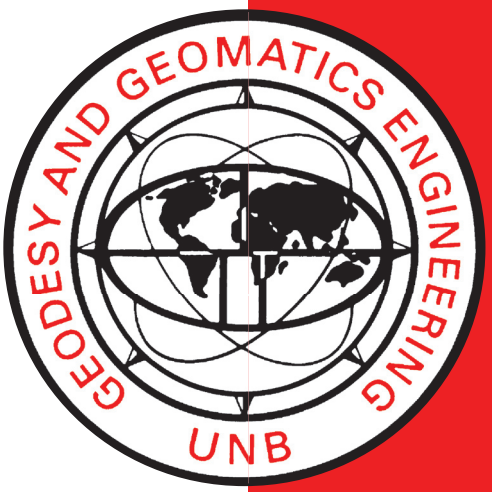
PREDICTION OF ERROR DUE TO TERRAIN SLOPE IN LIDAR OBSERVATIONS

TRISTAN GOULDEN

January 2009



TECHNICAL REPORT
NO. 265

PREDICTION OF ERROR DUE TO TERRAIN SLOPE IN LIDAR OBSERVATIONS

Tristan Goulden

Department of Geodesy and Geomatics Engineering
University of New Brunswick
P.O. Box 4400
Fredericton, N.B.
Canada
E3B 5A3

January 2009

PREFACE

This technical report is a reproduction of a thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Engineering in the Department of Geodesy and Geomatics Engineering, January 2009. The research was supervised by Dr. Peter Dare, and support was provided by the Geological Survey of Canada, Glaciology Division.

As with any copyrighted material, permission to reprint or quote extensively from this report must be received from the author. The citation to this work should appear as follows:

Goulden, Tristan (2009). *Prediction of Error Due to Terrain Slope in LiDAR Observations*. M.Sc.E. thesis, Department of Geodesy and Geomatics Engineering Technical Report No. 265, University of New Brunswick, Fredericton, New Brunswick, Canada, 138 pp.

ABSTRACT

Glacial environments around the world are experiencing decadal changes over time that result in a consistent decline in their volume. Currently there are many tools used to map and monitor the change in glacial surfaces. One tool that has proved successful is LiDAR (Light Detection and Ranging). However, this technology is relatively new and error parameters are not well defined. This makes it difficult to separate detected melt from the error. One error source in LiDAR observations that is particularly prevalent in glacial environments but has remained undefined is caused by the slope of the terrain. Terrain slope has long been known to add significant error in DSMs (Digital Surface Models), especially those determined by LiDAR. The terrain slope error will have significant consequences to any parameters derived from the DSM and so will cause detriment to subsequent analyses.

This paper presents an approach for predicting the level of error in LiDAR observations due to slope effects. The algorithm is based on initial error estimates obtained by propagating the error of each hardware component through the LiDAR direct geo-referencing equation. Once initial estimates of error are determined they are represented as 3D error ellipsoids at each LiDAR observation. The local slope of the observed terrain and the initial error estimates are combined to calculate a worst-case scenario prediction of the true terrain. Vertical errors are then extracted by comparing the elevation of the actual point and the predicted surface.

Two test sites were chosen to verify that the predictions matched observed errors on sloped terrain. One test site was located at a ski hill in Windsor, Nova Scotia and a

second from a sloped road in Fredericton, New Brunswick. Tests from the ski hill showed optimistic predictions of the observed error. Testing from the road site in Fredericton showed pessimistic predictions of error, which were expected since the algorithm was designed to predict worst-case scenario solutions. Both test sites showed an improvement over existing LiDAR error prediction capabilities. It was concluded that the algorithm would benefit from additional predictions of currently un-defined sources of error in LiDAR observations. The error predictions do serve as better estimates of final DSM error than predictions based solely on hardware errors and should be included in studies involving LiDAR DSMs.

Acknowledgements

In addition to the support I received from my supervisors and examining board I would like to highlight additional individual efforts that aided in the completion of the work presented here:

- Mike Demuth of the Geological Survey of Canada, Glaciology Division, for providing funding and support for both the academic work and fieldwork.
- The Applied Geomatics Research Group for providing the necessary software and equipment for the research
- Laura Chasmer for operating the LiDAR equipment during field campaigns.
- Peter Horne, Kevin Garroway, Harsil Jani and Frank Bailey for fieldwork assistance during validation surveys.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1
Introduction

The world's glaciers are melting. Every year glaciers experience a periodic extension and retreat as winter and summer seasons pass. As global temperatures steadily increase an interesting trend has begun; glaciers tend to thaw slightly more than they accumulate each year. Although almost visually imperceptible on a yearly basis, the occurrence becomes very evident after decades have passed. Often used as illustrations for bringing the effects of global warming into popular culture, melting glaciers also serve another purpose they have served for millennia: providing water.

Every summer as glaciers experience their melt, high volumes of water are expelled into surrounding streams and rivers. The highest volumes of water are experienced near the end of summer when temperatures are the hottest. This is also typically the driest time of the year, when water levels begin to decline from seasonal highs experienced from snow melt and spring rains. If glaciers are melting at rates not previously documented it's possible that they will soon be extinct and a critical water source, provided at a critical time, for many parts of the world will dry up.

In order to detect the melting of glaciers on a yearly basis, a mapping tool that provides accurate and dense three-dimensional point samples on the glaciers physical surface is required. Light Detection and Ranging (LiDAR) has the capability to be such a tool. If LiDAR is used to map a glacial surface on a yearly basis, constructed surfaces can be subtracted to determine changes in the glacial surface. In any change detection analysis, separating apparent change from error in the observation system is a vital

consideration. If not taken into account, reported change may simply be a result of errors in the observation system and not from physical change. Since LiDAR is a relatively new technology, the error parameters associated with its observation are not well defined. This thesis intends to provide a methodology for predicting the error based on terrain slope, one of the largest errors in LiDAR observations, which is particularly prevalent in glacial environments.

1.1 LiDAR

LiDAR, which stands for Light Detection And Ranging, is a developing technology in the remote sensing industry. LiDAR is considered an active remote sensing instrument because it provides its own energy source. It uses this energy source to illuminate a target to obtain reliable measurements without coming into direct physical contact. It is similar to other well known active remote sensing technologies such as RADAR (Radio Detection and Ranging) and SONAR (Sound Navigation and Ranging). The important distinction between each of these technologies is the energy source that is used to make measurements. While RADAR uses energy in the microwave portion of the electromagnetic spectrum and SONAR uses sound energy LiDAR uses laser (light amplification by stimulated emission of radiation) energy. Although the observation characteristics and geometry between each sensor are similar, as can be seen from the similarity in the acronyms, the energy source provides for advantages and drawbacks of each system.

The first functioning lasers are attributed to Arthur Schawlow and Charles Townes in 1958 and were used for medical purposes [Maune, 2007]. Lasers are unique in

that they produce a highly collimated and high energy beam of light that can be easily directed. This is advantageous in remote sensing since the illuminated footprint on the target will be relatively small, allowing for accurate coordination. This contrasts RADAR or SONAR where illuminated footprints are large. The resulting coordination in these systems becomes more difficult due to target position ambiguity within the footprint.

The use of lasers in remote sensing is not a new idea. Electronic Distance Measurements from terrestrial platforms have been common for the past thirty years and are used extensively in the surveying industry. Geodetic observations are made through satellite laser ranging and lunar laser ranging. These applications required stationary platforms to transmit the laser signal and properly coordinate targets. Although mounting the laser sensor to a moving platform was theoretically possible, it wasn't until the advent of several enabling technologies that lasers were mounted onto moving airborne and spaceborne platforms that led to current LiDAR systems.

In order to properly geo-reference ground targets from a moving platform the orientation and position of the platform and the direction of the emitted energy must be known in a common coordinate system. These requirements became recently available with the advent of kinematic Global Positioning System (GPS) technology and accurate inertial measurement units (IMU). Kinematic GPS is able to provide for accurate three-dimensional positioning on a moving platform while the IMU can determine platform orientation. Combining these technologies with a laser ranger enables accurate three dimensional coordinates to be determined on the physical terrain.

Once these enabling technologies became accessible, LiDAR systems quickly developed sophisticated scanning mechanisms and rapid output of laser pulses. The

scanning mechanisms are used to direct laser pulses into pre-defined swaths that sample large areas quickly. The high frequency repetition of laser pulses allows enormous volumes of data samples to be collected rapidly and provide a dense sampling of physical terrain. These benefits have lead LiDAR systems into wide-spread commercial use throughout the world with a large application base.

1.2 Commercial Application of LiDAR

LiDAR is typically used to provide digital terrain models (DTMs), although other applications do exist. DTMs were typically supplied in the past through terrestrial surveying observations, photogrammetric stereo pair analysis, and more recently through synthetic aperture radar (SAR). Terrestrial surveying, although accurate, is extremely time consuming and subject to suitable land access. Photogrammetric stereo pair analysis and synthetic aperture radar can both provide a quicker solution, although processing times are high, accuracy can be variable and point spacing is often sparse. LiDAR DTMs are able to achieve a point spacing of more than one point per square metre, are more accurate, and require less complicated processing routines. Considering these benefits, it is not surprising that airborne LiDAR has seen increased use in the commercial industry. [Maune, 2007] describes several commercial applications:

- Forestry – LiDAR observations have the unique ability to penetrate through small gaps in forest canopy cover and also to receive multiple returns form a single pulse. When observations are performed over vegetated areas the multiple pulse returns can be used to describe canopy structure, understory and forest floor. This gives an opportunity

for increased forest inventory analysis as tree heights and density of vegetation can be investigated. In addition, if enough pulse returns are received from the forest floor, a 'bare-earth' model of can be acquired. This gives an impression of the topography of the land if vegetation were removed. From this, improved hydrological analysis and watershed mapping can be facilitated.

- Coastal Engineering – Areas of limited contrast such as sandy areas prove problematic for determining DSMs (Digital Surface Models) through Photogrammetric approaches. In addition, inter-tidal zones often require timely and rapid observation to avoid tidal restrictions. These factors make LiDAR an excellent tool for acquiring coastal zone regions and aid in the management of environmentally sensitive coastal zones.

- Corridor or Right of Way Mapping – most LiDAR applications involve observing large areas of terrain with multiple flight lines. However, a need has also arrived to map linear corridors used to house gas pipelines, highways, railroads, utility poles and lines, and telecommunication equipment. This infrastructure requires surveys after completion or reconnaissance prior to construction. Usually performed with a helicopter as the platform is able to respond to sudden changes in corridor direction, LiDAR surveys can quickly obtain information for extensive corridors. The high reflectivity of power lines makes LiDAR an extremely effective tool for determining line sag, ground clearance, and encroachment. Alternatively these measurements typically require expensive and time consuming ground surveys.

- Flood Plain Mapping – one of the largest contractor's of LiDAR surveys in North America is the Federal Emergencies Management Agency (FEMA), a governmental organization in the United States tasked with managing the risk associated with artificial

and natural disasters. After the historic devastation caused by hurricane Katrina, FEMA realized the immediate need to establish Digital Surface Models (DSMs) of flood plains throughout the coastal United States. A large scale mapping project is currently being undertaken using LiDAR surveys for the majority of data collection. As a result of the project FEMA has also internally developed "LIDAR specifications for flood hazard mapping" which details standard practices and procedures required for detailed flood plain mapping using LiDAR.

- Urban Modelling – Three dimensional modelling of urban areas can be used for a variety of applications. Typically, geometric distortions and shadowing on high-rise structures causes disadvantages when attempting to model urban centres through SAR or Photogrammetric methods. LiDAR avoids these issues because of its dense point spacing and small footprint size. The models can be used for urban planning, disaster planning, line of site analysis for telecommunication equipment, and view shed analysis.

- Damage Assessment – LiDAR can be quickly deployed after a significant event has occurred damaging the environment or developed areas. The acquired DSMs can be used to assess damage for insurance purposes, aid in disaster management and reconstruction planning. A survey of this nature was undertaken shortly after the collapse of the World Trade Center in New York on September 11th, 2001.

This list proves that the technology has become an important tool in the remote sensing industry and has a large market potential. This point was furthered in Baltsavias [1999] and TMS [2005] which detail the growth of the LiDAR industry over the past two decades. Indications suggest the current trend will continue and the industry will continue to expand. This indicates that research into LiDAR technology will benefit from a wide

audience within the remote sensing community from the private, public and academic sectors.

1.3 Research Objective

The effect of terrain slope on LiDAR data has been documented as a significant sources of error [Morin ,2002], [Hodgson et al., 2005]. Comparison of reference data on slopes consistently shows larger discrepancies than on flat ground. Several authors have noted that theoretically there is justification for the existence of such errors [Morin, 2002], [Schaer et al., 2007]. To date, no attempt has been made to model these effects. This thesis will provide a methodology for predicting the magnitude of these errors. The developed algorithm will be tested to verify it is properly predicting the error. Test areas will focus on glacial surfaces which are characterized by heterogeneous terrain, absence of surface vegetation, and strong reflective properties. From this error analysis, research into the surface change of glaciers and predictions of their relationship with water run-off will benefit. This is of critical importance to management of water-resources, especially in those areas that rely largely on fall glacial run-off.

1.4 Significance of Research

LiDAR is a relatively new remote sensing technology. As mentioned, it required the addition of enabling technologies such as kinematic GPS and IMUs before it could reach its current potential. Commercial systems became available only in the early nineties with the first release of Optech's ALTM (Airborne Laser Terrain Mapper) [Maune, 2007].

Since then, the private industry has driven the market for LiDAR systems quickly, and although few organizations with the exception of FEMA and ASPRS (American Society for Photogrammetry and Remote Sensing) have responded with literature relating to errors, standards and best practices.

The largest source of information relating to errors is from system manufacturers such as Optech who quote system capabilities in commercial literature. Within these documents they do not provide conditions test data were observed under. Usually, only simplified quotes based on flying height are given. It is well known that LiDAR errors change dynamically as a result of several other survey parameters. Also, statistics are gathered under ideal survey conditions so that optimistic estimations of system performance can be used in advertisements. This makes it difficult for end-users that expect the system to provide similar results in a variety of survey settings. This misunderstanding can often lead to surveys that do not meet client's accuracy requirements. This is a costly consequence as supplemental information may have to be gathered that requires re-mobilization of aircraft and loss of sensor time. If users had access to appropriate error prediction models that considered all possible sources of error then they would be able to efficiently meet client specifications by designing appropriate flight plans and choosing appropriate survey parameters.

Since terrain slope is regarded as one of the largest sources of error in final LiDAR point positions, the proper modelling of its effects are a critical part of the complete error model. Although the research focuses at determining slope errors in glacial environments, it is pertinent to a broad cross-section of the LiDAR community as sloped areas will be encountered in most LiDAR surveys.

1.5 Previous Research

There is a consensus that similar to other positioning systems, several error sources contribute to the quality of final LIDAR point positions. These errors sources include terrain slope and aspect, terrain roughness, hardware system, calibration, laser incidence angle, laser beam divergence, and atmospheric effects [Lemmens, 1997], [Morin, 2002], [Schaer et al., 2007]. Extensive research has focussed on the determination and correction of systematic biases in LiDAR that result from several of these sources. Determination of systematic errors came as a result of ground strip matching which arbitrarily corrects flight strips to match relatively, with less attention on the global effect on the data and its comparison to external sources. The drive to properly match adjoining observation strips was natural as the usefulness of the data was limited if the terrain model was not continuous. Examples of strip matching procedures can be found in Filin and Vosselman [2003] and Vosselman and Maas [2001]. Since the solutions for strip matching didn't explicitly require error modelling, there has been less research into understanding the error phenomena behind the mismatch in strips.

Lemmens [2007] offers a list of all error sources within LiDAR observations, whether they were random or systematic and their approximate magnitude under six different ground classes. Although the list is useful in identifying all the error sources the magnitudes of error were not justified with theoretical or empirical evidence. Morin [2002] offers a partial list of error sources and magnitudes along with research for methods of determining angular boresight calibration values. While the study was

rigorous in the methodology and testing of the calibration parameters there was no evidence to support the error magnitudes of other sources. Schwarz et al. [2006] also investigated the error in boresight calibration and developed a methodology that included a least squares solution with error analysis.

Several studies have been done that attempt to identify errors in hardware components of the total LiDAR system such as Baltsavias [1999] and Schneck [2001] who focussed on systematic bias and Goulden and Hopkinson [2008] and Glennie [2007] who focussed on random error. The errors due to hardware components were found to increase at swath edges due to propagation of IMU errors, increased ranges, and larger pulse footprints. Schaer et al. [2007] added the effects of laser beam incidence angle to the errors in hardware components and noted that this error will increase on sloped areas depending on the vector of the laser pulse, although the effects of slope itself were not investigated.

Several studies have also been done that report observed errors between LiDAR observations and a reference surface. These studies give an impression of the total magnitude of error in the observations, but cannot separate the error source, error type, or give conclusions as to the consistency of the results. Several of these studies focussed on the error due to ground slope, however no predictive models were provided and the errors due to ground slope could not be effectively isolated from other error sources. The results still provide valuable information about the change in total error magnitude when a slope is introduced. Hodgson et al. [2003] isolated various slopes and reported the Root Mean Square (RMS) errors in Table 1.1.

Table 1.1 Terrain Slope vs RMS Error [Hodgson et al., 2003]

Slope	0-2°	2-4°	4-6°	6-8°	8-10°
Elevation Error (cm)	60	65	88	93	89

From this it was concluded that terrain slope was correlated with LiDAR elevation error and that increasing slope resulted in increased errors. They further predicted that larger slopes found in highly mountainous areas such as the Appalachians or Rockies would result in much larger elevation errors. Also found was that land cover is correlated with slope. Generally scrubs / shrubs and deciduous tree cover are found on highly sloped areas and this may be partially responsible for larger errors. Testing was performed to correlate errors between slope, surface features and the combination of slope and terrain coverage. It was concluded that there was a positive correlation between all three variables and a resultant increase in elevation error and terrain cover was found to be the dominant force behind the increase. However, no slopes greater than 8° were included in the study due to lack of reference data. Since error is known to increase with increased slopes there could be a point in which slope error becomes the dominant contributing factor to total elevation error which was not reached with slopes of only 8°.

Reutebuch et al. [2003] classified slope points into only two groups, those with slope less than 18% (Group one) and those with slope greater than 18 % (Group two). It was reported that there was very little correlation between slope and elevation error with mean errors of 21 cm and 22 cm for Group one and Group two respectively. Also included was a scatter plot, seen in Figure 1.1 showing the spatial distribution of the elevation error compared to slope.

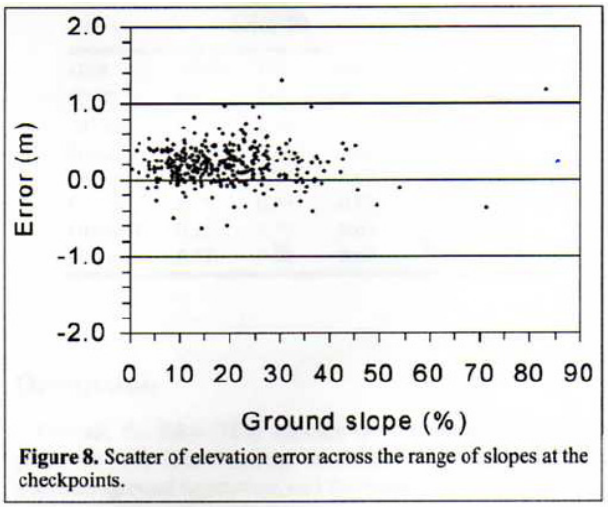


Figure 8. Scatter of elevation error across the range of slopes at the checkpoints.

Figure 1.1
Scatter Plot of Ground Slope vs RMS Error from [Reutebuch et al., 2003]

Unfortunately, there was no attempt to isolate the slope from other types of variables such as canopy cover or terrain roughness characteristics. Reference slopes were calculated from gradients along a DSM calculated with the inverse distance squared function which may have had an effect of over-smoothing the topography and biases reference data [Reutebuch et al. 2003]. Also mentioned is that mean and median of the slope for all measured points are 18% and 19% respectively. It appears from the scatter plot that most of the data falls between the 10% and 25% slope categories. Perhaps choosing 18% slope as the divide between the two groups was not an adequate choice since the mean error in each group could be biased toward the slope values that represented the majority of data. Choosing smaller variations of slope and creating more groups may have identified more descriptive results.

Huising and Gomes [1998] report in a table that originated from Lemmens [1997] that terrain slope causes centimetre level systematic errors on sloped surfaces in various

land covers. However, there is no detail given as to the degree of slope from which this statement is made and only reports that it is under the condition of paved surfaces, barren land, grass and shrubs.

Among the existing research it appears as if the effects caused by terrain affect the overall accuracy of the LiDAR observations although its causes are not well understood. Within the presented research the error increases with terrain slope although a definite correlation has not been established. Also, extreme slopes have not been investigated. This could be a result of the difficulty obtaining reference data. Once investigated, it's possible extreme slopes will dominate all other error sources. Although these previous studies pertaining to LiDAR data did not show strong correlations between slope and error it is worth noting that this phenomena has been identified in Maling [1989] in a discussion of DSMs, regardless of observation equipment.

