_Empir_cent=v_Empir_Orig-1.1375;    %Centers data on 0.  Spans -0.6375
to 0.6375m

v_Empir_horiz_shift=v_Empir_cent+shift_v_horiz; %Shifts v_empir to the
left or right
```

```matlab
v_Empir=v_Empir_horiz_shift*scale_v;   %Scales v_empir by a
multiplicative factor

basisscaled(:,:)=basis(:,:)*scale_basis;  %Scales magnitude of basis
vectors

v_spacing=length(v_Indexa);

%Analytic Basis Vectors%
v=linspace(-1,1,v_spacing);

%% Gegenbauer Polynomials

[~, Gegen_basis, ~, w] = make_basis_2010_04_30(v, lambda);


%% Root Mean square Error Routine Gegenbauer

%truncate emprical basis vectors to be same length as analytic%
basis_trunc=basisscaled(v_Indexa,:);

sum_total=zeros(length(v),4);

for h=1:3
    for i=1:length(v)
        sum(i,h)=(Gegen_basis(i,h)-basis_trunc(i,h))^2;
        sum_total(i,h)=sum_total(i,h)+sum(i,h);
    end
    RMS_gegen(:,h)=sqrt(sum_total(:,h)/length(v));
end

for i=1:length(v)
    sum(i,3)=(Gegen_basis(i,4)-basis_trunc(i,3))^2;
    sum_total(i,3)=sum_total(i,3)+sum(i,3);
end
RMS_gegen(:,3)=sqrt(sum_total(:,3)/length(v));

for i=1:length(v)
    sum(i,4)=(Gegen_basis(i,5)-basis_trunc(i,4))^2;
    sum_total(i,4)=sum_total(i,4)+sum(i,4);
end
RMS_gegen(:,4)=sqrt(sum_total(:,4)/length(v));

%% Plotting Gegenbauer

%Color scheme for plots
clr1={'b*','r*','g*','k*'};
clr2={'b','r','g','k'};

disp(['For lambda= ' num2str(lambda)])

figure; title('Gegenbauer vs. Empirical')
```

```matlab
%title(['Lambda = ( ', num2str(lambda), ')'])
subplot(2,2,1),hold on, grid on, plot(v_Empir,
basisscaled(:,1),clr1{1}), plot(v,Gegen_basis(:,1),clr2{1}),
xlabel('Normalized Transverse Distance, v, [m]')
ylabel('Amplitude'), legend('Empirical q=1', 'Analytic
p_1','Location','Best'), hold off;
subplot(2,2,2),hold on, grid on, plot(v_Empir,
basisscaled(:,2),clr1{2}),
plot(v,Gegen_basis(:,2),clr2{2});xlabel('Normalized Transverse
Distance, v, [m]')
ylabel('Amplitude');legend('Empirical q=2', 'Analytic
p_2','Location','Best'), hold off;
subplot(2,2,3),hold on, grid on, plot(v_Empir,
basisscaled(:,3),clr1{3}),
plot(v,Gegen_basis(:,4),clr2{3});xlabel('Normalized Transverse
Distance, v, [m]')
ylabel('Amplitude');legend('Empirical q=3', 'Analytic
p_4','Location','Best'), hold off;
subplot(2,2,4),hold on, grid on, plot(v_Empir,
basisscaled(:,4),clr1{4}),
plot(v,Gegen_basis(:,5),clr2{4});xlabel('Normalized Transverse
Distance, v, [m]')
ylabel('Amplitude'), legend('Empirical q=4', 'Analytic
p_5','Location','Best'), hold off;
%legend('Gegenbauer Basis q=0','EmpiricalBasis q=0', 'Gegenbauer Basis
q=1','EmpiricalBasis q=1','Gegenbauer Basis q=3','Gegenbauer Basis
q=4', 'EmpiricalBasis q=2','EmpiricalBasis q=3','Location','Best')



%% Chebyshev Polynomials

[Chebyshev_basis]=Chebyshev(v);



%% Root Mean square Error Routine Chebyshev

sum_totalCheby=zeros(length(v),4);

for h=1:3
    for i=1:length(v)
        sumCheby(i,h)=(Chebyshev_basis(i,h)-basis_trunc(i,h))^2;
        sum_totalCheby(i,h)=sum_totalCheby(i,h)+sumCheby(i,h);
    end
    RMS_chebyshev(:,h)=sqrt(sum_totalCheby(:,h)/length(v));
end

for i=1:length(v)
    sumCheby(i,3)=(Chebyshev_basis(i,4)-basis_trunc(i,3))^2;
    sum_totalCheby(i,3)=sum_totalCheby(i,3)+sumCheby(i,3);
end
RMS_chebyshev(:,3)=sqrt(sum_totalCheby(:,3)/length(v));

for i=1:length(v)
    sumCheby(i,4)=(Chebyshev_basis(i,5)-basis_trunc(i,4))^2;
    sum_totalCheby(i,4)=sum_totalCheby(i,4)+sumCheby(i,4);
```