Further studies have shown how the uncertainty in DSMs affected their properties such as slope and aspect [Fisher, 1991], [Florinsky, 1998], [Heuvelink et al., 1989]; [Oksanen and Sarjakoski, 2005]. Although these studies can't be used to predict error based on slope they provide information on the dependence between error in point positions and the errors in slope. These relationships aid in forming proper models of the DSM error. Also, they show the importance of properly formulating error models in DSMs so that the quality of information that is extracted from the DSM can be properly assessed.

1.6 Proposed Method

Hodgson and Bresnahan [2004] diagrammed their opinion on the nature of error due to terrain slope on LiDAR observations in Figure 1.2. It is apparent from the figure that problems are caused when a LiDAR point contains horizontal error. The horizontal error causes the apperent *X-Y* position of the LiDAR point, represented by the black dot to be shifted horizontally. The elevation value associated with the apparent position will be referenced to where the laser enegy reflected from the terrain, represented by the arrows. This causes a preceived elevation error of the apparent LiDAR point. Therefore, we can view this error as having one-dimensional component in the vertical direction.

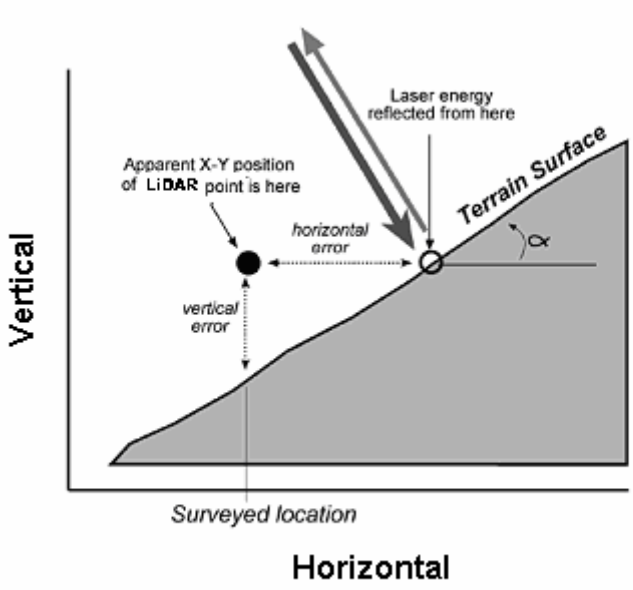


Figure 1.2

Vertical error caused by horizontal error on sloped terrain.

With some a-priori knowledge about the size and orientation of the error associated with the apparent *X-Y* position of the LiDAR point, we can statistically predict a limited

space the real terrain is within. With this knowledge, a probability surface of the real terrain be generated and the difference between the elevation of the LiDAR point and the surface can be used to define the vertical error. Figure 1.3 displays this phenomenon. The solid black line represents the terrain as determined by LiDAR, the dashed gray line indicates two possibilities of where actual terrain based on the predicted error (grey ellipse) in the LiDAR observation. The vertical difference between the elevation of the LiDAR point and 'probability surface', represented by the vertical dashed black line, will represent the predicted vertical error in the LiDAR observation.

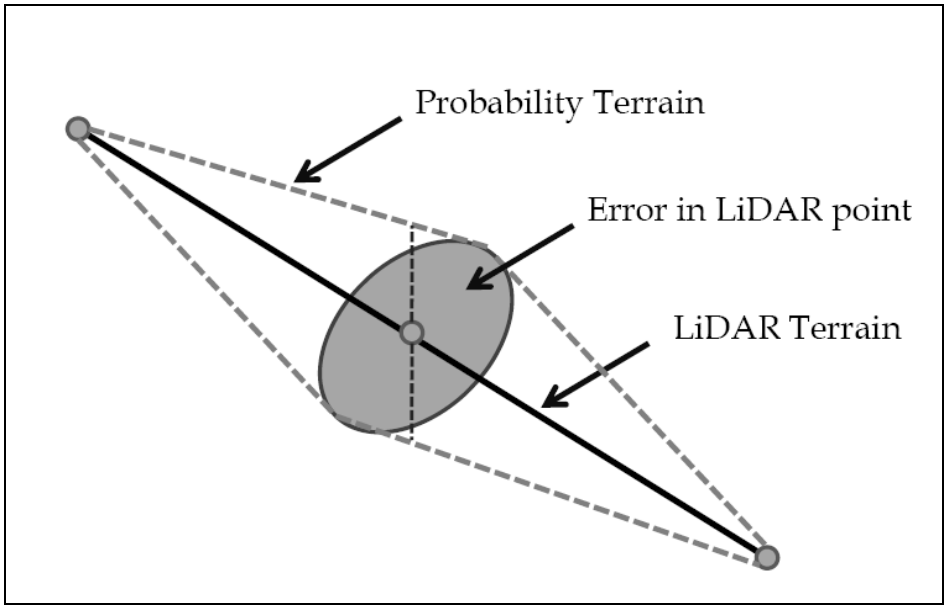


Figure 1.3

Calculation method for determining vertical error based on terrain slope.

It becomes apparent that with the appropriate information the problem can be solved geometrically. Prior to formulating the geometrical relationship we must have knowledge

of the initial error in the LiDAR observation and determine how to construct the probability surface. The initial error in the LiDAR point can be calculated from the combined error of the hardware components in the LiDAR system. A foreknowledge of how these components interact and their respective error quantities can be used to propagate the error into final point positions. Studies of this nature have been undertaken by Goulden and Hopkinson [2006] and Glennie [2007]. The methods described in these studies will be used to determine the initial point position error.

Although there are an infinite number of possible surfaces that could represent the true terrain it is prudent to determine the surface that will create the most pessimistic results, often termed the 'worst-case-scenario.' In error modelling this is advantageous as it ensures errors will be less than predicted. Although this will tend to bias predicted error quantities to be more pessimistic than actual results, this is considered to be more favourable than optimistic predictions. Under these conditions if results show that errors are higher than predictions it can be assumed other sources of error have contaminated results and will need to be considered separately. With the constructed surface and initial error estimate the geometrical relationship can be used to solve for additional error caused by the slope in the terrain.

1.7 Thesis Outline

Chapter 2 will review the error characteristics of various components of a combined LiDAR system. This will give the reader the appropriate background information to understand the flow of errors from instruments to final point positions. From this, details

on how the user can modify survey settings to affect the outcome of surveys can be interpreted. The different components of the system that will be discussed are: GPS, IMU, laser scanner and laser ranger. Following their descriptions, how the components are integrated to determine final point positions will be explained with a direct georeferencing mathematical model. A brief outline of the propagation of errors through this model to determine point position error will be included to describe the process of determining initial uncertainty values. Also, suggested error estimates for each of the hardware components are included.

Chapter 3 describes unique characteristics of LiDAR DSMs such as point spacing and obtainable accuracy. Comparisons will be drawn between existing DSM models and how the properties of LiDAR DSMs make the appropriate error prediction more critical to final products. Through error propagation an analysis of how the quality of slope measurements are affected by properties of LiDAR DSMs will be carried out. It will be shown graphically how the properties of traditional DSMs made error considerations negligible, while the properties of LiDAR DSMs make error considerations significant.

Chapter 4 will describe how the model for determining errors based on slope is formulated. It begins with the calculations for the slope between each point in the dataset. This is facilitated by the creation of the TIN model using the Delaunay triangulation method. After slope calculations the determination of a probability surface using the *a – priori* error knowledge is described. From the probability surfaces and original LiDAR points the vertical error calculation is given. Output from simple simulated terrain models is given for descriptive purposes.

Chapter 5 begins by explaining the test data sets. The first test data set comes from a ski hill in Windsor, Nova Scotia and the second data set from Windsor Street near the University of New Brunswick in Fredericton, New Brunswick. The survey parameters, flight conditions and terrain characteristics will be outlined so that when results are presented they can be cross-referenced. The procedures for acquiring reference data and how they were used to generate results are described. Following this, a description of the tests that were used to determine the validity of the model is provided. These tests include a test for normality of the acquired data set, testing the percentage of residuals falling within predicted ranges, and statistically testing the compatibility the a-priori predictions with the variation in residuals. Following this, comments on the results and their applicability to the study will be mode.

Chapter 6 offers conclusions about the project. It provides a synopsis of the entire project including a description and interpretation of results. A section on how this work contributes to the larger body of LiDAR research and how it can be used within the project workflow of LiDAR surveys is included. Finally recommendations to correct deficiencies in the current study and for future work are given.

Chapter 2

Determination of Initial Error Estimation

The initial error estimate for the slope calculations is a combination of errors that are introduced from observations of the hardware components in the LiDAR system. A review of the errors in each hardware component is given here to facilitate an understanding of how each component contributes to a combined error estimate. The hardware components that are going to be reviewed are the GPS, IMU, laser scanner and laser ranger. The combination of the hardware components observations into a direct geo-referencing equation forms the mathematical model used for error propagation.

The General Law of Propagation of Errors (GLOPOV) is applied to the mathematical model in combination with individual hardware component error estimates to provide error predictions for each acquired position. Error estimates applied in the formulation must be a result of random processes as systematic biases cannot be propagated in this manner. Therefore, the error discussion has a focus on random sources of error and will not consider systematic biases. This will not adversely affect initial error predictions for the slope calculation as systematic biases would affect the terrain similarly in a local area resulting in identical slope calculations and error predictions. Once formulated, fully propagated errors will take on the form of three dimensional error ellipsoids represented by a fully populated covariance matrix. In-depth details on the quantification of individual error components, functional covariance propagation of the LiDAR direct geo-

referencing formula, and testing of the error predictions, can be found in Goulden and Hopkinson [2008].

2.1 Global Positioning System (GPS)

The GPS antenna is typically located on the roof of the aircraft. It receives signals transmitted from satellites orbiting the earth in order to determine its spatial location within the WGS-84 datum. The system is dependant on redundant information from multiple satellites to accurately determine its position. Random error in the position received from the GPS antenna is affected by many factors including the number of visible satellites in the sky, and their position relative to the receiver. Satellite geometry and the number of satellites are often represented by a value termed the position dilution of precision (PDOP). The PDOP value is commonly used in the GPS industry to describe the predicted precision of the GPS observations. Users should expect that high PDOP values will lead to solutions with higher instances of random error and poorer results. Optech specifies that at least four satellites and a PDOP of less than four must be observed continuously throughout the entire survey to guarantee their accuracy specifications are met (Optech, 2004). Geometric dilution of precision (GDOP) can also be used to describe the quality of the GPS observations, users familiar with this term can generally expect similar results with larger GDOP values.

In addition to satellite geometry, GPS observations are also affected by atmospheric influences within the ionosphere and troposphere. To minimize these effects, surveys are post-processed through double differencing techniques. This is achieved by setting up a base station over a known survey control point and then processing the base station data

simultaneously with the data from the receiver mounted on the aircraft. Since these atmospheric effects usually vary little over short distances they are virtually eliminated when differencing the data gained at both the base receiver and the aircraft receiver. If the survey area is in excess of 30 km from the base station the atmospheric effects can begin to have noticeable detrimental effects on the accuracy of the data [Mostafa, 2001].

2.2 Inertial Measurement Unit (IMU)

The IMU sub-system within the integrated LiDAR system measures the instantaneous acceleration and angular motion of the vehicle. This is done through a strap-down system that contains three gyroscopes and three accelerometers. A strap-down system is hard mounted to a surface and rotates with the platform. Observations are used to determine the rotational parameters of exterior orientation of the vehicle, and to fill in gaps of position information between GPS data epochs and during periods of loss of lock.

A portion of the error in the IMU is dependant on the initialization parameters which are defined when the unit is first activated Schenk [2001]. During this period, the three axes of the IMU three-dimensional coordinate system are defined. The most important of the three are those which orient toward the sensed direction of local gravity and the direction of astronomic north. The third axis is set to complete a right handed system. During the initialization process various external factors such as temperature variations can contribute to the random error of the system [Müller et al., 2002].

The initialization process has three stages; coarse levelling, coarse heading alignment and fine heading alignment. The coarse levelling resolves roll and pitch error in the IMU frame to one to two degrees. During coarse heading alignment, the heading is roughly

determined in the GPS reference frame. Finally, the three attitude parameters are precisely resolved for best performance in fine alignment mode [Mostafa et al., 2001]. The IMU frame must be referenced with that of GPS to coordinate the final observations. This is realized by the implementation of a transformation by an amount equal to the deflection of the vertical at the initial point [Schwarz, 1983], [Schenk, 2001]. It is presumed that a global geo-potential model is utilized within the processing of the positioning and orientation software to properly account for geoidal undualtions. Although there is inevitably some uncertainty resulting from its use its details are proprietary and therefore could not be included in the analysis. Vaughn et al. [1996] claim that the systematic error introduced by ignoring this effect is rarely significant, from this it is assumed the same is true for the random component.

The second significant factor that contributes to the uncertainty of the IMU is drift. Since an IMU calculates position by double integration of acceleration data, the accuracy of the integration algorithm is critical to the overall uncertainty. Through time, errors occurring in the integration process will accumulate. The longer the unit is left running without an independent correction the larger the random error will become [Lee, 2004]. In an integrated system this problem is partially overcome by including GPS observations as an independent reference. Subsequently the drift will only accumulate between GPS observations and during periods of loss of lock to the GPS satellites. Generally, if the GPS signal is temporarily lost then on the fly (OTF) cycle ambiguity calculations are required once signals are reacquired. During this period, the errors in the GPS system can reach metres and will do very little to bound the accumulated drift of the IMU. During

this period the entire LiDAR system is compromised and LiDAR observations must be suspended until cycle ambiguities are resolved.

2.3 Laser Ranging System

Typically solid state lasers are used in the ranging unit of commercial LiDAR systems because they can transmit at high frequencies with high power output [Wehr, 1999] and are compact in size with long lifetimes [Barnes, 1986]. LiDAR systems typically transmit energy in the near infrared (NIR) portion of the electromagnetic spectrum. This wavelength is selected for its propagation characteristics through the lower atmosphere, range resolution, eye-safety considerations, and to avoid ambient noise reflected by solar radiation in the visible spectrum [Cruickshank and Harney, 1986].

Upon reception of the laser signal an optical telescope is used to collect and direct the returned energy to an avalanche photo-diode. The avalanche photo-diode (AVD) photon detector measures the incoming photons of energy as a function of voltage and time. Avalanche photo-diodes have been widely accepted within the LiDAR industry for their high sensitivity, high-speed detections, and quantum efficiencies [PerkinElmer, 2003]. To place the sensitivity of the AVD into perspective imagine a pulse that is generated with two kW (kilowatts) of energy and has a two way travel distance of 1500 metres. If the pulse makes contact with a surface with good reflective properties approximately 2.6 x 10^{-6} W of energy will be returned [Baltsavias, 1999].

Recall that the laser range is converted into a distance based on the reception time of the laser pulse. By graphing the returned voltage of the laser pulse against time the required stop time can be interpreted. Details for determining stop times are often considered proprietary information by LiDAR manufactures. Vaughn et al. [1996] describe a signal processing technique in which a fixed signal threshold is used with a correction that is based on the return intensity of the pulse. Keeping all other system settings equal, the combined system accuracy improves with peak pulse power [Hopkinson, 2007], and therefore it can be deduced that so does ranging accuracy.

In addition to the time, typical LiDAR systems can also record the peak intensity values returned from targets. This value represents the reflective properties of the target and provides additional information about target properties. LiDAR intensity images are similar to RADAR images in that they represent the amount of backscattered energy. For example, still water typically absorbs energy in the near infrared portion of the electromagnetic spectrum leading to low or non-existent intensity values. These are termed “dropouts” since not enough energy is returned to calculate a suitable stop time for the pulse. On the contrary, targets that reflect energy well such as ice or snow can inundate the sensitive AVD. A sudden spike in intensity will cause the system to stop the time measurement too early and register a short range. Subsequently the coordinate point will be misplaced from its true position in a phenomenon termed ‘range walk’. To partially overcome this error the AVD contains a real-time gain control. An intensity image of a portion of the University of New Brunswick campus can be seen in Figure 2.1. The image is greyscale and displays high intensities as white and low intensities as black.

From this, it is apparent that high intensity returns are generated from vegetated areas while low intensities are generally returned from building roofs or asphalt.

Figure 2.1

Intensity Image of a portion of the University of New Brunswick campus

The most obvious source of hardware related error in the laser ranging system is contributed by the time interval meter (TIM) and its ability to identify the start and stop time of the pulse. The clocks used in measuring the two way travel time of the pulse are accurate to approximately ±0.1 nanoseconds [Baltsavias, 1999] leading to range errors of approximately 2 cm that are independent of the magnitude of the range.

Atmospheric affects are also known to contribute systematic and random sources of error into a laser range observation [Marini and Murray, 1987]. The size of these errors will be directly proportional to the length of the range. The error will depend on the atmospheric conditions (e.g. air temperature, humidity and pressure) during the survey and how well the measurement of temperature and pressure on the ground and in the air can be performed [Vaughn et al., 1996].

2.4 Scanning Unit

The scanning unit is a mechanism within the integrated system that measures the scan angle of the mirror as it directs outgoing pulses across and beneath the vehicle trajectory. The ALTM 3100 scanning mechanism is an oscillating mirror. The oscillating mirror scanner is centred on a rotation point and mechanically oscillates back and forth directing the laser beam in a measured direction. This creates scan lines known in the industry as "saw tooth" in which every second line is parallel which can be seen in Figure 2.2. The system is driven by a galvanometer that controls the amount of angular motion of the mirror.

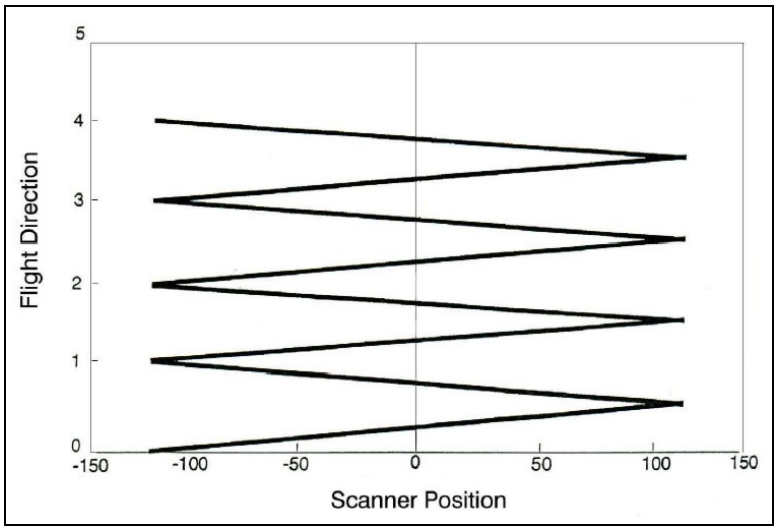


Figure 2.2

Saw tooth patters of laser pulses created by an oscillating mirror

The angle is measured with an internal laser scanning mechanism which reads an etched glass plate located on the side of the rotating cylinder. The etches on the glass

plate form a unique pattern similar to a bar code that specifies the angular distance the glass plate has been rotated. The uncertainty results from the bar code reader's ability to correctly interpret the bar code. This ability derives from the construction and design of the angular encoder. Variations in temperature and vibrations within the external environment can affect the random error of the system Reinshaw [2005], Heidenhain [2006]. Additional systematic effects in the scanning mechanism result from torsion in the rotating cylinder caused by the varying angular accelerations during scanning and encoder latency, a mis-registered timing scheme between laser pulses and scan angle observations [Morin, 2002].

In addition to the error due to the physical scanning mechanism mentioned there is also an additional error in scan angle due to the effects of beam divergence. Immediately after a pulse is transmitted it begins to expand. The area of the pulse upon impact is dependent on the flying height of the aircraft and a beam width mode that can be selected within the hardware. The ALTM 3100 has two beam mode widths that can be selected, either 0.3 mrad (narrow beam) or 0.8 mrad (wide beam). These values indicate the area of the pulse (footprint) that will illuminate the ground upon impact and contain approximately 68% (1 sigma) of the total pulse energy. It is theoretically possible that the returned distance is from anywhere within the footprint. However, it is assumed that the return comes from a vector that begins from the transmission point and in the direction defined by the scan angle at the centre of the beam footprint. The ambiguity of the location of the returned distance within the footprint creates an additional uncertainty.

Glennie [2007] identifies that the pulse energy is at its peak in the centre and becomes weaker in an approximate Gaussian distribution. Figurere 2.3 displays how

pulse energy would look theoretically. It is intuitive that with power levels at their highest in the centre, strong returns would generally be identified there. Glennie [2007] indicates that the uncertainty introduced by the beam divergence will generally be one quarter of area defined by the area that contains nearly 100% (3 sigma) of the pulse energy.