```matlab
end
RMS_chebyshev(:,4)=sqrt(sum_totalCheby(:,4)/length(v));

%% Plotting Chebyshev

%Color scheme for plots
clr1={'b*','r*','g*','k*'};
clr2={'b','r','g','k'};

disp(['For lambda= ' num2str(lambda)])

figure; title('Chebyshev vs. Empirical')
%title(['Lambda = ( ', num2str(lambda), ')'])
subplot(2,2,1),hold on, grid on, plot(v_Empir,
basisscaled(:,1),clr1{1}), plot(v,Chebyshev_basis(:,1),clr2{1}),
xlabel('Normalized Transverse Distance, v, [m]')
ylabel('Amplitude'), legend('Empirical q=1', 'Analytic
T_1','Location','Best'), hold off;
subplot(2,2,2),hold on, grid on, plot(v_Empir,
basisscaled(:,2),clr1{2}),
plot(v,Chebyshev_basis(:,2),clr2{2});xlabel('Normalized Transverse
Distance, v, [m]')
ylabel('Amplitude');legend('Empirical q=2', 'Analytic
T_2','Location','Best'), hold off;
subplot(2,2,3),hold on, grid on, plot(v_Empir,
basisscaled(:,3),clr1{3}),
plot(v,Chebyshev_basis(:,4),clr2{3});xlabel('Normalized Transverse
Distance, v, [m]')
ylabel('Amplitude');legend('Empirical q=3', 'Analytic
T_4','Location','Best'), hold off;
subplot(2,2,4),hold on, grid on, plot(v_Empir,
basisscaled(:,4),clr1{4}),
plot(v,Chebyshev_basis(:,5),clr2{4});xlabel('Normalized Transverse
Distance, v, [m]')
ylabel('Amplitude'), legend('Empirical q=4', 'Analytic
T_5','Location','Best'), hold off;
%legend('Gegenbauer Basis q=0','EmpiricalBasis q=0', 'Gegenbauer Basis
q=1','EmpiricalBasis q=1','Gegenbauer Basis q=3','Gegenbauer Basis
q=4', 'EmpiricalBasis q=2','EmpiricalBasis q=3','Location','Best')

%% Fourier Polynomials

[~, Fourier_basis] = make_basis_fourier_2010_09_27(v, q);


%% Root Mean square Error Routine Fourier

sum_totalFourier=zeros(length(v),4);

for h=1:3
    for i=1:length(v)
        sumFourier(i,h)=(Fourier_basis(i,h)-basis_trunc(i,h))^2;
        sum_totalFourier(i,h)=sum_totalFourier(i,h)+sumFourier(i,h);
    end
    RMS_fourier(:,h)=sqrt(sum_totalFourier(:,h)/length(v));
```

```matlab
end

for i=1:length(v)
    sumFourier(i,3)=(Fourier_basis(i,3)-basis_trunc(i,3))^2;
    sum_totalFourier(i,3)=sum_totalFourier(i,3)+sumFourier(i,3);
end
RMS_fourier(:,3)=sqrt(sum_totalFourier(:,3)/length(v));

for i=1:length(v)
    sumFourier(i,4)=(Fourier_basis(i,4)-basis_trunc(i,4))^2;
    sum_totalFourier(i,4)=sum_totalFourier(i,4)+sumFourier(i,4);
end
RMS_fourier(:,4)=sqrt(sum_totalFourier(:,4)/length(v));

%% Plotting Fourier

%Color scheme for plots
clr1={'b*','r*','g*','k*'};
clr2={'b','r','g','k'};

disp(['For lambda= ' num2str(lambda)])