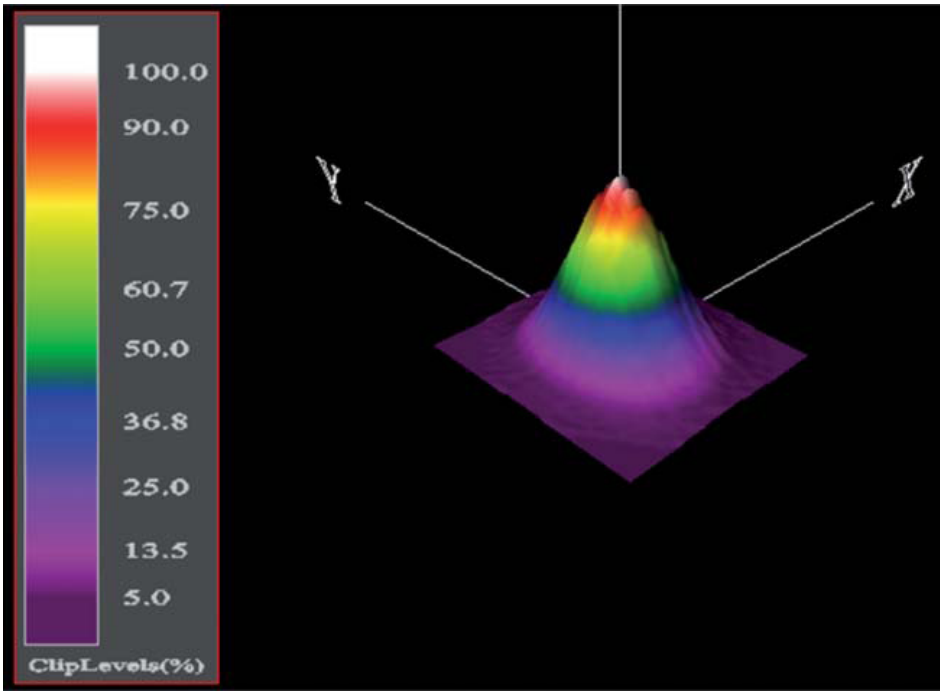


Figure 2.3

Approximate output power distribution of ALTM 3100 laser pulse

[Glennie, 2007]

2.5 Measurement System Integration

When considering the total error within the hardware components consideration must be given to combining the system measurements. The location of the GPS antenna is typically on the roof of the aircraft and it is necessary to determine positional coordinates

of the IMU unit and the scanning mirror with respect to the GPS antennas. To register the position of these three pieces of equipment, linear distances between them are measured. This can be done through terrestrial surveys or during an in-flight calibration procedure [Vaughn et al., 1996], [Schwarz et al., 1983]. Both procedures will contain some random error in their calculations that must be taken into account in the estimation of the total propagated error. Strictly speaking, relative positional integration errors are systematic in nature. However, they typically constitute an unknown magnitude and direction of error that varies slightly each time the LiDAR system is initialized because of local deflection of the vertical components and initialization parameters. Therefore we must treat these affects as random error sources.

To properly integrate the IMU and laser scanning reference frames, rotational offsets must be calculated through in-flight calibration procedures [Vaughn et al. 1996], [Schenk, 2001], [Klaus and El-Sheimy, 2004], [Schwarz et al., 1993], [Cramer, 1999]. Since LiDAR observations that inherently contain all sources of error must be used to determine the values, the random errors associated with them cannot be isolated. It is assumed that the random errors associated with these values can be accounted for in conservative estimates of the IMU roll, pitch and yaw values since they are similar in nature.

Time stamping represents another important source of the uncertainty resulting from measurement system integration. The primary source of the time stamp is the GPS time obtained from satellites for the GPS and IMU components (collectively referred to as the position orientation system – POS), or an internal computer clock within the LiDAR sensor. It is possible; therefore, that internal temporal drift in the electrical circuitry,

inaccurate time stamping of measurements, or inaccurate temporal registration between system component time stamps could lead to a level of uncertainty in the final integrated solution. However, because such intricate knowledge of the system temporal integration methodology is usually highly proprietary and since the integration of components is precisely calibrated by sensor manufacturers, this aspect of the propagated system error must be either ignored or partially accounted for by adopting conservative uncertainty estimates in the individual system components. Figure 2.4 from Goulden and Hopkinson [2008] diagrams the integration of components through a flowchart.

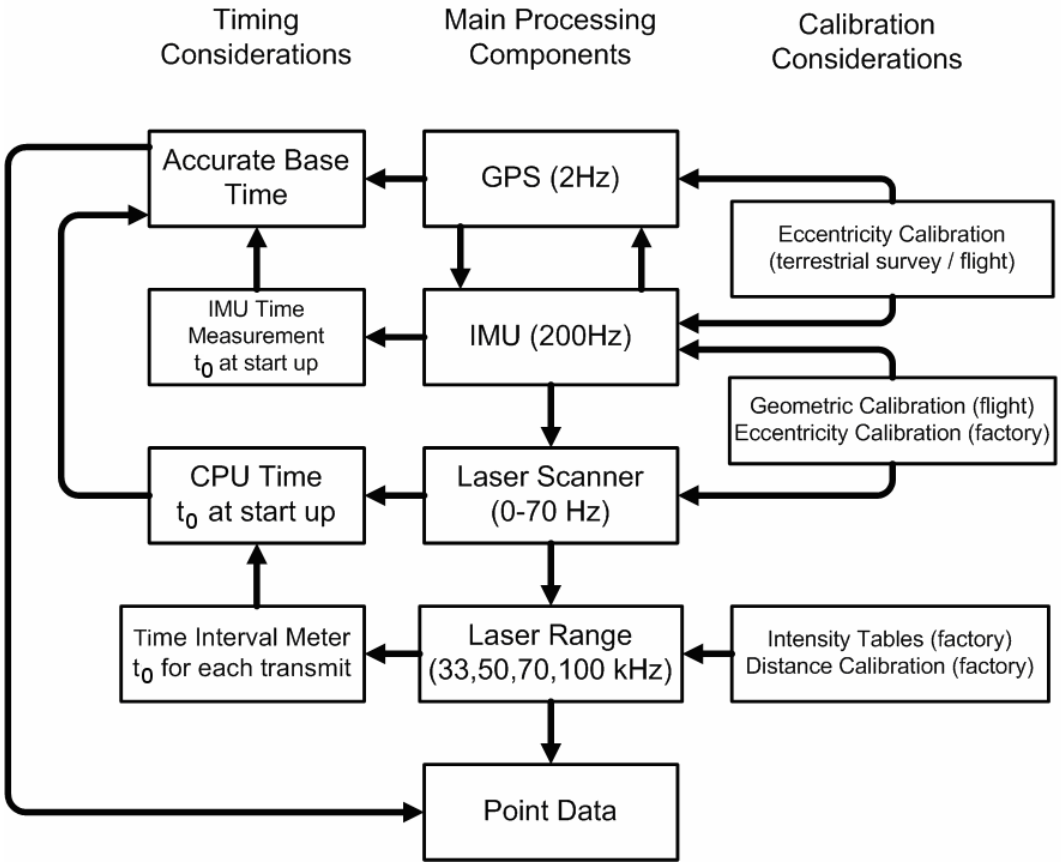


Figure 2.4

Flow chart Diagramming LiDAR System Components from [Goulden and Hopkinson, 2008].

The main components listed in Figure 2.4 can be seen in the centre column of the flow chart. They represent the observational components of the system. Also included with the components are their respective observation rates. The final block culminates in the combination of all the system components to produce the observed point data. The integration of the components is included on the right and on the left side of the flowchart, representing calibrations and time synchronizations, respectively. Literature relating to bore-sight misalignments between the IMUs and imaging devices is readily available as it has been an issue for the direct geo-referencing of aerial imagery. Information on the misalignments can be found in Cramer [1999], Shwarz et al. [1983] and for specific LiDAR applications in Morin and El-Shiemy, [2002], Skaloud and Lichti [2006], Schenk [2001], Vaughn et al. [1996] and Filin [2003].

2.5 Formulating Initial Error Prediction from the Direct Georeferencing Equation

The direct geo-referencing equation for LiDAR observations was first published in Vaughn et al. [1996]. It includes all of the aforementioned hardware component observations and reads as follows:

$$\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} = \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \end{bmatrix} + R_w R_G R_{INS} \left(\begin{bmatrix} \delta_x \\ \delta_y \\ \delta_z \end{bmatrix} + R_m R_s \begin{bmatrix} 0 \\ 0 \\ -\rho \end{bmatrix} \right) + \begin{bmatrix} \overline{e}_x \\ \overline{e}_y \\ \overline{e}_z \end{bmatrix} \tag{2.1}$$

Where x_1, y_1, and z_1 represent the location of the observed point, X_0, Y_0, and Z_0 represent the location of the phase centre of the GPS antenna, R_w is the rotation matrix from the WGS-84 datum to a local ellipsoidal reference frame, R_G is the rotation from the local gravity frame to the ellipsoidal frame, R_{INS} is the rotation from the body frame to the local gravity frame, δ_x, δ_y, and δ_z are the offsets between the laser transmission point and the phase centre of the GPS antenna in the body frame, R_m is the boresight angular values which rotate between the body frame and laser scanning frame, R_s is the rotation by a value equal to the observed scan angle, ρ is the observed range observation and e_x, e_y, e_z are the random error components of the observation in the same reference frame as the laser point. [Vaughn et al. 1996], [Schenk, 2001], [Filin, 2003]. Figure 2.5, adapted from Habib et al. [2006], displays the rotation frames involved in the direct geo-referencing. Where subscript '*b*' represents the body frame, '*lu*' represents the laser scanning frame and '*lb*' represents the laser beam frame the range is initially transmitted within.

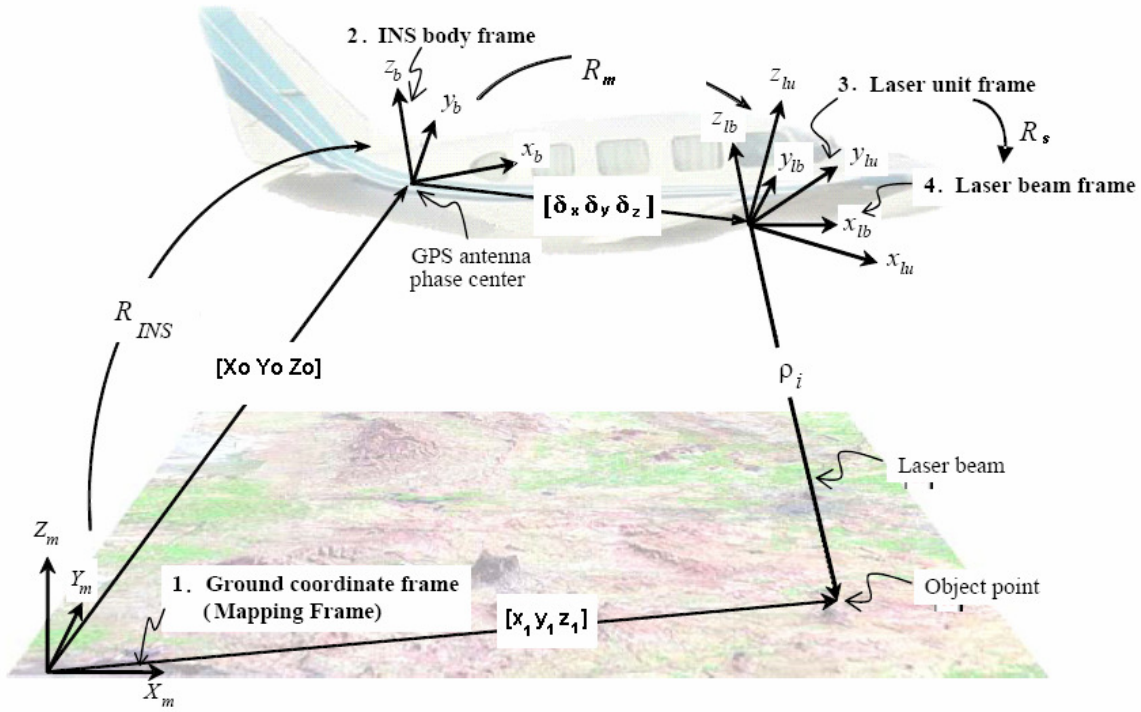


Figure 2.5
Direct geo-referencing of LiDAR observations. Adapted from Habib et al. [2006].

When the functional covariance propagation is performed it is in search of values to represent e_x, e_y, e_z. These will be used as the initial error predictions in the slope model. These values are treated as purely stochastic, meaning they are random in nature and it is assumed that they contain no systematic bias. Therefore we cannot solve for their true values, only statistically predict a probable range, represented by an uncertainty, they will be within. The uncertainty must be accompanied by an associated confidence interval which indicates how often observations will fall within the range. To calculate the probable ranges the well-known general law of propagation of variances (GLOPOV) can be implemented [Wolf and Ghilani, 1997], [Mikhail and Gracie, 1981]. Consider the linear equation:

$$y = Ax + b\,, \tag{2.2}$$

where y is a random vector containing randomly distributed variables, A is a matrix of coefficients and b contains constants. The uncertainty in the quantities in y can be expressed as:

$$\Sigma_{yy} = A\Sigma_{xx}A^T\,, \tag{2.3}$$

where Σ represents a covariance matrix containing the uncertainty values. From this, the law can be described as the sum of the squares of each observable quantity, multiplied by the observation's uncertainty, will result in the total propagated uncertainty when errors are contributed from statistically independent observations. When dealing with non-linear equations the law appears as follows:

$$\Sigma_{yy} = J_{yx} \quad \Sigma_{xx} \quad J_{yx}^T\,, \tag{2.4}$$

where J is termed the *Jacobian* matrix and consists of the partial differentiation for all quantities we are predicting and each observable. Partial differentiation is done to linearize the non-linear equation allowing it to be solved [Mikhail and Gracie, 1981]. In general form the two matrices will appear as follows;

$$J = \begin{bmatrix} \frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial x_2} & \cdots & \frac{\partial F_1}{\partial x_n} \\ \frac{\partial F_2}{\partial x_1} & \frac{\partial F_2}{\partial x_2} & \cdots & \frac{\partial F_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial F_3}{\partial x_1} & \frac{\partial F_3}{\partial x_2} & \cdots & \frac{\partial F_3}{\partial x_3} \end{bmatrix} \quad \text{and} \quad \Sigma = \begin{bmatrix} \sigma_{x1}^2 & \sigma_{x1x2} & \cdots & \sigma_{x1xn} \\ \sigma_{x2x1} & \sigma_{x2}^2 & \cdots & \sigma_{x2xn} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{xnx1} & \sigma_{xnx2} & \cdots & \sigma_{xn}^2 \end{bmatrix}, \tag{2.5}$$

Where F is a function, x are the observables and σ represents the uncertainty. The function (F) being analyzed is (2.1), the Σ_{yy} contains the uncertainties of the solution and includes final x, y, and z coordinates and the Σ_{xx} contains the uncertainties for each individual hardware component including: roll, pitch, yaw, x mirror coordinate, y mirror coordinate, z mirror coordinate, scan angle, and range.

The covariance matrix of the solution, Σ_{yy}, will be fully populated including co-variances (off diagonal elements) between each of the final coordinate errors resulting from the dependence of each coordinate solution to the common observations. These co-variances must be appropriately taken into account to calculate correct estimates of the total random error in each dimension. The covariance matrix of the observations in this case can be simplified to:

$$Cov_{Observations} = diag\begin{bmatrix} \sigma_{x1}^2 & \sigma_{x2}^2 & \cdots & \sigma_{xn}^2 \end{bmatrix}, \tag{2.6}$$

as the observations from each hardware component are statistically independent from each other since they are generated from individual sources. It should be mentioned that this is not explicitly true for the covariance values of the position and orientation system as in a tightly coupled system they are dependant. However, values for populating the positioning and orientation uncertainties in the covariance matrix of observations are taken directly from the software that calculates the aircraft trajectory during a survey. The software does not output appropriate covariance values and therefore they are assumed to be zero. Consequences of this assumption are that these error estimates will be optimistic and not in their proper orientation. Despite these drawbacks, Hare [2001] and Grejner-Brzezinska and Wang [1998] indicate it is the most appropriate method for acquiring error estimates along a trajectory of this nature. Further details on the error estimates of

the positioning and orientation system and their effect on the final point position errors can be found in Goulden and Hopkinson [2006].

Table 2.1 details the values that are chosen for propagating the errors into final point position. Error estimates for the position and orientation system will vary throughout the survey depending on satellite availability, satellite geometry and the flying conditions. The quantities that have been reported in Table 2.1 for these errors (GPS and IMU) are typical values obtained from the Applanix software used for processing the trajectory. Interesting to note is that these quantities reflect the achievable error reported in commercial literature with the exception of the error in yaw which is approximately twice the reported value.

Table 2.1
Error Parameters for hardware components

OBSERVATION	ERROR MAGNITUDE
X coordinate	3 cm
Y coordinate	3 cm
Z coordinate	5 cm
Roll	0.005 degrees
Pitch	0.005 degrees
Yaw	0.02 degrees
Scan Angle	10.6 seconds
Laser Range	2 cm

Beam Divergence	~0.3 mrad

Armed with the knowledge of the system related errors and the formulas developed from the functional covariance propagation it is possible to calculate initial estimates for point precision. Figure 2.6 displays an output of the error predictions across a portion of single scan line. The predicted error is at a minimum when displayed in blue and a maximum when displayed in red. It becomes immediately apparent that the largest errors will appear on the edges of the scan lines and the smallest at nadir. This is a result of the errors propagating more significantly with increased ranges and scan angles.

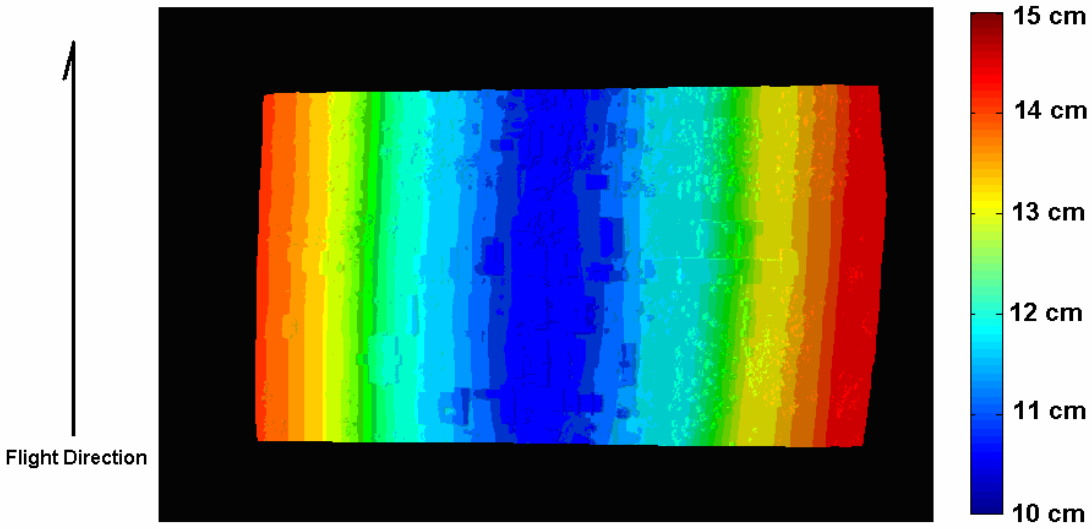


Figure 2.6
Change in predicted hardware errors along a single scan line

Chapter 3

LiDAR DSM Characteristics

Digital Terrain Models are discrete representations of the physical terrain. They consist of triplets of *x,y,z* points, often at regular intervals in the *x,y* plane. A greater number of coordinate triplets will create a dense and more realistic representation of the terrain. However, large data volumes will always come at the price of more complicated and time-consuming processing routines, increased storage requirements and a need of powerful hardware for manipulating and viewing data. Technically, the term DTM is reserved for representations of the 'bare earth' terrain, void of surface entities such as buildings or trees. The more specific digital surface model (DSM) is reserved for describing terrain with surface features included [Maune ,2007].

There are several methods for obtaining DTMs in practice. Originally, DTM representations were spot heights observed with traditional land surveying equipment. Although probably the most accurate, this is also the most time consuming method available. It has a place for accurate small scale DTMs or for quality control and quality assurance procedures. Typically, this method is not considered for large scale DTM acquisitions being discussed here. Systems in commercial use today that can provide large scale DTMs with similar point spacing and levels of accuracy are photogrammetric stereopairs, interferometric SAR (IFSAR) and LiDAR. Each of these systems require airborne or space borne platforms with sophisticated positioning and orientation techniques, either on board the aircraft or by rigorous ground control techniques.

Although all the technologies can obtain similar levels of point spacing and accuracy, LiDAR DSMs have several advantages over the competing technologies:

- Multiple returns from single pulses have the ability to observe multiple layers in vegetated areas including canopy structure, understory and true ground. This makes a better representation of bare earth DSMs possible.
- Determination of elevation is not dependant on the texture of the terrain which is required for aligning features in photogrammetric stereopairs. This is especially significant in glacial areas with significant snow and ice coverage where differences in texture are extremely slight yet laser reflectivity is strong [Hopkinson and Demuth ,2006].
- The highly collimated laser beam is less sensitive to obstructions in off-nadir observations which traditionally cause shadowing or geometric distortion in photogrammetric or IFSAR observations.
- The project workflow from raw data capture to data delivery is much faster due to the direct observation of elevation data without the need for complex processing routines to extract elevations.
- Compared to photogrammetry, which requires a passive observation system ,missions are less dependant on incumbent weather conditions.
- There is an ability to obtain elevation values on features that are smaller than the footprint size such as power lines which would not be visible in optical images [Baltsavias, 1999] and would not generate returns from high wavelength SAR observations.

In addition to the aforementioned advantages, there are also several disadvantages that require consideration

- The swath width of typical LiDAR systems, typically 0 to 20 degrees off nadir in each direction, is narrower than the field of view in aerial photographs and IFSAR systems. To cover equal areas more flight lines, and subsequently more flight time and cost is required.
- LiDAR observations rely heavily on accurate GPS and IMU measurements since each individual pulse is calculated independently and falls in a psudo-random pattern on the ground. This makes quality assurance and quality control more difficult as isolating suitable comparison structures is not straight forward.
- When performing filtering of raw data, photogrammetry has the advantage of high resolution optical images for comparison. From analyzing raw LiDAR data point clouds it is difficult to identify and classify features for removal.