figure; title('Fourier vs. Empirical')
%title(['Lambda = ( ', num2str(lambda), ')'])
subplot(2,2,1),hold on, grid on, plot(v_Empir,
basisscaled(:,1),clr1{1}), plot(v,Fourier_basis(:,1),clr2{1}),
xlabel('Normalized Transverse Distance, v, [m]')
ylabel('Amplitude'), legend('Empirical q=1', 'Analytic
f_1','Location','Best'), hold off;
subplot(2,2,2),hold on, grid on, plot(v_Empir,
basisscaled(:,2),clr1{2}),
plot(v,Fourier_basis(:,2),clr2{2});xlabel('Normalized Transverse
Distance, v, [m]')
ylabel('Amplitude');legend('Empirical q=2', 'Analytic
f_2','Location','Best'), hold off;
subplot(2,2,3),hold on, grid on, plot(v_Empir,
basisscaled(:,3),clr1{3}),
plot(v,Fourier_basis(:,3),clr2{3});xlabel('Normalized Transverse
Distance, v, [m]')
ylabel('Amplitude');legend('Empirical q=3', 'Analytic
f_3','Location','Best'), hold off;
subplot(2,2,4),hold on, grid on, plot(v_Empir,
basisscaled(:,4),clr1{4}),
plot(v,Fourier_basis(:,4),clr2{4});xlabel('Normalized Transverse
Distance, v, [m]')
ylabel('Amplitude'), legend('Empirical q=4', 'Analytic
f_4','Location','Best'), hold off;
%legend('Gegenbauer Basis q=0','EmpiricalBasis q=0', 'Gegenbauer Basis
%q=1','EmpiricalBasis q=1','Gegenbauer Basis q=3','Gegenbauer Basis
q=4', 'EmpiricalBasis q=2','EmpiricalBasis q=3','Location','Best')


%% Plotting RMS to compare each basis vector

figure
```

```matlab
grid on
title('RMSE of Analytic vs. Empirical (1^s^t Basis Vector)')
hold on
plot(v,RMS_gegen(:,1),'g')
plot(v,RMS_chebyshev(:,1),'r')
plot(v,RMS_fourier(:,1),'b')
xlabel('Normalized Transverse Distance, v, [m]')
ylabel('Root Mean Square Error')
legend('C_1 vs. b_1','T_1 vs. b_1', 'f_1 vs. b_1')

figure
grid on
title('RMSE of Analytic vs. Empirical (2^n^d Basis Vector)')
hold on
plot(v,RMS_gegen(:,2),'g')
plot(v,RMS_chebyshev(:,2),'r')
plot(v,RMS_fourier(:,2),'b')
xlabel('Normalized Transverse Distance, v, [m]')
ylabel('Root Mean Square Error')
legend('C_2 vs. b_2','T_1 vs. b_2', 'f_2 vs. b_2')

figure
grid on
title('RMSE of Analytic vs. Empirical (3^r^d Basis Vector)')
hold on
plot(v,RMS_gegen(:,3),'g')
plot(v,RMS_chebyshev(:,3),'r')
plot(v,RMS_fourier(:,3),'b')
xlabel('Normalized Transverse Distance, v, [m]')
ylabel('Root Mean Square Error')
legend('C_4 vs. b_3','T_4 vs. b_3', 'f_3 vs. b_3')

figure
grid on
title('RMSE of Analytic vs. Empirical (4^t^h Basis Vector)')
hold on
plot(v,RMS_gegen(:,4),'g')
plot(v,RMS_chebyshev(:,4),'r')
plot(v,RMS_fourier(:,4),'b')
xlabel('Normalized Transverse Distance, v, [m]')
ylabel('Root Mean Square Error')
legend('C_5 vs. b_4','T_5 vs. b_4', 'f_4 vs. b_4')

%% Terrain Components Empirical

[m,n]=size(TrueSurface);
basis=basis(1:48,:);

%Calculate the projection of the True Surface onto the principle
    %directions
  sigma = zeros(m,4);
    for i=1:m,
        for l=1:4,
                sigma(i,l)=TrueSurface(i,:)*basis(:,l);
        end
    end
```

```matlab
    %Calculate the truncated surface
    sTruncSum=zeros(m,n);
    sTrunc=zeros(m,n);
    for i=1:m,
        for l=1:4,
            sTrunc(i,:)=sigma(i,l)*basis(:,l)';
            sTruncSum(i,:)=sTruncSum(i,:)+sTrunc(i,:);
        end
    end

%% Terrain Components Analytical
vIndex=linspace(1,48,20);
vIndex=floor(vIndex);
[a,b]=size(Gegen_basis);
TrueSurfaceTrunc=TrueSurface(:,vIndex);