Considering the advantages and disadvantages mentioned above, LiDAR proves to be an exceptional tool for observing DSMs. This chapter will briefly introduce LiDAR DSM characteristics as several features of these DSMs are important when considering the derivation of error models. Focus will be given to how mission planning characteristics will affect the final LiDAR DSM point spacing and point quality.

There are several quantities that are commonly derived from DSMs for analysis such as slope and aspect. To interpret how the characteristics of LiDAR DSMs affect the quality of derived quantities the calculation of slope from simulated LiDAR scenarios is presented. Slope was chosen as a test calculation because it is one of the most common

properties that is obtained from a DSM and because of the relationship with the larger body of work presented here. It is shown that special consideration must be taken when reporting derived DSM characteristics and performing an error analysis with LiDAR DSMs.

3.1 Mission Planning

Effective mission planning of LiDAR surveys is critical to ensure final DSM products provide the required information and are of sufficient quality. Planning LiDAR missions consists of choosing several different types of hardware specifications and designing a flight plan. The decisions made will affect the outcome of the survey in both point spacing and accuracy. Each of the choices that are left to the instrument operator are discussed here to give the reader an impression of how dependant the final DSM product is on these initial selections and how one can control the final DSM characteristics. There are noticeable differences from photogrammetric missions where there are very few options available to modify the outcome of the DSM since cameras have a well defined planer field of view that cannot be easily manipulated. Focus is given to the LiDAR ALTM 3100 system manufactured by Optech since it was used in the survey missions reported within his study.

3.1.1 Pulse Repetition Frequency (PRF)

The pulse repetition frequency (PRF) defines how many laser pulses are going to be transmitted in one second. When LiDAR systems were first commercialised the PRF had a maximum value of 2 kHz, allowing for 2000 pulses per second [Maune ,2007]. As

systems developed this number has increased and currently sits at 167 kHz, which is available on the Optech Gemini system. The ALTM 3100 has PRF options of 33 kHz, 55 kHz, 70 kHz and 100 kHz.

Maintaining a high pulse repetition has the obvious advantage of increasing point density, however it comes at the cost of low pulse energy. A high PRF reduces the time the laser diode can recharge since pulses are being generated in rapid succession. With less time to recharge each individual pulse will contain less energy. The power loss has its greatest consequence upon pulse reception. The low energy returns create a lower probability of attaining multiple returns from individual pulses and increase the likelihood that pulses will not return energy at all. With all other settings being equal, Figure 3.1 displays how changing the PRF affects point spacing and pulse energy. Notice that the pulses follow similar paths, however when a low PRF is implemented there is significantly more spacing between the points; however, the dots are much larger representing the increase in transmitted energy.

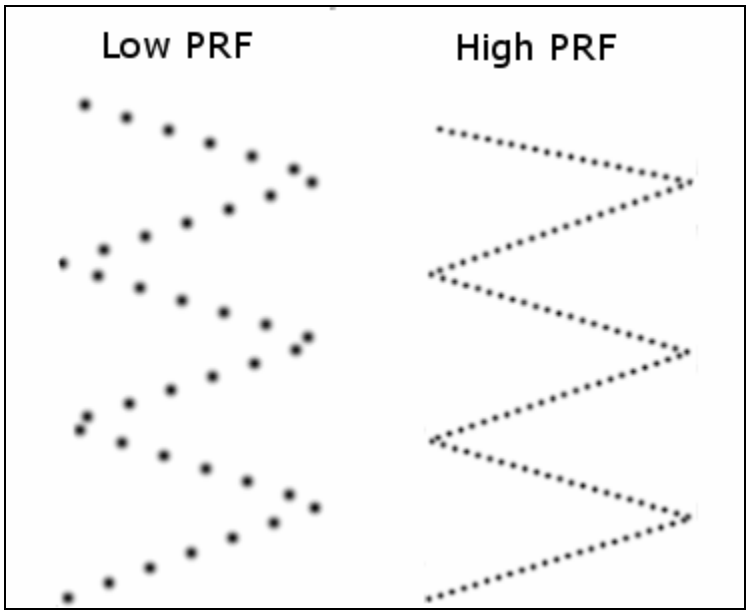


Figure 3.1
Variation of Pulse Repetition Frequency (PRF).

3.1.2 Scan Frequency

The scan frequency determines how quickly the scanning mirror will complete one scan cycle. A cycle is completed when the scanning mirror returns to the same angular location and is travelling in the same direction. If the scan frequency is set at its upper limit, each scan line will be closer together which will decrease the along track point spacing. However, the across track spacing of the points will be spread further apart as larger angular movement of the mirror occurs between pulses. Therefore, increasing the scan frequency has the overall effect of decreasing the along track spacing, but increasing across track spacing. In the ALTM 3100 the operational window for the scan frequency is from 0 to 70 Hz. Setting the scan frequency to zero creates observations only at nadir as the mirror will be stationary. Figure 3.2 shows the effects of scan frequency on the pulses if all other settings remain equal. The swath in the left portion of the diagram shows a

low scan frequency and subsequently the pulses within the scan line are close together. The swath on the right of the diagram shows a high scan frequency which has the effect of scan lines which are closer together, but contain fewer pulses in each individual line.

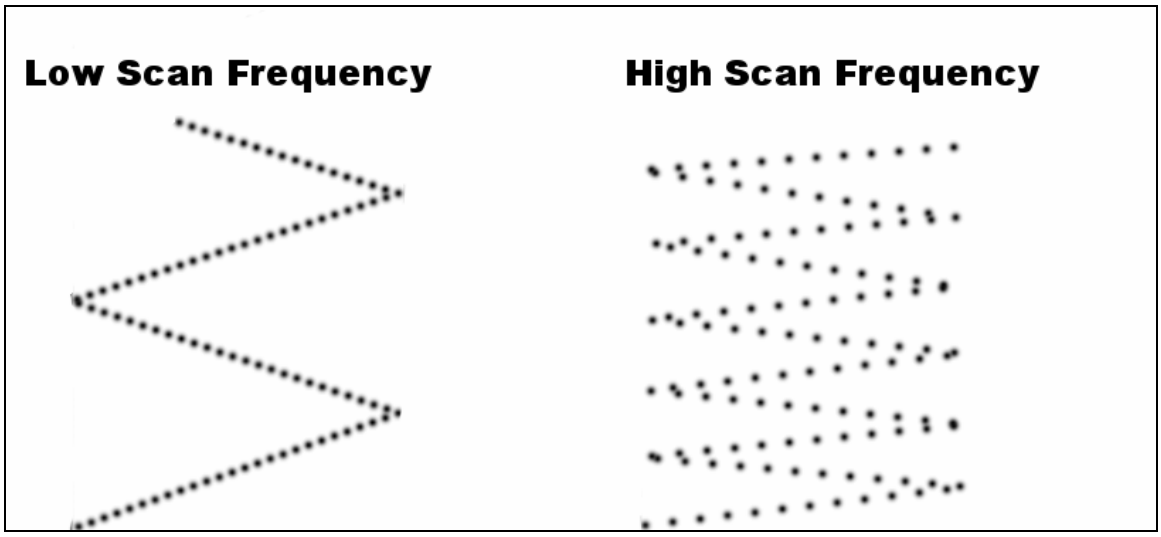


Figure 3.2
Variation of the scan frequency.

3.1.3 Scan Angle

The scan angle allows for control of the area of ground that is being observed in a single flight line. The value defines the limit of angular rotation of the scanning mirror from nadir. The swath width is generally twice the reported scan angle as it is applies to each of the port and starboard directions. The ALTM 3100 has an operating scan angle from 0 to 25 degrees. Often this upper limit of the scan angle is not utilised as it is well known that errors will increase as the scan angle increases [Schaer et al., 2007], [Goulden and Hopkinson, 2006]. Errors increase due to increased spreading of the pulse energy, increased ranges, propagation through the atmosphere, and increased ambiguity in return

signals due to 'smearing' of the pulse across the terrain. The smearing effect, which results in an increase in the pulse footprint size is caused by the incidence angle between the laser pulse vector and the terrain, and is described in [Schaer et al. ,2007]. This effect is shown in the Figure 3.3.

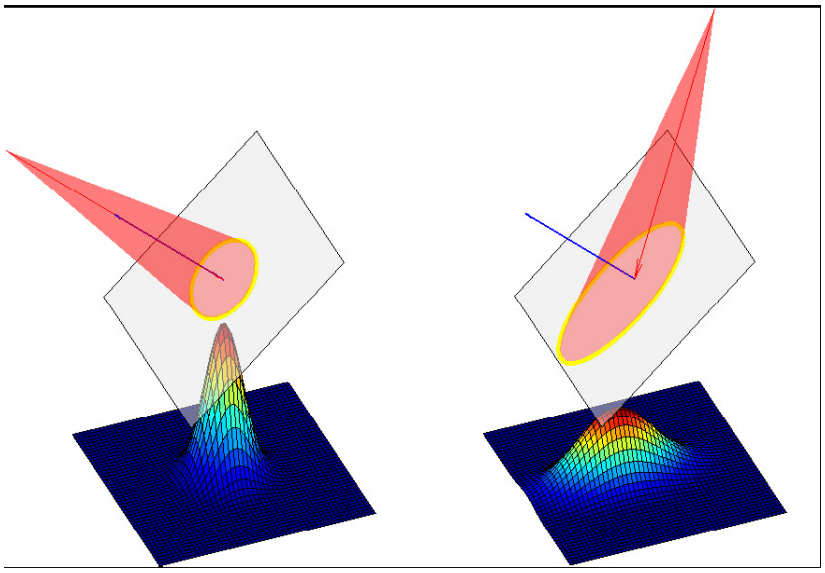


Figure 3.3
Smearing effect of LiDAR pulse from Schaer et al. [2007].

In addition to the physical errors of the laser pulse, significant error is also introduced due to the mechanics of the scanner system at maximum scan angles. As the scanner reaches its maximum extent it is decelerating to a stop point, and then accelerates in the opposite direction. The deceleration of the mirror is not constant across the entire swath width and is only present for a short period before the mirror reaches its stop position. During this intense period of deceleration significant error is introduced and several pulses are bunched together during the brief stop. To compensate the scanning mirror scans further than the specified scan angle. The observations that are recorded outside of the scan angle range at the swath edges are not used in processing.

Point spacing is directly affected by scan angle as small scan angles will result in dense points within a scan line. The compromise with a small scan angle is that the amount of ground coverage is reduced. Therefore, reduced scan angles result in longer flight times and an increase in the cost of the survey. Figure 3.4 displays narrow and wide scan angles and how this affects the ground point spacing. The left swath in the Figure shows a narrow scan angle which creates a dense point spacing in the scan line and covers a thin area. The swath in the right portion of the diagram displays a wide scan angle which increases point spacing in the scan line and covers a larger area.

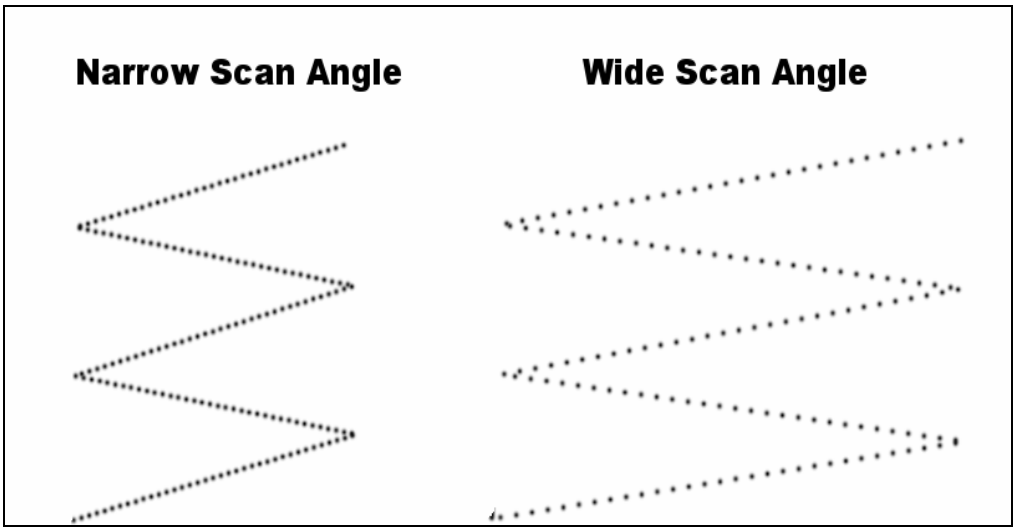


Figure 3.4
Variation of the Scan Angle.

3.1.4 Beam Divergence

The beam divergence is defined as the angular area that contains 68% of the laser pulse energy. There are two settings within the ALTM 3100 hardware for beam

divergence, 0.3 mrads and 0.8 mrads. This setting will affect the size of the laser footprint and the spread of energy throughout it. Narrow beam divergence settings contain a higher concentration of energy within a smaller pulse footprint which is beneficial for penetrating through canopy or for high altitude surveys. The wide angle setting spreads the pulse energy over a larger area which increases the likelihood it will contact desired objects such as power lines or find gaps within tree canopy. Beam divergence does not affect point spacing in a predictable way, however if the energy level is not sufficient to provide a return then significant dropouts will occur and the point density will suffer.

3.1.5 Aircraft Speed

Although aircraft speed can be considered in mission planning it is often determined by the weather conditions on the day of the flight. All aircraft will have a safe operating range of speeds that surveys must be done within. In general, a slower aircraft speed will allow for a denser point spacing in the along track direction since individual scan lines will be condensed. However, it will take longer to cover a survey area, thereby increasing the cost. Figure 3.5 shows the effects of aircraft speed on point spacing. The left swath in the diagram displays a aircraft travelling at a slow speed resulting in scan lines that are condensed. The swath on the right side of the diagram displays a greater spacing between scan lines due to increased speed.

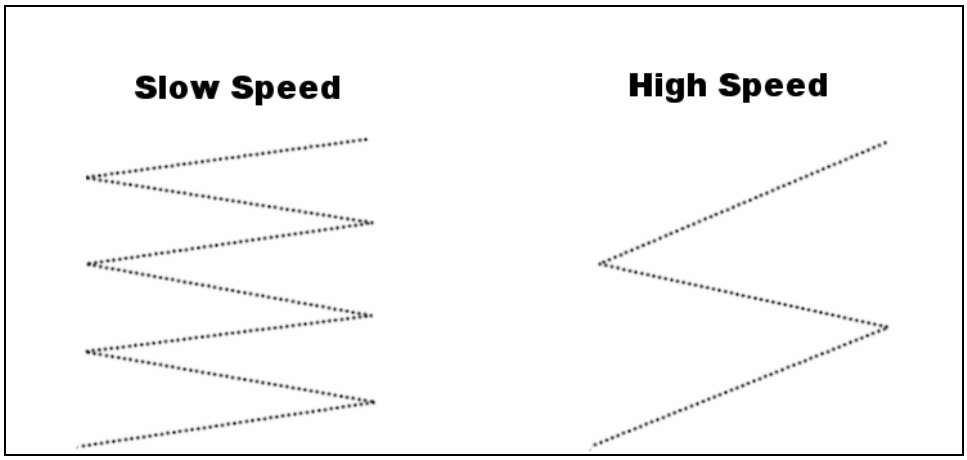


Figure 3.5
The effect of aircraft speed on point spacing.

3.1.6 Overlap

For surveys that intend to cover large contiguous areas overlap of flight strips is recommended to avoid gaps or holes in the data set. If strips were designed to be flown with no overlap then turbulent conditions causing angular motion of the aircraft could cause a section of the terrain to be missed. Usually flight strips are designed so that there is at least 50% of overlap. This is implemented by planning the edge of each swath to be the flight path of the following strip. With this design every piece of the survey area is observed twice. Also, areas covered by swath edges will subsequently be observed by nadir observations. This has two advantages, data gaps will usually appear at swatch edges where aircraft motion most severely affects observations, and low accuracy observations at swath edges will be overlapped with higher accuracy observations from nadir. The amount of overlap selected will directly affect the point density and the speed of the survey. High overlap will increase the point density and the time of the survey.

3.1.7 Flying Altitude

Like aircraft speed the flying altitude can be designed during mission planning but often is decided on the day of the survey based on weather conditions. Laser pulses are not able to penetrate through cloud cover or thick fog. Therefore, surveys must be flown underneath the cloud ceiling. In addition, flying heights must be planned above eye-safe limits based on the type of laser in the system. The laser in the ALTM 3100 is classified by the FDA as a class IV laser, indicating it could be a hazard to both the skin and eyes [FDA, 2006]. As the flying height increases the amount of energy propagating to the ground will decrease. There are minimum altitude limits the aircraft must maintain in order not to injure individuals on the ground. The altitude affects point spacing similarly to modifying the scan angle as lower altitudes will result in higher point densities, but will also increase survey time. Figure 3.6, from Optech [2004] shows several swath strip boundaries represented by coloured polygons that are a result of low altitude and high altitude surveys. Notice that fewer strips were required with the high altitude survey, however the low altitude strips would contain a higher point density.

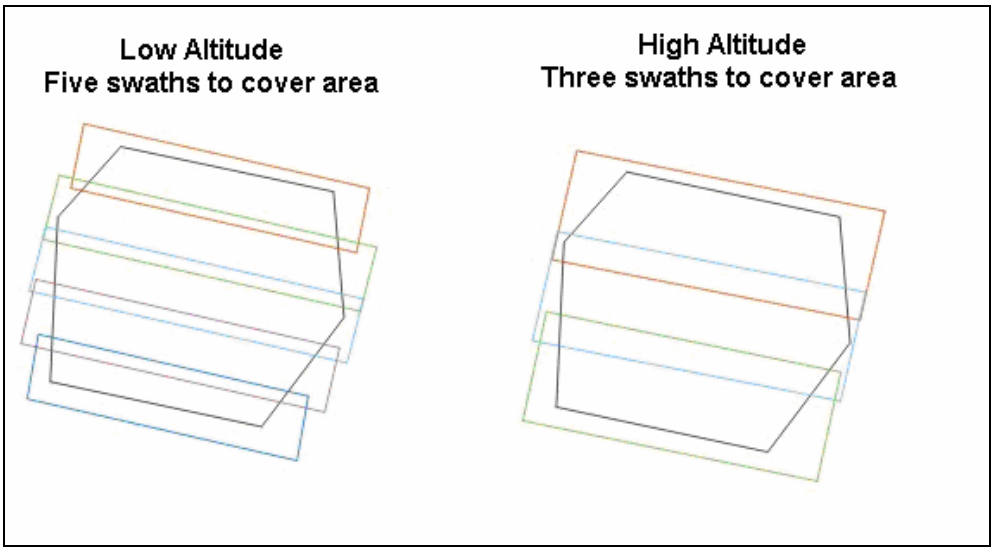


Figure 3.6 [Optech, 2004]
Effect of Altitude on Swath Width.

3.1.8 Combined Effect of Mission Parameters

Table 3.1 summarizes the each of the above mentioned properties and whether it affects point spacing, the energy of the returned pulse, and the time of completing the survey.

Table 3.1
Summary of each LiDAR observation parameter used in mission planning and whether this affects the point spacing, pulse energy and survey time.

Survey	Point Spacing	Pulse Energy	Survey Time
PRF	●	●	
Scan Frequency	●		
Scan Angle	●		●
Beam Divergence		●	
Aircraft Speed	●		●

Aircraft Altitude	•	•	•
Overlap	•		•

Figure 3.7 displays how actual observations will distribute with typical survey settings. The PRF and scan frequency were varied to display the effect of these parameters on the point spacing. Although overlap affects point spacing it is not included as data was taken from single flight strips. The grid spacing in figure 3.7 is one metre in both the *x* and *y* directions. Therefore, point density in terms of square metres can easily be determined by the number of grey points within each square. Notice the example in the top portion of the figure contains a closer along track spacing due to the increased scan frequency, however the across track spacing is larger. The example in the bottom portion has clearly defined scan lines due to the lower scan frequency which also caused a wider spacing in the along track direction.