    %Calculate the projection of the True Surface onto the principle
     %directions
  sigma_Gegen = zeros(m,20);
    for i=1:m,
        for l=1:20,

sigma_Gegen(i,l)=TrueSurfaceTrunc(i,:)*w*Gegen_basis(:,l);
        end
    end

    %Calculate the truncated surface
    sTrunc_Gegen=zeros(m,a);
    sTrunc_GegenSum=zeros(m,a);
    for i=1:m,
        for l=1:18,
%
sTrunc_Gegen(i,:)=sTrunc_Gegen(i,:)+sigma_Gegen(i,l)*Gegen_basis(:,l)';
            sTrunc_Gegen(i,:)=sigma_Gegen(i,l)*Gegen_basis(:,l)';

sTrunc_GegenSum(i,:)=sTrunc_GegenSum(i,:)+sTrunc_Gegen(i,:);
        end
    end
return


function [coeff, b, p, w] = make_basis_2010_04_30(v, lambda)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%
% DESCRIPTION%
%
% returns the transition matrix so that sigma = invV*s
%
% Assumptions
%
% 1) Uses WEIGHTED inner-product (w=diag(1/nv); b'*w*b)
% 2) Changes v to a column vector
%
% lambda = 1/2 => Legendre
%
```

```
%
% INPUTS
%
% 1) v: row vector with lateral locations of the grid with respect to
path (meters)
% 2) lambda: parameter defining the Gegenbauer polynomials
%
% OUTPUTS
%
% 1) coeff: a matrix of vectors representing the coefficients for the
analytical polynomials
% 2) b: a matrix of analytical basis vectors drived from the Gegenbauer
% polynomials evaluated at the locations defined by v
% 3) p: a matrix of polynomials evaluated at locations defined by x =
% vector v
% 4) w: weighting matrix for inner-product
%
% INTERNAL VARIABLE DEFINITONS
%
%   .nv: number of v values
%
% TASKS
% 1) Initialize variables
% 2) Form first two polynomial vectors and coefficients directly
% 3) Form additional vectors from recursive method
% 4) Check rank of coeff and form inverse
% 5) Form Vandermonde matrix, V
% 6) Check rank of V and form inverse
% 7) Check p = V*coeff
% 8) Form orthonormal basis vectors
% 9) Form the weighting matrix w for the inner-product
% 10) Check w is positive defninite
%    10a) Check that w is symmetric
%    10b) Check that w has positive eigenvalues


%
% VALIDATION
% v=[0 1 2 3 4] - validated up through p4 (C3)
% v=[-1 -.5 0 .5 1]; lambda = .5;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%

% 1) Initialize variables
sizeV = size(v);
if sizeV(2)>sizeV(1), v=v'; end;    %Force v to be a vector
nv = length(v);
p  = zeros(nv,nv);
coeff = zeros(nv,nv);
errorTol = (nv^2)*1.0e-10;  % effectively an acceptable "zero"

% 2) Form first two polynomial vectors and coefficients directly
p(:,1) = ones(nv,1);
coeff(1,1) = 1;

p(:,2) = 2*lambda*v;
coeff(2,2) = 2*lambda;
```

```matlab
% 3) Form additional vectors from recursive method
for i=3:nv,
    n=i-1;                      % 'n' in documentation is i-1...
    c1 = 2*(n+lambda-1)/n;  % coefficient multiplying vP(n-1)
    c2 = -(n+2*lambda-2)/n;  % coefficient multiplying P(n-2)
    p(:,i) = c1*(v.*p(:,i-1)) + c2*p(:,i-2);
    % coeff calculated here...
    coeff(1,i) = c2*coeff(1,i-2);
    for j=2:nv,
        coeff(j,i) = c1*coeff(j-1,i-1)+c2*coeff(j,i-2);
    end

end


% 4) Check rank of coeff and form inverse
rankCoeff = rank(coeff);
if (rankCoeff<nv),
    disp('coeff matrix is rank deficient');
    disp(['rank coeff =    ' num2str(rankCoeff)]);
    disp(['required rank = ' num2str(nv)]);
    invC=[];
else
    invC  = inv(coeff);
    invCt = inv(coeff');
end



% 5) Form Vandermonde matrix, V
V = zeros(nv,nv);
for i=1:nv,
    V(:,i)=v.^(i-1);
end

% 6) Check rank of V and form inverse
rankV = rank(V);
if (rankV<nv),
    disp('V matrix is rank deficient');
    disp(['rank V =        ' num2str(rankV)]);
    disp(['required rank = ' num2str(nv)]);
    invV=[];
else
    invV  = inv(V);
    invVt = inv(V');
end


% 7) Check p = V*coeff
error = abs(max(max(p-V*coeff)));
if error> errorTol,
    disp('p does not equal V*coeff');
    disp(['min(min(p-V*coeff)) =    ' num2str(error)]);
    disp(['maximum difference  = ' num2str(errorTol)]);
    coeff=[];
    V=[]; v=[];
```

```matlab
    p=[];
end


% 8) Form orthonormal basis vectors
b=V*coeff;


% 9) Form the weighting matrix w for the inner-product
w=invVt*invCt*invC*invV;        % TODO use '/' or '\' to take inverse
instead of 'inv'