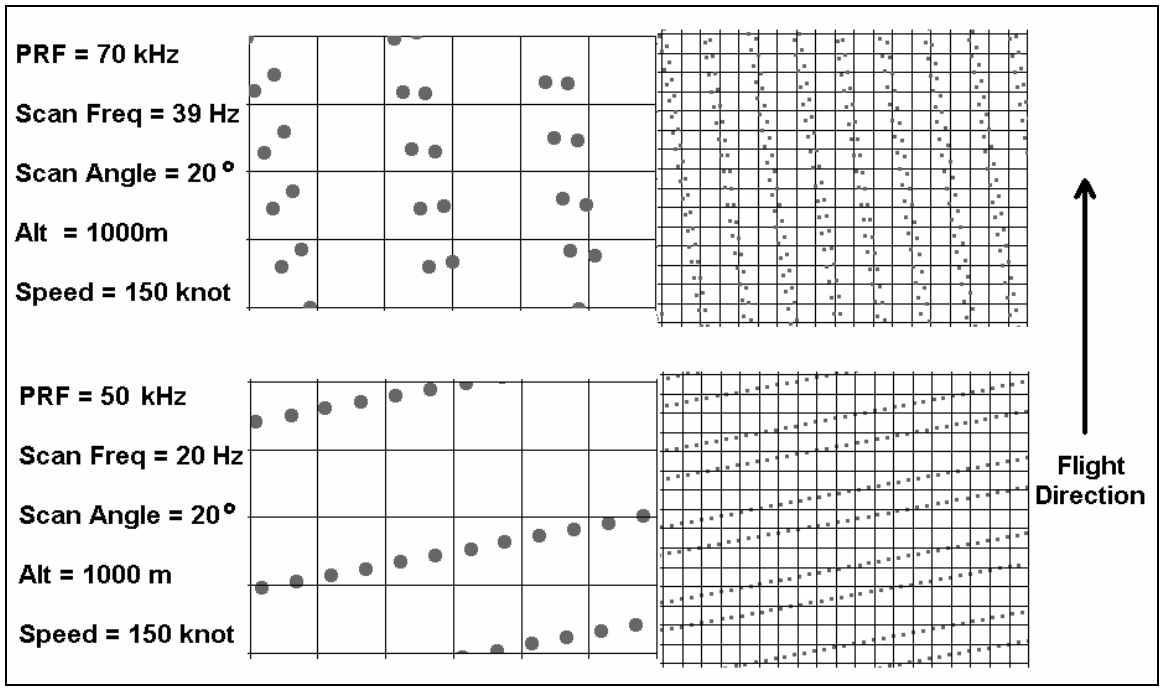


Figure 3.7

Effect of LiDAR observation parameters on point spacing.

Figure 3.7 also displays the pseudo-random point spacing that occurs as the result of the chosen specifications and flight conditions. Generally, DSMs are processed into regular grids using interpolation routines. Common routines include inverse weighted distance, spline and kriging. These routines result in additional error being introduced into the final results. For the purposes of slope error analysis the irregularly spaced data is being analyzed directly to avoid unwanted effects of the interpolation routines.

3.2 Effect of Point Spacing on Slope Calculations

Natural Resources Canada provides DSMs covering most of the nation. The DSMs are distributed through GeoBase, a partnership between federal, provincial and territorial governments to distribute current geospatial information describing the country. The point spacing in the DSMs vary from a minimum of 0.75 arc seconds to a maximum of 12 arc seconds depending on the area of the country. This translates into a minimum point spacing of approximately 23.2 metres and a maximum point spacing of 371 metres in the north-south direction. These DSMs are considerably coarse compared to typical LiDAR DSMs. The ability to acquire dense DSMs (on the order of several points per square metre) calls for a change in the outlook of DSM error. Consider Figure 3.8, in the upper portion of the diagram, a typical publicly available DSM with wide point spacing is shown. When the points are connected they form a coarse estimate of the terrain. When analyzing the DSM the error in the points is not a great concern as the coarse resolution

causes interpolation errors which are more significant, especially when the estimation bridges valleys or cuts peaks.

When DSM properties such as slope are determined the error in the points would have little effect on the final determination. The estimates of slope would be sufficiently coarse that the error in the points could be considered negligible. Inspecting the LiDAR DSM in the bottom portion of Figure 3.8 show the points have a dense spacing, and the error level in the points begins to approach the level of spacing. With this dense realization of the terrain it is natural to assume that DSM properties are more accurate because of the reduced point spacing. This is a dangerous assumption as the errors in the points are going to propagate much more severely into derived quantities. Although the interpolation error has been drastically reduced, the previously negligible error in the points begins to have a significant affect on the derived quantities. Therefore, there is an inherent difference in how errors must be viewed. In coarse DSMs the significant errors can be viewed as a result of the interpolation, and not as a result of the errors in the observations. As DSMs become denser the errors in the observations become more significant. Therefore, when utilizing coarse DSMs there is a tendency to ignore a rigorous error analysis of the points as it is negligible compared to the interpolation routines. This same attitude cannot be taken with dense DSMs as the errors in the points will contribute much more significantly to the error in derived quantities.

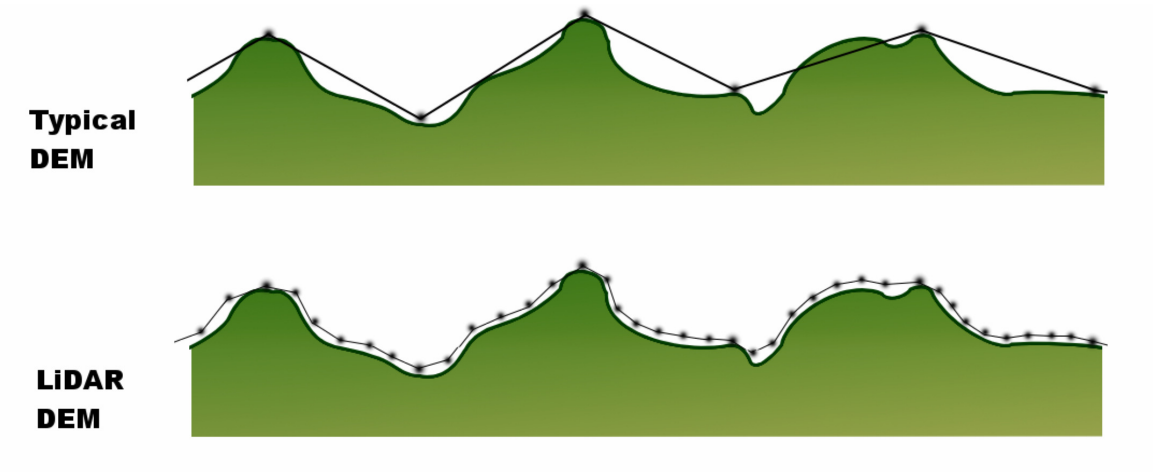


Figure 3.8
Difference in DSMs derived from LiDAR and typical publicly available DSMs.

A similar effect has been noted in studies of the derived quantities of high resolution DSMs. [Oksanen and Sarjakoski ,2005] and [Oksanen and Sarjakoski ,2006] described errors in slope based on the value of slope and aspect of the terrain. These were described with a DSM that had a point spacing of either 5 or 10 metres. The latter study considered errors in a DSM derived from contours while using LiDAR data as a reference. It was found that the error in slope and aspect both increased with the terrain steepness. Since the point density was consistent throughout the study, its effect on the error in slope was not considered. Although these studies performed error analysis on dense DSMs, they have not included a rigorous analysis of the error in individual points and how this propagates into derived quantities of dense DSMs. Performing simple error propagation on the formula for slope using the Special Law of Propagation of Variances (SLOPOV), we can develop the error in slope as follows:

$$slope = \left(\frac{z_1 - z_2}{x_1 - x_2} \right) \qquad (3.1)$$

followed by,

$$\sigma^2_{Slope} = \left[\frac{\partial slope}{\partial x_1}\right]^2 \sigma^2_{x_1} + \left[\frac{\partial slope}{\partial x_2}\right]^2 \sigma^2_{x_2} + \left[\frac{\partial slope}{\partial z_1}\right]^2 \sigma^2_{y_1} + \left[\frac{\partial slope}{\partial z_2}\right]^2 \sigma^2_{y_2} \qquad (3.2)$$

where x and z represent the horizontal and vertical coordinates of a point respectively, and σ represents the respective error in a quantity. Analyzing this formula with conservative error estimates of a decimetre for the position of each coordinate value and varying the point spacing and terrain steepness results in Figure 3.9. Blue, red and green values represent errors when terrain steepness in 0, 45 and 70 degrees respectively. The grey portion represents a point spacing that is typically achieved with LiDAR DSMs.

It can be seen that the error in slope reduces drastically with an increase in point spacing. This is an intuitive conclusion as the slope calculation contains the point spacing in the denominator of the expression. If this value is small, a minute change will greatly affect the outcome of the calculation. On the contrary, a large value in the denominator with a slight change will have very little effect on the final result.

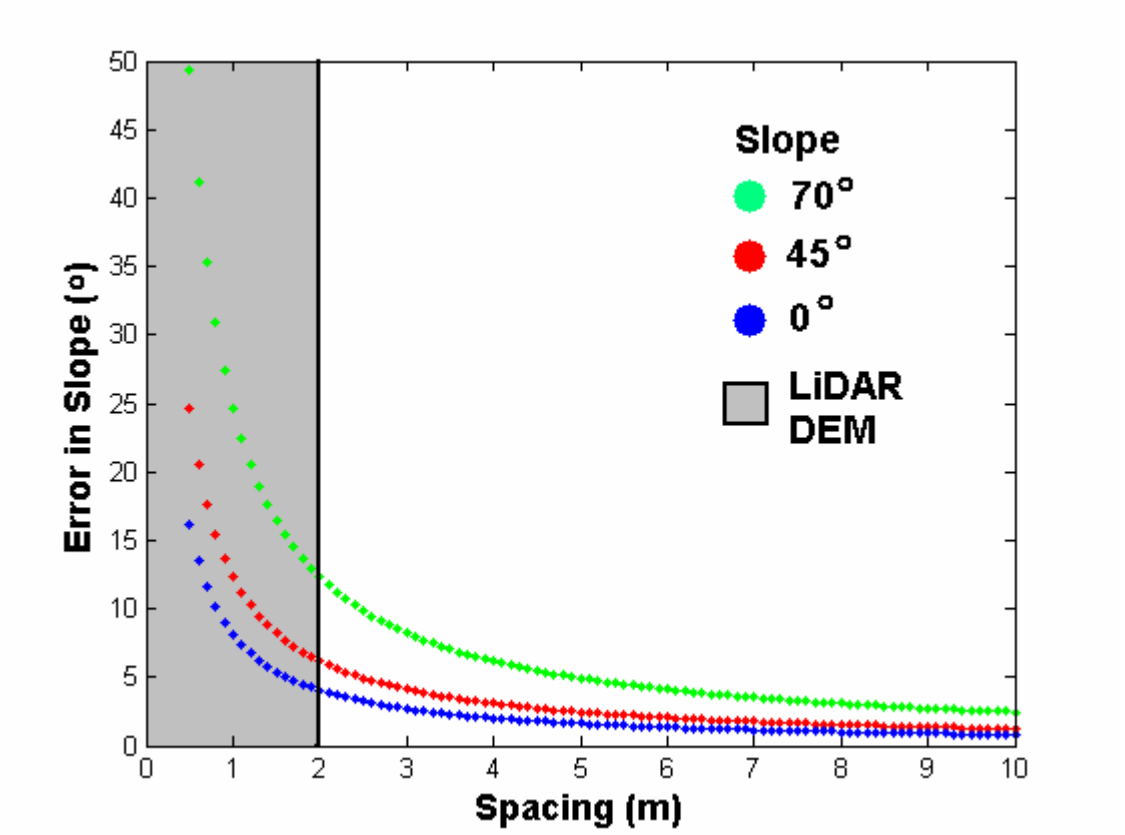


Figure 3.9
Error in slope determination as horizontal point spacing and terrain steepness is varied.

Consider the grey portion of Figure 3.9 where results from a typical LiDAR DSM point spacing are displayed. It is in this area of the graph that values begin to tend toward infinity and can generate extremely poor results. With typical DSMs, such as those publicly available by Natural Resources Canada, the spacing would fall past the 10 metre mark on the graph where it begins to reach a limit along a horizontal asymptote. The error magnitude of the slope calculations at this spacing has less importance as it has become constant. This indicates that when dealing with LiDAR DSMs special caution must be taken when reporting or utilizing any derived parameters. Although they are dense and describe the physical terrain more accurately than coarse DSMs they are much

more sensitive to the DSM characteristics. Now consider Figure 3.10, in which the vertical error was modified along with point spacing with slope left constant at 25°. Blue, red and green values represent errors in slope when vertical error is 0.1, 0.3 and 0.5 metres respectively. The grey portion represents a point spacing that is typically achieved with LiDAR DSMs.

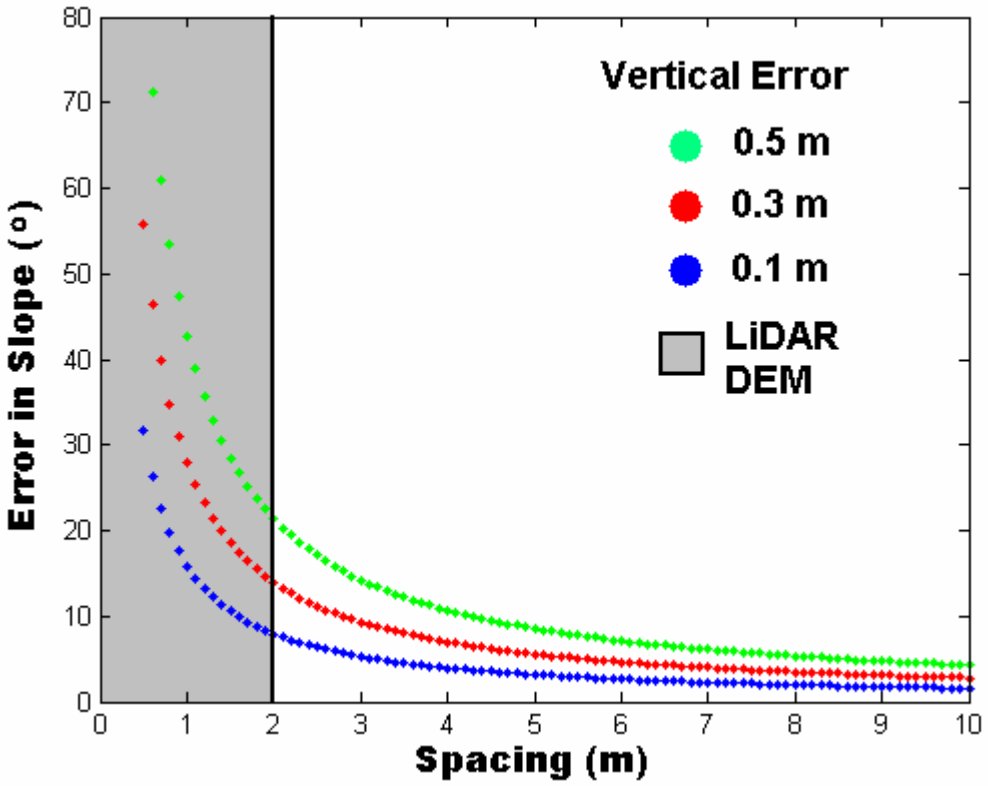


Figure 3.10
Error in slope determination as point spacing and vertical error is varied.

From the diagram it is evident that changes in vertical error also significantly affect the derived slope quantity. The vertical error was initially set at 0.1 metres as this is a conservative estimate for the lower limit of vertical error in the LiDAR system. This was varied to 0.5 metres as a value attained in sloped terrain. From this, we can conclude

that an appropriate analysis of the error in the point observations within a DSM is an important consideration when reporting errors in derived quantities. When dealing with LiDAR DSMs this consideration has greater importance since the point spacing drastically affects the sensitivity of these quantities. Therefore, a rigorous error analysis of LiDAR point position errors has a critical role in predicted errors in a DSM, making the analysis of errors based on terrain slope, which are known to be significant, of utmost importance.

Chapter 4

Development of Terrain Slope Error Model

This chapter discusses the derivation of the slope error model for LiDAR DSMs. Details of the required theory and algorithmic processes are given. For reference, a flow chart has been included after the explanation of the derivation and is labelled as Figure 4.10. It can be referenced throughout the Chapter to gain perspective on the entire algorithmic process. The chapter begins by discussing surface generation using a TIN model, and the calculation of slopes from this model. From this, the implementation of the hardware error model discussed in Chapter Two is presented. The use of the hardware errors to calculate a slope-based error estimate is formulated. Following this, testing on a simple sample data set is performed. To complete the chapter the algorithm is tested on a LiDAR dataset of a ski hill in Windsor, Nova Scotia. From this, it is proven that the algorithm performs as expected and is capable of handling large LiDAR DSMs.

4.1 Delaunay Triangulation and Slope Calculation

The first step of the slope error model is creating a Delaunay triangulation to calculate connections between all points in the data set. A Delaunay triangulation creates a surface without interpolating or smoothing the original data points allowing their true position and integrity to be maintained. The surface is composed of several triangular faces connected to three data points. The Delaunay model is designed to ensure that no points will exist within a triangular face and that triangles are nearly equilateral. Maintaining the integrity of original data points is preferable to interpolation or

smoothing routines because it does not allow an artificial increase in the error of the data points. Under these circumstances the initial error in the data points can be correctly predicted and then errors can be propagated through an interpolation or smoothing routine to determine final point position errors. Triangulation is also considered the best model for calculation of true slopes [Maune, 2007] which is a critical step in the slope error model.

To implement the Delaunay triangulation the Matlab function *Delaunay* is used. The function accepts the *x* and *y* data coordinates pairs as input and creates the surface of triangular faces. The function returns a matrix in which each row contains three index references to the input point coordinates. These three referenced points create one triangle within the TIN model. A sample of input data containing four points and output data that resulted in two triangles is provided below and diagrammed in Figure 4.1:

$$Input = \begin{matrix} P1 \\ P2 \\ P3 \\ P4 \end{matrix}\begin{bmatrix} 1 & 0 \\ 2 & 2 \\ 3 & 2 \\ 2 & 0 \end{bmatrix} \qquad Output = \begin{matrix} Tri1 \\ Tri2 \end{matrix}\begin{bmatrix} 4 & 2 & 1 \\ 3 & 2 & 4 \end{bmatrix},$$

where P1, P2, P3, P4, Tri1 and Tri2 are as shown in Figure 4.1.

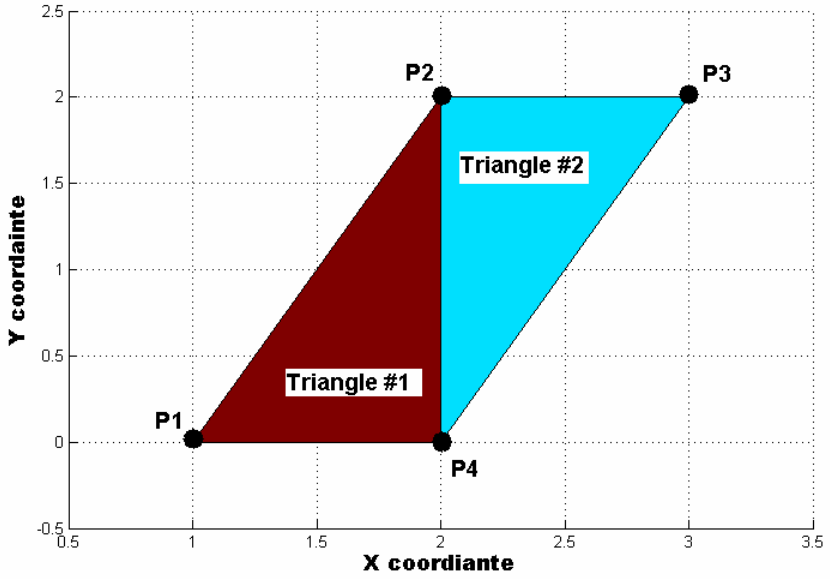


Figure 4.1
Representation of TIN surface using just four points generating two triangles

From Figure 4.1 the first triangle can be identified as being created from *P4*, *P2*, and *P1* as indicated by the first row in the output matrix. The same follows for *P3*, *P2* and *P4* which form the second triangle. This output matrix details all of the triangular edge connections in the TIN model. Knowledge of these connections allows the slope to be calculated between any two adjacent points along a triangular edge. The calculation of slope between points is performed through equation (3.1). However, when dealing with three dimensional data the point spacing is calculated as the Euclidian distance between the horizontal x and y coordinates. Therefore equation (3.1) must be modified to:

$$slope = \left(\frac{z_1 - z_2}{\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}} \right), \tag{4.1}$$

where x and y are the horizontal coordinates and z is the vertical coordinate.

With the Delaunay Triangulation complete the calculation of the error due to slope begins by sequentially analyzing each point in the dataset for the connections to any other points. This is performed algorithmically by accessing each point in the coordinate data and querying the output from the Delaunay triangulation to isolate the triangles that include the point. Once all triangles the point contributed to are found a list of all other points that were used to construct the triangles is compiled. This list details all of the connections of the point to any other point in the data set. The slope error is calculated along each of these connections, and the maximum error of all connections is output. To properly quantify the error due to slope, the initial error prediction based on hardware components is determined along the plane parallel to the connection vector between points.

4.2 Calculation of 2D ellipse based on covariance information

Each point in the coordinate dataset is associated with a covariance matrix of errors that is obtained from the hardware error parameters discussed in Chapter Two. This covariance matrix contains errors in three dimensions, populated with variances in the *x*, *y* and *z* coordinate directions respectively along the main diagonal, and covariance information in the off-diagonal elements as follows:

$$Covariance_{XYZ} = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} & \sigma_{xz} \\ \sigma_{xy} & \sigma_y^2 & \sigma_{yz} \\ \sigma_{xz} & \sigma_{yz} & \sigma_z^2 \end{bmatrix} \tag{4.2}$$

Form the covariance information a three-dimensional error ellipsoid can be constructed that defines the volume in three-dimensional space a percentage of observations will fall

with a defined confidence. A simulation of point data and associated error ellipsoids can be visualised in Figure 4.2. Within the figure a TIN surface has been generated and forms the terrain. Each point has an associated error ellipsoid. From this, the problem of the slope error can begin to be understood as the actual terrain could take on many forms depending on where the true terrain points are located within the volume of the error ellipsoid.

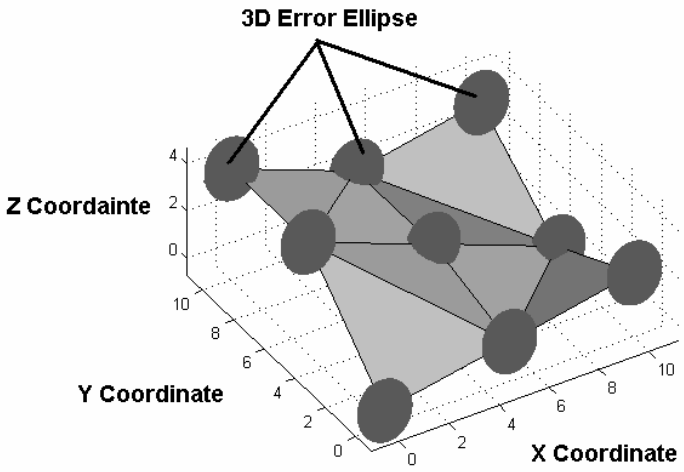


Figure 4.2
Representation of TIN surface with 3D error ellipsoids located on each point.

To analyze the error along a triangular edge connection between two points, a two-dimensional error ellipse in the same plane as the triangular edge, and perpendicular to the *x-y* plane is required. This two-dimensional error ellipse is extracted from the three-dimensional ellipsoid using the direction vector of the connection and the orientation of the error ellipsoid. The resulting two dimensional error ellipses are also represented by a covariance matrix of errors. Figure 4.3 displays the extraction of the 2D ellipse. From the diagram it is evident that the 2D error ellipse is contained within the

three-dimensional error ellipsoid. Knowing the direction of the triangular edge and the orientation of the semi-major and semi-minor axes of the error ellipsoid the two-dimensional error ellipse can be determined.

The direction of the triangular edge can be calculated from the horizontal coordinates between the two points as follows:

$$\theta = \text{atan}\left[\frac{y_2 - y_1}{x_2 - x_1}\right] \tag{4.3}$$

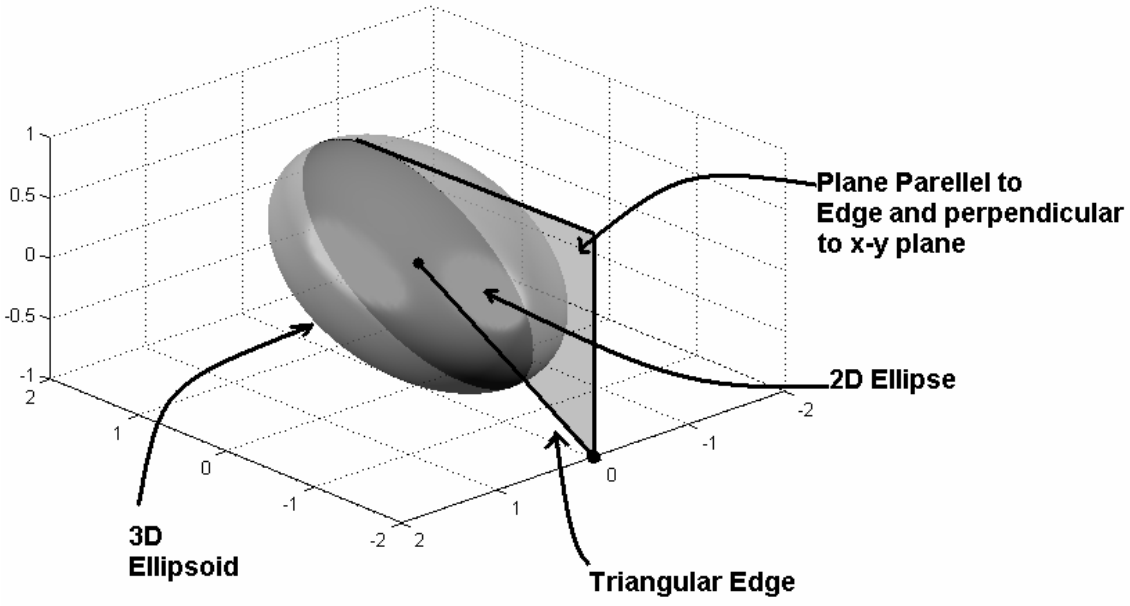


Figure 4.3
Two Dimensional extraction of an error ellipse from three dimensional error ellipsoid. The extracted error ellipse exists in a plane that is parallel to the triangular edge and perpendicular to the *x-y* plane between two points in the TIN surface.

where x and y are the horizontal coordinates of the two points along the triangular edge and θ is the direction measured counter-clockwise from the x axes that the triangular edge follows. If the original axis are rotated by this angle around the z axis the resulting axes

(*x',y'z'*) will contain the required variance along the new *x'* axis as it will fall along the direction of triangular edge. The variance along the *z'* axis will represent the vertical error and the rotated covariance values will be in the appropriate orientation. Following this, the required 2D ellipse can be determined by removing the variances and co-variances from the y-axis. The process is formulated as follows:

$$[Cov'] = [rot]_z [Cov][rot]_z^T, \tag{4.4}$$

where square brackets indicate a matrix, rot_z represents a rotation matrix around the z axis, *Cov* is the original covariance matrix and *Cov'* is the rotated covariance matrix. The details of the equation can be rewritten as:

$$\begin{bmatrix} \sigma_{x'}^2 & \sigma_{x'y'} & \sigma_{x'z'} \\ \sigma_{x'y'} & \sigma_{y'}^2 & \sigma_{y'z'} \\ \sigma_{x'z'} & \sigma_{y'z'} & \sigma_{z'}^2 \end{bmatrix} = \begin{bmatrix} \cos(\theta) & \sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \sigma_x^2 & \sigma_{xy} & \sigma_{xz} \\ \sigma_{xy} & \sigma_y^2 & \sigma_{yz} \\ \sigma_{xz} & \sigma_{yz} & \sigma_z^2 \end{bmatrix} \begin{bmatrix} \cos(\theta) & \sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}^T \tag{4.5}$$

The final error quantities we are interested in come from only the *x-z* plane, therefore the final two dimensional error ellipse will take on the form:

$$Cov_{xz} = \begin{bmatrix} \sigma_{x'}^2 & \sigma_{x'z'} \\ \sigma_{x'z'} & \sigma_{z'}^2 \end{bmatrix}. \tag{4.6}$$

Once the new 2D error ellipse has been calculated in the appropriate orientation for both points the situation can be analyzed in two dimensions where the resulting two points and their associated error ellipses will appear as in Figure 4.3:

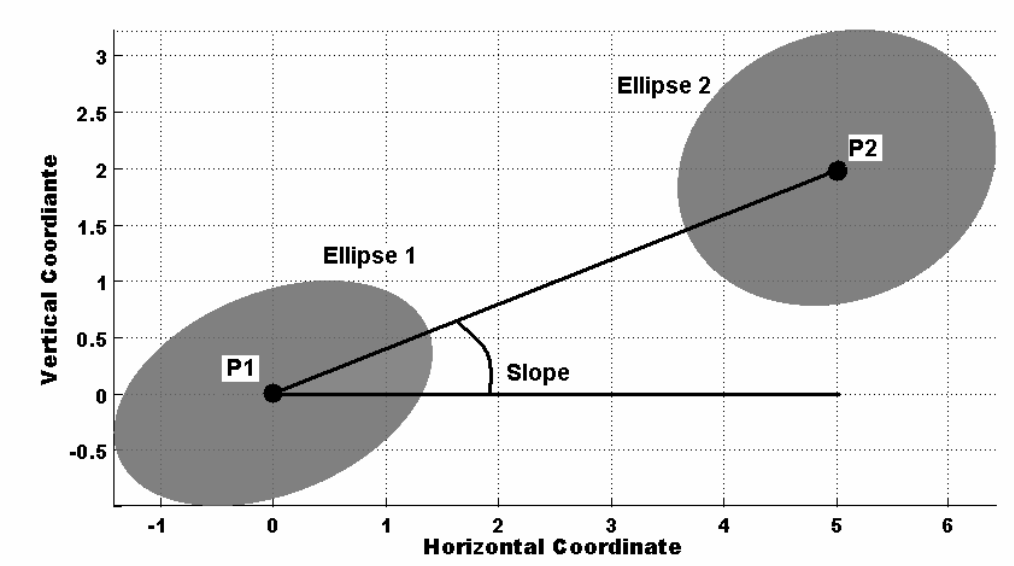


Figure 4.4
Representation of two points and their associated error ellipses shown in the two-dimensional plane of the connection along a triangular edge.

4.3 Calculation of Error Based on 'Worst Case' Probability Surface

Analyzing Figure 4.3 it becomes evident that the terrain could take on many realizations, but are limited by the extents of the error ellipses that statistically predict the area the point could fall within. Figure 4.4 displays several possibilities the true terrain could take from the point and its associated error ellipse.

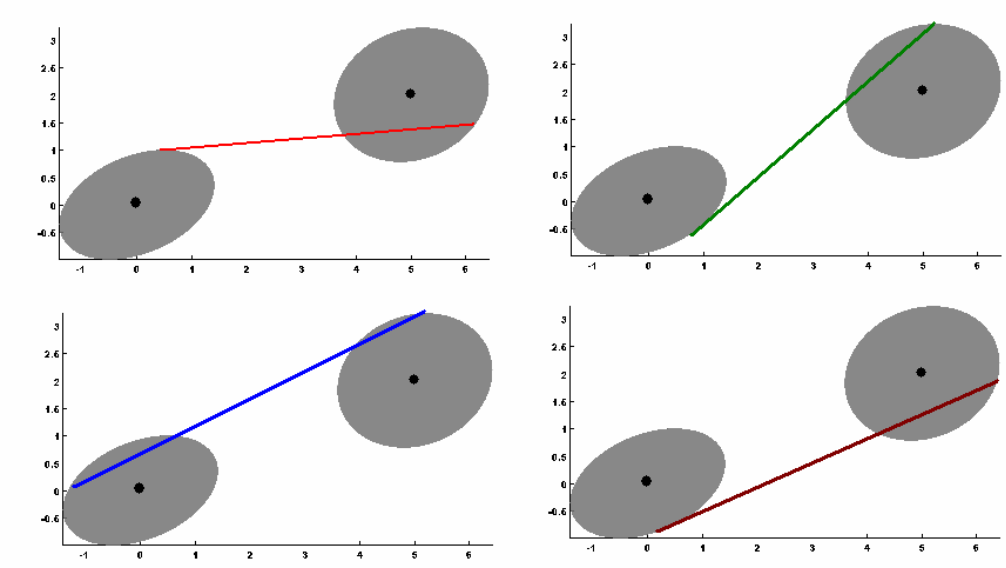


Figure 4.5
Multiple Realizations of the terrain based on the error ellipses at each point. Theoretically there are an infinite number of realizations.

Although there is obviously an infinite number of connections that can be made within the error ellipses there is one realization that will result in the largest amount of error between the acquired point and the possible realizations of the terrain. This will give an estimation of the worst-case scenario, or maximum amount of vertical error that could exist between the actual point and physical terrain. To identify the situation in which the maximum amount of possible error is achievable, examine Figure 4.5. From the figure, a conclusion can be drawn that the realization of the terrain that does not intersect the ellipse will result in the largest possible vertical error. The black lines represent the possible realizations of the terrain, while the red line represents the vertical error. The red line is the largest when the realization of the terrain is tangent to the edge of the ellipse and not secant.

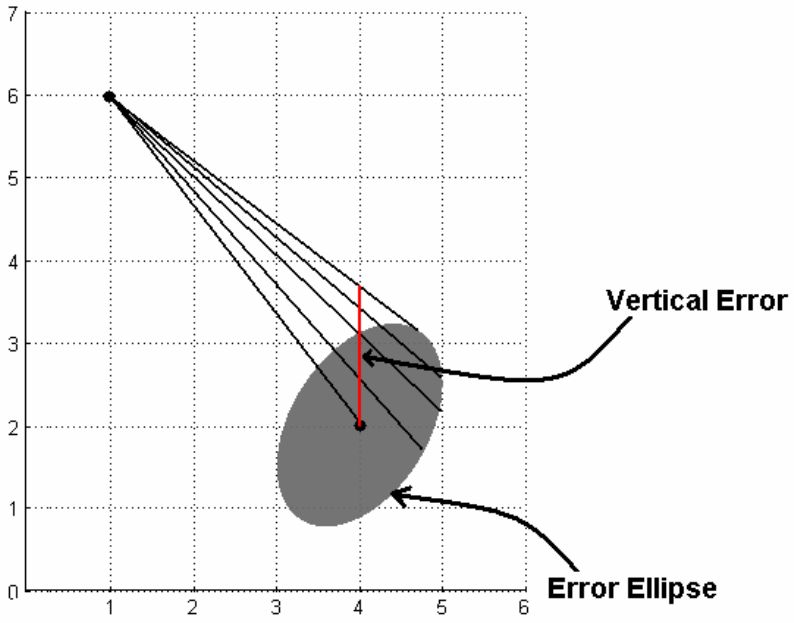


Figure 4.6
The maximum difference between the centre point of the ellipse and a line from another point occurs when the line is tangent to the ellipse.

Intuitively, it can be visualised that there are two possible solutions which are tangent to the ellipse. This situation can be seen in Figure 4.6, which clearly displays the phenomenon with the original point (*p*) in space being labelled as Point 1, and the two points on the ellipse that form a tangent line with Point 1 being labelled in maroon as Tangent 1 (*t1*) and Tangent 2 (*t2*).

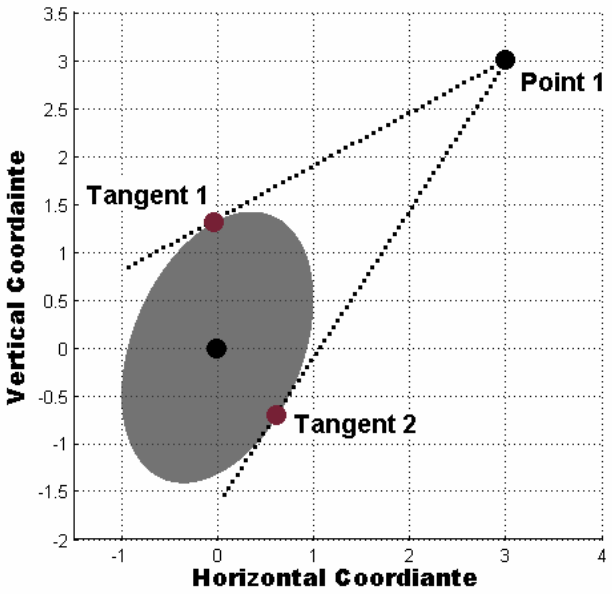


Figure 4.7
Developing a tangent line from Point1 to the surface of the ellipse always yields two solutions represented by Tangent 1 and Tangent 2.

To determine the line from a given point in space that is tangent to an ellipse, the parametric equations of a line and ellipse must be investigated.

The general parametric equation of an ellipse can be written as:

$$\begin{aligned} x(t) &= c_x + a\cos(t)\cos(\alpha) - b\sin(t)\sin(\alpha) \\ y(t) &= c_y + a\cos(t)\sin(\alpha) + b\sin(t)\cos(\alpha) \end{aligned} \tag{4.7}$$

Where, c represents the coordinate of the centre of the ellipse, α is the counter-clockwise rotation angle of the semi-major axis of the ellipse from the x axis and t is the angular parameter defining the location on the edge of the ellipse. The parametric equations for a line can be written as follows

$$\begin{aligned} x(s) &= p_x + m_x s \\ y(s) &= p_y + m_y s \end{aligned} \tag{4.8}$$

where p is the point of origin of the line, m is the slope of the line and s is the parameter defining the location on the line. Since the line of interest must be tangent to the ellipse, the first derivative of the parametric equations of the ellipse results in the slope of the desired line. To simplify the derivation, it can be assumed that the ellipse is located at the origin and is not rotated. This can be easily implemented practically by shifting the centre of the ellipse (c) and p by c, and then rotating each by α. After the tangent point has been calculated, it can be shifted back into the proper location by rotating it by $-\theta$ and translating by $-c$. Having the centre of the ellipse at the origin with no rotation allows the allows the parametric equation of the ellipse to be simplified to

$$\begin{aligned} x(t) &= a\cos(t) \\ y(t) &= b\sin(t) \end{aligned}, \tag{4.9}$$

And the first derivative becomes

$$\begin{aligned} \frac{\partial x}{\partial t} &= -a\sin(t) \\ \frac{\partial y}{\partial t} &= b\cos(t) \end{aligned} \tag{4.10}$$

These values represent the slope of the tangent line we are interested in and can by substituted into the parametric equation of a line as follows

$$\begin{aligned} x(s) &= p_x - a\sin(t)s \\ y(s) &= p_y + b\cos(t)s \end{aligned}, \tag{4.11}$$

Since the point on the line and the point on the ellipse representing the tangent are equivalent the equations can be written as;

$$a\cos(t) = p_x - a\sin(t)s \tag{4.12}$$

$$b\sin(t) = p_y + b\cos(t)s \tag{4.13}$$

From the equation (4.12), *s* can be defined as

$$s = \frac{a\cos(t) - p_x}{-a\sin(t)} \tag{4.13}$$

which can be substituted back into equation (4.13) to achieve

$$b\sin(t) = p_y + \frac{b\cos(t)\left(a\cos(t) - p_x\right)}{-a\sin(t)} \tag{4.14}$$

This equation leaves only the parameter *t* to be solved which will define the location of the tangent line on the surface of the ellipse. Through some re-arrangement the equation is as follows and is fully derived in Appendix A;

$$a^2b^2 - p_y^2 - 2p_x ab^2\cos(t) + \left(p_x^2b^2 + p_y^2a^2\right)\cos^2(t) = 0. \tag{4.15}$$

If *cos(t)* is written as *x*, equation (4.15) can be thought of as a quadratic. Using the well-known quadratic formula the solution becomes;

$$\begin{aligned} t &= \operatorname{acos}\left(\frac{-b \pm \sqrt{b^2 - 4ac}}{2a}\right) \\ a &= p_x^2b^2 + p_y^2a^2 \\ b &= 2p_x ab^2 \\ c &= a^2b^2 - p_y^2a^2 \end{aligned} \tag{4.16}$$

and there are two solutions, as expected, to the calculation. To determine which tangent point to choose is not a straightforward problem. It depends on the location of the ellipse, its orientation and the location of the original point *p*.

To ensure the worst-case scenario solution is chosen the tangent point that is furthest from the origin point is selected. This is the point that will create the largest discrepancy between the probable terrain and the point located at the centre of the ellipse (*c*). Generally, the tangent points will be located on each side of *c* in the *x* direction. Since

we are analyzing only two points, *p* and *c,* respectively representing the connections along a single triangular edge in the TIN model, we cannot predict the connections of these points to other vertices in the TIN model. Without this knowledge, the vertical error prediction cannot be done on a tangent point that does not cross the horizontal location of point *c*. Consider Figure 4.8; the tangent point '*A*', is further from the origin point and will cross the horizontal coordinate at the centre of the ellipse (*c*). With this information we are able to measure the vertical error between the point and realization of the terrain, represented by the vertical dashed line extending from the centre point of the grey ellipse. Now consider point *B*, since the tangent line does not cross the horizontal coordinate there is no knowledge of where the probability surface will continue, as represented by several diverging lines in Figure 4.8. Therefore, the vertical error cannot be predicted from this information.

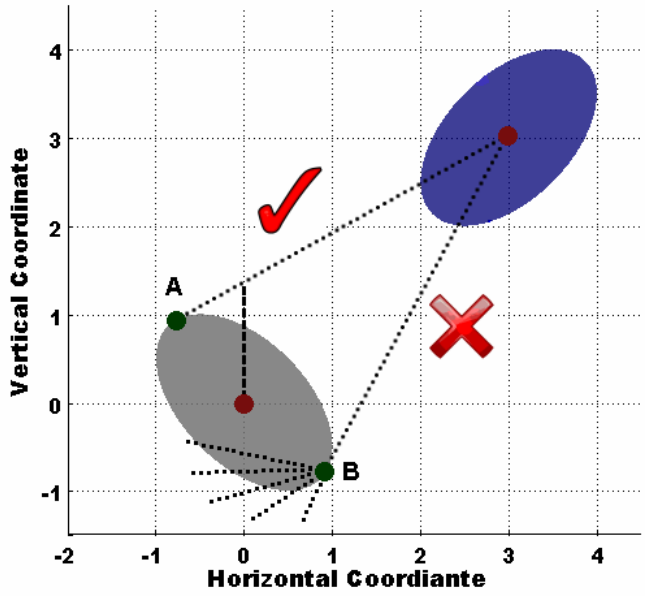


Figure 4.8
Selection of worst case scenario tangent line. The point 'A' is selected because it is further from the centre of the blue ellipse. Point 'B' does not cross the horizontal

coordinate at the centre of the grey ellipse therefore the continuation of the surface to a subsequent point cannot be predicted.