% 10) Check w is positive defninite
%   10a) Check that w is symmetric
error = abs(max(max(w-w')));
if (error>errorTol*1.0e+5),     % W is a very poorly conditioned
matrix, implement '/' function and change this errorTol
    disp('w is not symmetric');
    disp(['errorTol = ' num2str(errorTol)]);
    disp(['error = ' num2str(error)]);
    w=[];
end

%   10b) Check that w has positive eigenvalues
error = min(eig(w));
if (error < 0),
    disp('w does not have positive eigenvalues');
    disp(['min eig = ' num2str(error)]);
    w=[];
end


%Development items - plot up to the first 5 basis vectors
% figure; plot(v,p(:,1:min(5,nv)),'-'); xlabel('v'); ylabel('p');
title(['lambda = ', num2str(lambda)]); hold on;
plot(v,p(:,1:min(5,nv)),'o');
figure; plot(v,b(:,1:min(5,nv)),'-'); xlabel('Normalized Transverse
Distance, v,[m]'); ylabel('Amplitude'); title(['Gegenbauer Basis
Vectors (lambda = ', num2str(lambda), ')']); hold on;
plot(v,b(:,1:min(5,nv)),'o');

return

 function [a, b, f] = make_basis_fourier_2010_09_24(v)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%
% DESCRIPTION%
%
% Assumptions
%
% 2) Changes v to a column vector
%
%
```

211

```matlab
% INPUTS
%
% 1) v: row vector with lateral locations of the grid with respect to
path (meters)
% 2) n: number of basis vectors desired to plot
%
% OUTPUTS
%
% 3) p: a matrix of polynomials evaluated at locations defined by x =
% vector v
%
% INTERNAL VARIABLE DEFINITONS
%
%   .nv: number of v values
%
% TASKS
% 1) Initialize variables


%
% VALIDATION
% v=[0 1 2 3 4] - validated up through p4 (C3)
% v=[-1 -.5 0 .5 1]; lambda = .5;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%

% 1) Initialize variables
sizeV = size(v);
if sizeV(2)>sizeV(1), v=v'; end;     %Force v to be a vector
v = sort(v);                         %make sure v is monotonically
increasing
nv = length(v);
p  = zeros(nv,nv);
a = zeros(nv,nv+1);
b = zeros(nv,nv);
%errorTol = (nv^2)*1.0e-10;  % effectively an acceptable "zero"

theta = 2*pi*(v-v(1))/(v(nv)-v(1));    % Normalize 0 < theta < 2*pi
maps to vector v

 a(:,1) = ones(nv,1);    % this is the constant term... a_sub_0
 for j = 1:nv,
     a(:,j+1) = cos(j*theta);
     b(:,j) = sin(j*theta);
 end
figure(1); hold all;
xlabel('v');
ylabel('Fourier curves - cosines');
plot(v,a(:,1))
for j = 1:nv,
     plot(v,a(:,j+1));
end

figure(2); hold all;
xlabel('v');
ylabel('Fourier curves - sines');
for j = 1:nv,
```

```matlab
        plot(v,b(:,j));
end

%% Fourier Series %%

f=ones(nv,1);
term1=0;
term2=0;

for j=1:nv
    term1=term1+(a(:,j+1)).*cos(theta*v);
    term2=term2+b(:,j).*sin(theta*v);
end
f=a(:,1)./2+term1+term2;
return

 function [chebyshev]=Chebyshev(x)
%Input row vector and polynomial order for chebyshev polynomials of
order 1
chebyshev(:,1)=ones(length(x),1);
chebyshev(:,2)=1.*x;
chebyshev(:,3)=2.*x.^2-1;
chebyshev(:,4)=4.*x.^3-3*x;
chebyshev(:,5)=8.*x.^4-8*x.^2+1;
chebyshev(:,6)=16.*x.^5-20.*x.^3+5.*x;
chebyshev(:,7)=32.*x.^6-48.*x.^4+18.*x.^2-1;
chebyshev(:,8)=64.*x.^7-112.*x.^5+56.*x.^3-7.*x;

figure
plot(x,chebyshev(:,1),'o-b');
hold on
plot(x,chebyshev(:,2),'g');
plot(x,chebyshev(:,3),'-^m');
plot(x,chebyshev(:,4),'-xr');
plot(x,chebyshev(:,5),'--c');
xlabel('Normalized Transverse Distance, v, [m]')
ylabel('Amplitude')
legend('Analytic T_1','Analytic T_2','Analytic T_3','Analytic
T_4','Analytic T_5')
grid on
return
```