It is possible to check all the connected vertices in the TIN from Point *B* to determine where the probability surface will lead, however there are considerations associated with this implementation. Namely, the connections to other vertices will not exist in the same two-dimensional plane; therefore the projection of the connection onto the same plane would have to be found to properly perform the analysis. This is a complicated procedure as the projection would have to be extracted from the surface of the TIN model. Also, the other connections to this point are considered in the subsequent analyses of the remaining points in the TIN model. Those connections are also analyzed for worst-case scenario conditions, which will undoubtedly be more pessimistic then an analysis associated with the continuation of the terrain from *B*. Therefore, this is an unnecessary calculation as it will undoubtedly be disregarded in favour of more pessimistic solutions. From these conclusions Point *A*, represented as the furthest tangent point from *p,* is selected.

Once this tangent point is selected, a second tangent is calculated back to the error ellipse around *p*, the original point. To ensure that this also creates the worst-case scenario, if the first tangent point was located vertically above the point *c*, then the higher tangent vertically on the ellipse surrounding point *p* will be selected. If the first tangent point was located vertically below *c* then the lower vertical tangent point around *p* will be selected. This is shown in Figure 4.8 and is represented by the line associated with a large checkmark. If this procedure is not followed, and the tangent point associated with the ellipse around point *c* is vertically higher and connected with the vertically lower tangent

on the ellipse surrounding point then the probability surface will significantly underestimate the worst-case scenario of the vertical error. This situation is represented with a line associated with the large '**X**' in Figure 4.8.

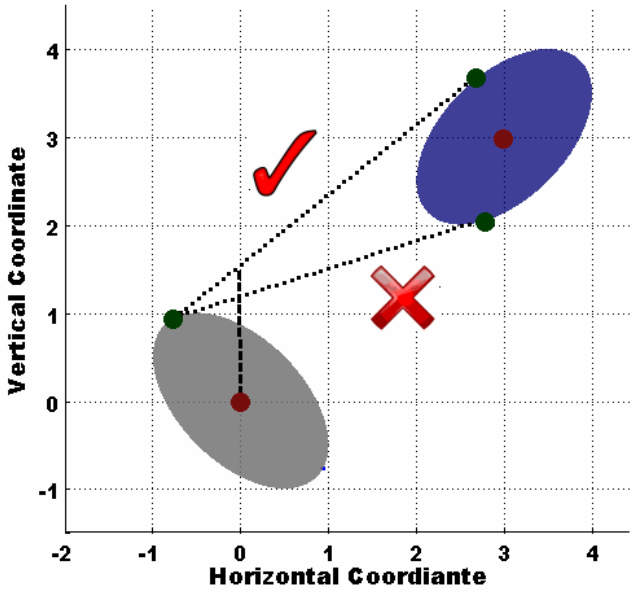


Figure 4.9
Selection of second tangent point. The vertical error represented by the black vertical dashed line is largest when connected to the point with a larger vertical coordinate.

Once the two correct tangent lines have been calculated the error can be found through the following sum:

$$\sigma_{Zslope} = \left(Tan1_z - c_z\right) + \left\|Tan1_{xy} - c_{xy}\right\|\left(Slope\right) \tag{4.17}$$

where:

$$\left\|Tan1_{xy} - c_{xy}\right\| = \sqrt{\left(Tan1_x - c_x\right)^2 + \left(Tan1_y - c_y\right)^2} \tag{4.18}$$

and σ_{Zslope} is the maximum possible error due to the slope, z represents a vertical coordinate, *Tan1* is the tangent point selected from the first tangent calculation point, x and y are horizontal coordinates, p is the origin point and will have the error applied to it, and *Slope* is the calculated slope between the two selected tangent points calculated from

(4.1). The slope error calculation is performed on all possible connections to each point and the maximum resulting error is saved for output. Once this has been completed, the algorithm accesses the next sequential point in the DSM and the process repeats.

4.4 Initial Tests on Simulated and Real Data

To easily assess the results of the algorithm a simple set of simulated test data is analyzed. The test data set consists of nine evenly spaced points forming a square. The spacing between the points is exactly five metres, and the maximum slope contained within the data set is 45°. The following matrix displays the point coordinates:

$$TestSet = \begin{matrix} x \\ y \\ z \end{matrix}\begin{bmatrix} 0 & 5 & 10 & 0 & 5 & 10 & 0 & 5 & 10 \\ 0 & 0 & 0 & 5 & 5 & 5 & 10 & 10 & 10 \\ 2.5 & 5 & 2.5 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Covariance matrices to represent each data point were chosen to be identical for further ease of analysis and contain typical LiDAR errors, namely 30 centimetres for each of the x and y coordinates and 10 centimetres for the z coordinate. Figure 4.10 displays the terrain in a TIN model with the error ellipses in their proper scale.

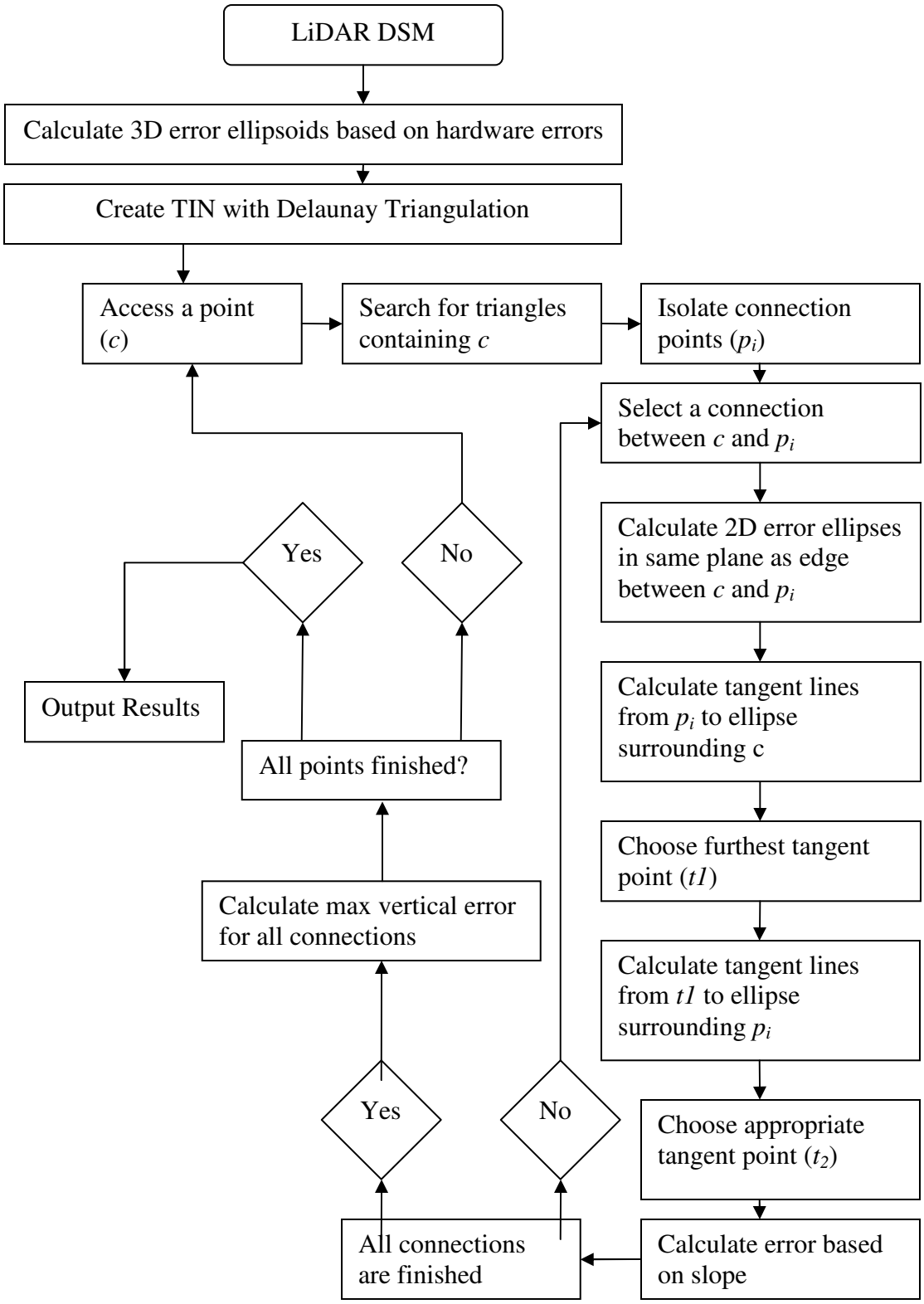


Figure 4.10
Flow-chart of the algorithmic process of slope based error calculation

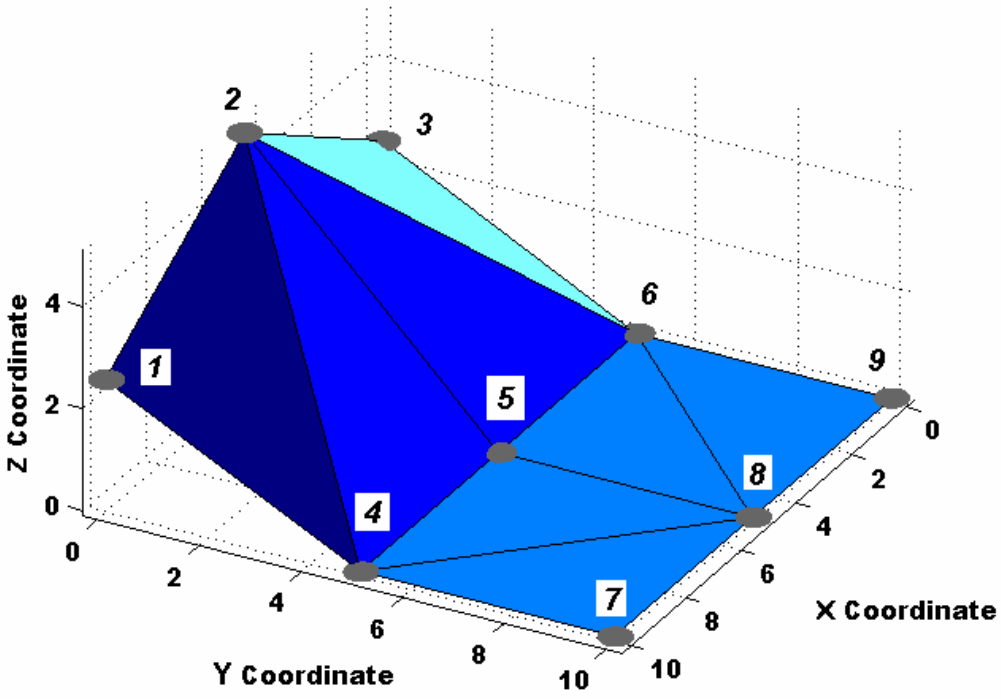


Figure 4.10
Diagram of the simulated test data with error ellipsoids designed following typically LiDAR errors drawn to scale

If the algorithm is applied to the sample terrain the output vertical errors due to slope for each of the points are displayed in Figure 4.11. Within the Figure the height of the bars represent the magnitude of the predicted vertical error and the horizontal red line represents the original vertical error prior to slope considerations.

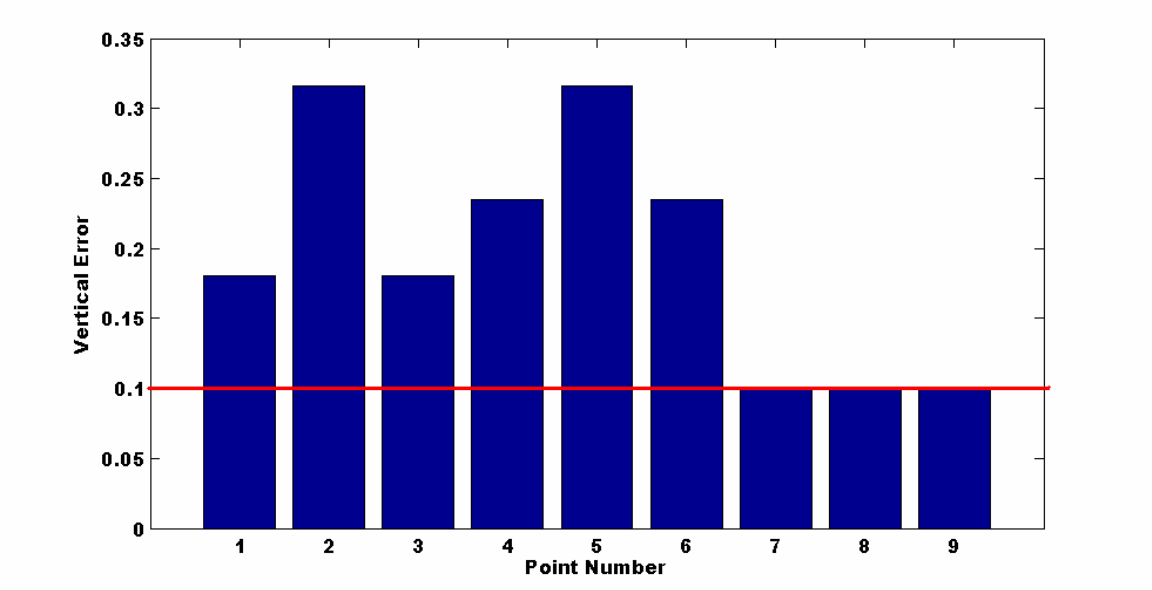


Figure 4.11
Value of the vertical error for each point in the simulated data. Each point number represents the numberes points in Figure 4.10.

Several conclusions can be drawn from the bar graph shown in Figure 4.11. The largest vertical error occurs on points two and five. By analyzing the TIN in Figure 4.9 it can be seen that points two and five are connected by the steepest slope as there is a vertical and horizontal change of five units resulting in a slope of 45°. The resulting vertical error is equivalent in both cases since the algorithm only outputs the maximum possible error which is equivalent when identical error ellipses and slopes are implemented. Also of note is that the error associated with points seven, eight and nine is equivalent to the initial error. This is what is expected since these points were located along flat ground where no additional error would be introduced as a result of slope.

The algorithm was applied to an actual data set of LiDAR observations of a ski hill in Windsor Nova Scotia. A ski hill is ideal test location as it contains steep slopes and has clear trails void of vegetative cover. Therefore, several points can be observed on the

true ground without the need to filter non-ground points. Details of the data set will be explained further in Chapter 5. The results of the slope error algorithm can be seen in the bottom portion of Figure 4.12 along with a slope map in the top portion of Figure 4.13.

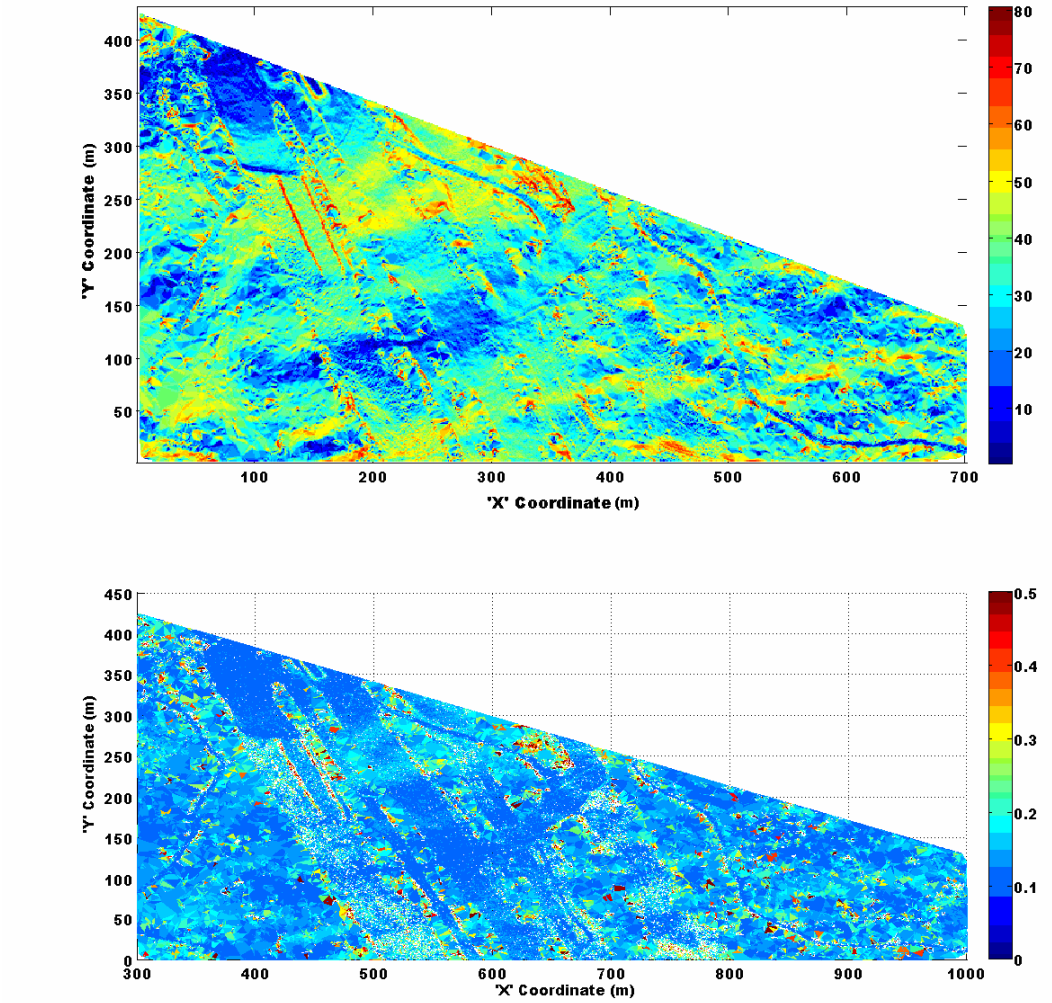


Figure 4.12
Vertical Error magnitudes of a LiDAR dataset over a ski hill. In general, errors range from original error estimations of approximately 10 centimetres up to 50 centimetres. Errors of up to 19 metres were seen but is not displayed so maximum color variation can be seen in the majority of data.

In the predicted error results, the range is approximately 0.1 metres to 0.5 metres. In actuality, errors ranged up to a maximum of 19 metres, however displaying the full range would compress the color variation in the majority of the data. In the slope solution the range is from 0° up to 80°. The high slope values are the result of ground points existing next to tree canopy. The correlation between highly sloped areas and increased vertical error can be easily seen by comparing the images. Theoretically, these diagrams show the algorithm is working as is expected and is successful for large LiDAR data sets. Chapter 5 will detail the comparison of the results with observations of higher accuracy, GPS specifically to determine whether the error estimates are appropriate.

Chapter 5

Testing Slope Error Prediction

In order to properly validate whether the designed slope error model is properly predicting observed errors on sloped ground a testing scheme must be implemented. This is achieved by comparing results from LiDAR observations against observations of higher accuracy. To obtain higher accuracy observations kinematic GPS observations were collected over various validation test sites. Kinematic GPS is chosen because of its high accuracy and quick collection rate. Two separate sites were sampled, one over a ski hill in Windsor, Nova Scotia, which most closely resembles glacial characteristics, and a second on a sloped road in Fredericton, New Brunswick. This Chapter will detail the characteristics of the datasets, the collection of the validation data and the resulting discrepancies between the two.

5.1 Ski Martock Dataset

The first test site chosen was Ski Martock, a ski hill in Windsor, Nova Scotia. A ski hill site is ideal for testing because it contains an array of varying slopes which are void of vegetation and are easily accessible. These areas of the terrain are similar to glacial environments and therefore conclusions drawn here can be extrapolated to the glacial environment. The ski hill was observed with the Optech ALTM 3100 LiDAR system on August 10th, 2006. The complete survey time was approximately 25 minutes and included four strips. Table 5.1 outlines the survey parameters used in the flight. The

compiled dataset used for testing used last return pulses only to better isolate ground positions.

Table 5.1
Parameters used in the Martock LiDAR survey

PRF (kHz)	Scan Frequency (Hz)	Scan Angle (degrees)	Flying Height (m a.g.l)	Flying Speed (knots)	Resolution (p/m^2)
50	31	15	1200	120	1

The entire Martock dataset consisted of over three million LiDAR point observations. The interest lies in the section of the survey that contains the ski trails. Clipping the portion of the survey to contain only the section with the ski trails reduced the dataset to just under 300 000 points. The slope error prediction routine was applied to the data set and the results are displayed in Figure 5.1.

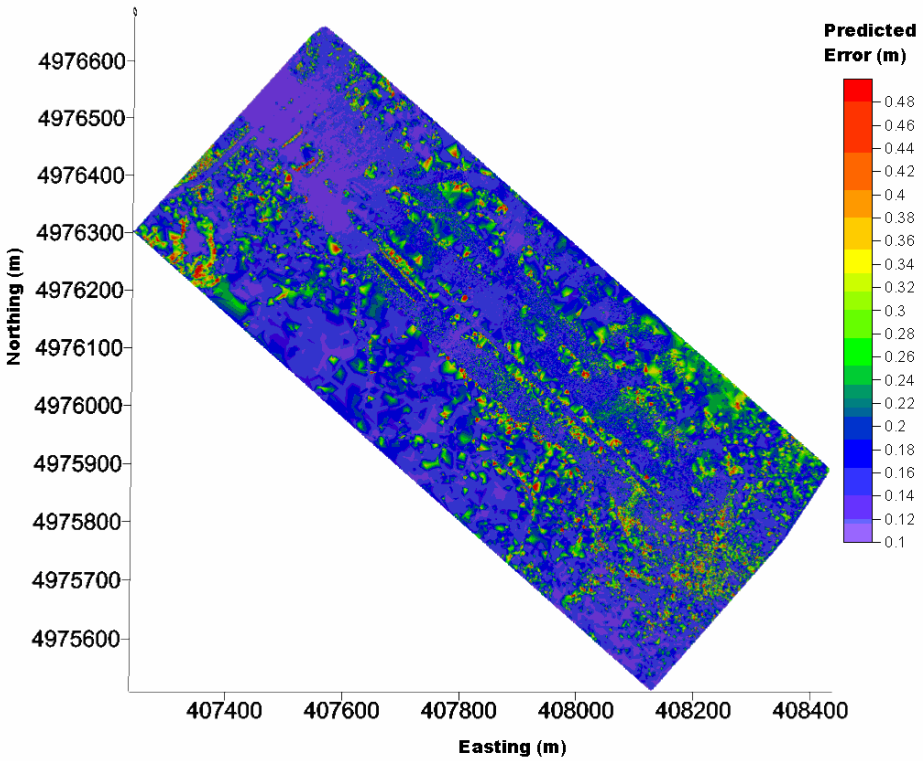


Figure 5.1

Vertical Error Predictions for Martock Dataset. The majority of errors range from 0.1 metres to 0.5 metres as is displayed on the colour bar on the right of the figure.

Figure 5.1 displays the clipped portion of the total dataset and is colour coded by the magnitude of vertical error. The colour range is set to vary between 0.1 m and 0.5 m as this contains the majority of the dataset. This allows for the largest possible colour variation within the ski trails. The majority of the high predicted vertical error values occur in wooded areas where the slope could be large if returns on adjacent points came from trees as well as ground points. Within the ski trails, the vertical error predictions remain relatively consistent between 0.15 metres and 0.24 metres although instances of errors above 0.4 metres do exist within these areas.

Unfortunately, reference GPS observations could not be completed at the time of the LiDAR survey although they were acquired during October of 2007. Therefore, both LiDAR observations and reference observations were taken in similar seasons when there was no snow accumulation. The grass on the hill had been recently cut prior to each survey and it is assumed that there hadn't been any significant changes in the terrain shape. However a slight systematic shift could exist because of different levels of vegetative growth at the time of each survey. GPS observations were taken using real-time kinematic observations with corrections being broadcast from a base station located less than three kilometres from the survey area. Observations were taken in the parking lot as well as on the ski slopes. Neither area had significant sky blockage which facilitated continuous lock to several satellites throughout the survey. This allowed GDOP to remain low and provided high accuracy observations. The GPS observations

are divided into four different categories, the first on the parking lot and the following three on separate trails on the ski hill labelled, Trail 1, Trail 2, and Trail 3 respectively.

5.1.1 Parking Lot Observations

The parking lot is a flat area and therefore can be used as a control surface to compare the observations on the ski slopes. These observations are important as they replicate typical quality control procedures for testing LiDAR surveys as outlined in Flood [2004]. The quality control procedures call for validation points for LiDAR surveys be located randomly throughout a flat area and statistics are generated between these points and associated surfaces generated from LiDAR observations. The advantage of observing flat areas is that it gives an impression of the error in the observations without additional error caused by terrain slope. Consequently, any horizontal error in the LiDAR observations will have no effect on the subsequent vertical error and the expected error void of this effect can be observed. Errors on sloped areas can then be compared with errors on the control surface. Figure 5.2 displays the locations of the GPS points in the parking lot on a shaded relief map of the entire clipped portion of the Martock LiDAR survey.

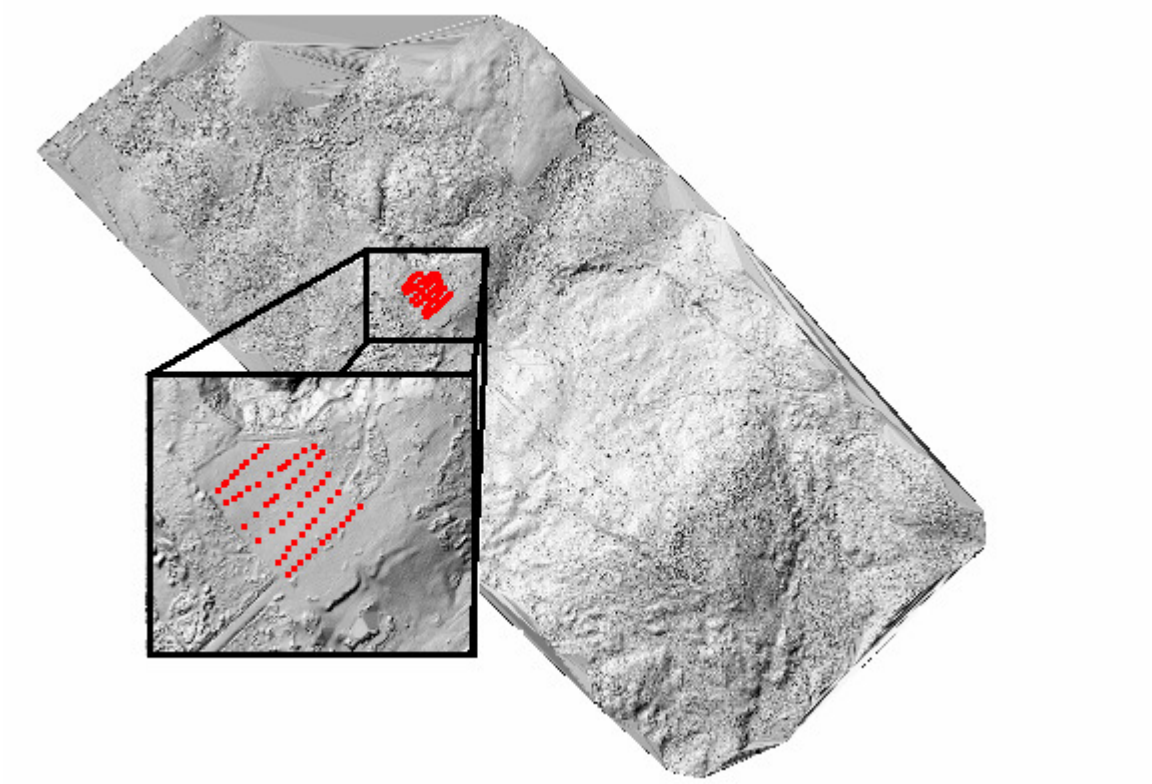


Figure 5.2
Shaded relief map of the clipped Martock dataset showing the parking lot area containing 65 GPS validation points shown as red dots.

The dataset of GPS observations consisted of 65 points. By comparing the vertical elevation values of these GPS points with the surface created from the LiDAR points the mean offset and standard deviation of the errors can be determined. The mean offset is likely due to some systematic influence to the observations, while the standard deviation represents the random components of the error. Once determined, this mean offset can also be removed from the rest of the dataset including the points on sloped areas, a practice which is common in LiDAR post-processing. Figure 5.3 displays the variation of the errors, termed residuals once the mean offset has been removed. The colors indicate

a range in centimetres of errors the residuals fell within. The intervals are designed to contain an equal number of points.

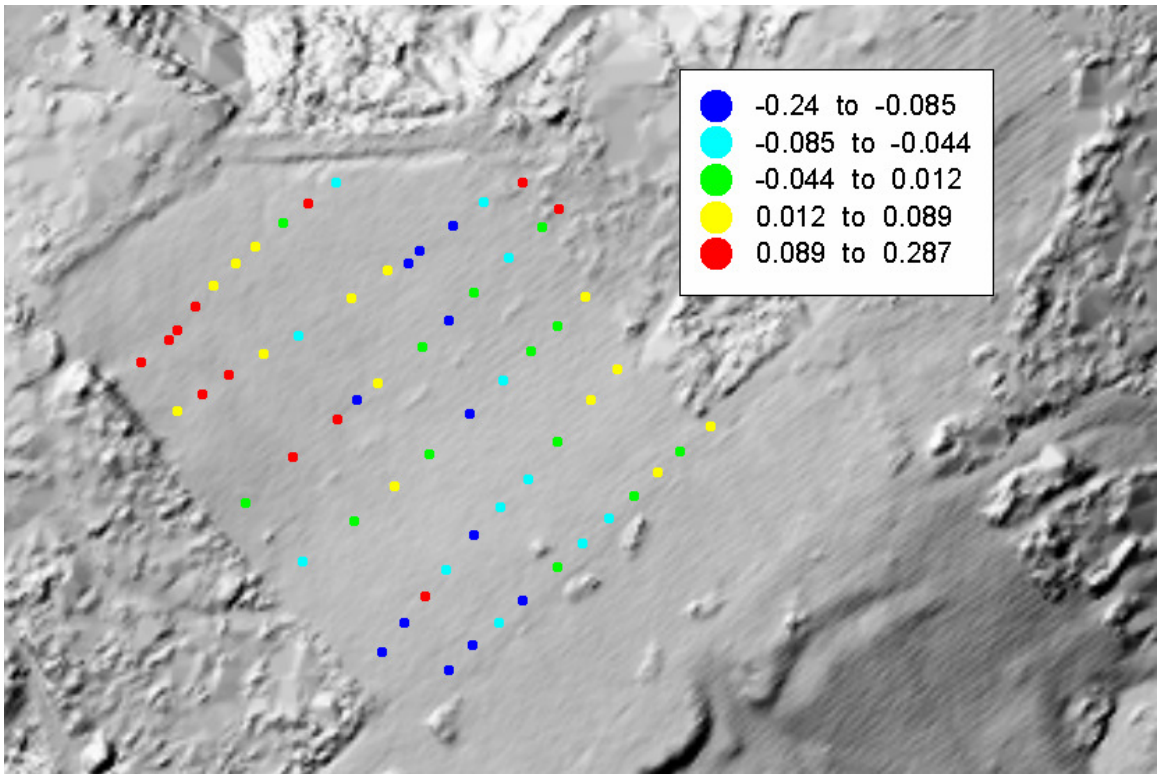


Figure 5.3
Residual Errors in Ski Martock Parking Lot

The resulting standard deviation of these residuals is 0.111 metres which is not atypical for LiDAR points on flat ground. The hardware error predictions for the points in the parking lot generally resulted in values of around eight centimetres. Considering that it is plausible that up to five centimetres of random error could be introduced through the RTK GPS observations, the hardware error predictions are reasonable. Although some residuals in the parking lot show magnitudes higher than what is predicted it should be remembered that the predictions are only at the standard confidence and several other sources of error affecting the LIDAR observations are unaccounted. From this it can be

concluded that the hardware predictions are reasonable estimates of the initial error, prior to the introduction of sloped ground.

5.1.2 Slope Observations

Observations on sloped terrain are divided into three trails, Trail 1, Trail 2 and Trail 3 which are located south to north on the ski hill respectively. Each trail has unique properties that affect the outcome of the testing results. Figure 5.4 shows the GPS observations taken on each of the three trails at Ski Martock.

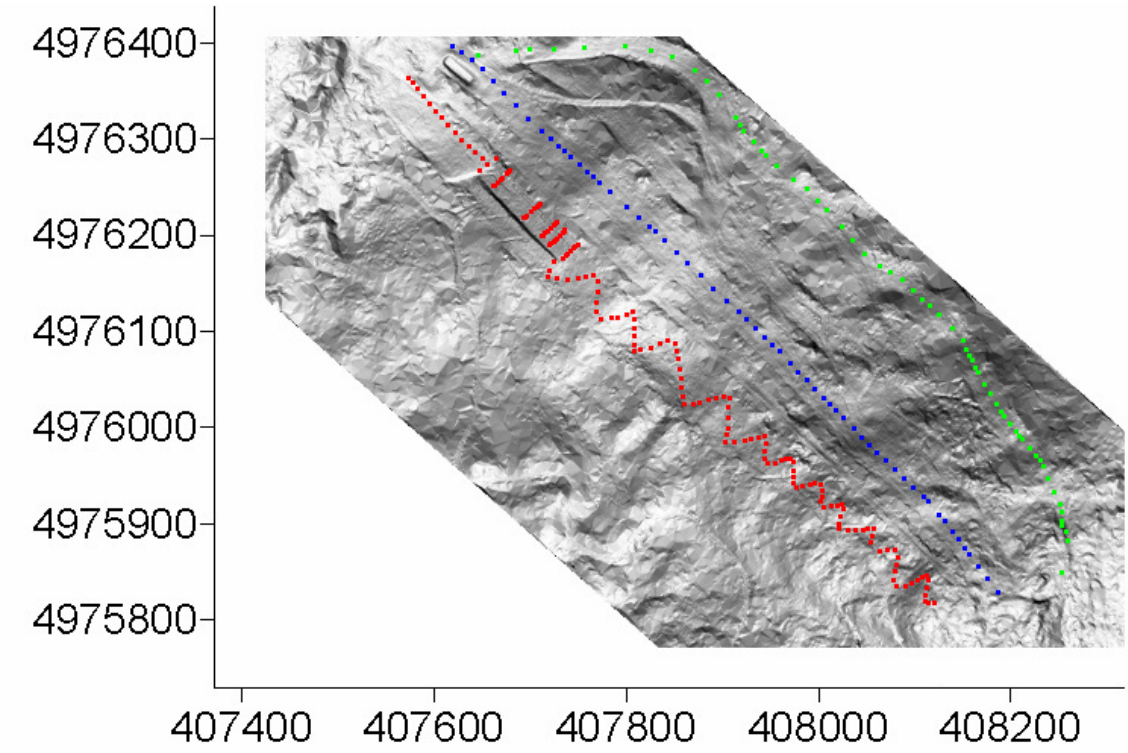


Figure 5.4
Shaded Relief Map of the Ski Martock Trails showing the GPS observations on each Trail. Trail 1 is represented by red dots, Trail 2 represented by blue dots and Trail 3 by green dots.

Trail 1 contained the most variable terrain with several large boulders throughout the slope as well as quickly changing slopes throughout. Also, near the end of the trail, a

'half-pipe' had been constructed. This feature is shaped as a 'U' and is approximately two metres from the top to bottom. Along its edges there are extremely steep slopes which approach 90°. Trail 2 cut the hill perpendicular to the contours throughout and did not contain any significant turns. The slope was smooth and consistent throughout the trail with a short horizontal section midway. Trail 3 cuts the hill at an angle to the contours and therefore contained slopes slightly less than trail one or two. Although the slopes changed throughout the trail the changes were smooth and generally consistent slopes persisted for large stretches.

One of the main problems that exists when validating LiDAR observations is that points do not generally fall in the same location as an observed validation point. Therefore, calculating the error between observed LiDAR points and observed reference points may result in differences that are due to the non-coincidence of position as opposed to errors in the observations. This is not a problem on flat terrain, as the points will contain the same elevation even if the reference point is not in the same location as the LiDAR observations. However, this becomes a significant problem on sloped terrain, where an offset in the horizontal position will create a large difference in the vertical calculation, as was discussed in Figure 1.2.

The proposed method for avoiding this shortcoming is to create a surface from the LiDAR points by linearly interpolating between them. If the LiDAR points are sufficiently dense, the surface should be an accurate representation of the terrain. However, it is possible that the terrain does not form a linear connection between all points and reference points located on these assumed connections may be in error due to

the interpolation and not actual errors in the observations. Despite this drawback, it is the best available method for testing purpose and is utilised here.

The first step in the testing procedure is the creation of the TIN surface. This was performed in Golden Software's Surfer program. Once the TIN was created, errors are generated by subtracting the vertical difference between each of the reference points and the generated TIN surface. This created a set of errors equal to the number of reference points. Figure 5.5 displays this calculation.

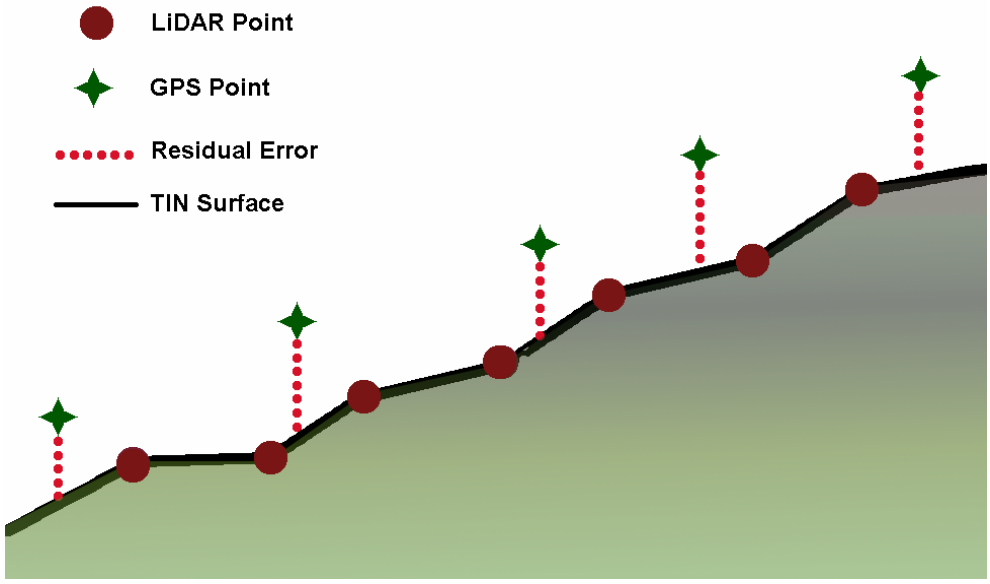


Figure 5.5
Testing procedure for calculating vertical errors between GPS validation points and LiDAR surface. LiDAR points diagrammed as red dots, with a linear surface shown as black lines connecting them. GPS validation points shown as red stars and the error is calculated as the vertical difference between the GPS point and LiDAR surface.

Acquiring this set of residuals is facilitated in Surfer software with the Residual function. This allows input of the TIN surface as well as the three-dimensional

coordinates of the GPS observations and returns the vertical distances between the GPS points and the surface, represented by the dotted red line in Figure 5.5.

Once the vertical error values are compiled several tests could be run to display the success of the slope algorithm. These same procedures were performed for each of the trails to facilitate comparisons between them and the parking lot data. The analysis will hereafter be divided into each trail as the unique characteristics of each trail warranted individual analyses.

5.1.3 Trail 1

Trail 1 contained 165 GPS reference points that were collected in a zig-zag pattern for most of the trail, and then in straight lines across the 'half-pipe' near the end of the trail. Table 5.2 outlines summary statistics of the dataset.

Table 5.2
Summary Statistics for Error Values on Trail 1

Statistic	Value (cm)
Mean	-10
StDev	53
Max	238
Min	-178

The mean and standard deviation indicate that the errors have increased compared to the control parking lot data as an additional systematic influence exists as well as significantly more random error. Several residuals near the end of the samples, between sample numbers 115 and 150, show extremely high magnitudes. These samples have occurred on the top of the half pipe half pipe next to a near vertical edge. The high residual values are most likely a result of the beam divergence. Since the last pulse

returns were used for processing, the position of the pulse could have been located on the top of the 'half-pipe', however a portion of the energy near the edge of the pulse may have missed the top and reached the bottom of the feature. Since this energy would contribute to the last pulse returns, the elevation would subsequently suffer. Therefore, these residuals can be considered a blunder as opposed to some systematic or random influence and are removed. The magnitude of all of the errors on the trail, including those on the 'half-pipe' are shown in Figure 5.6. This Figure is designed similarly to Figure 5.3 in which colours correspond to an error range that contains an equal number of points.

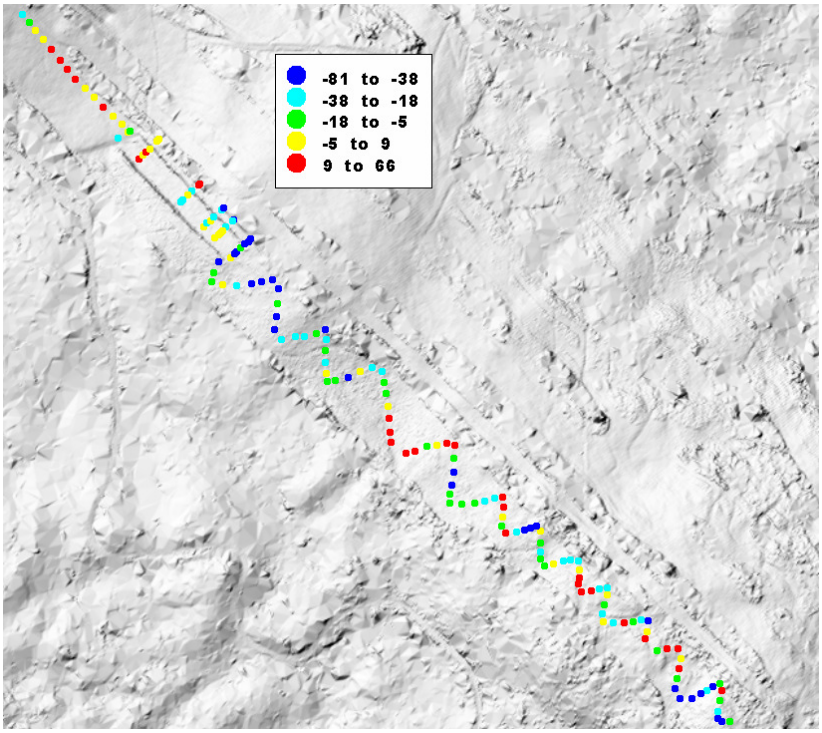


Figure 5.6
Plot of errors on Trail 1. Colors designate the range in centimetres the error falls within.

Figure 5.7 shows a histogram of the residual values. The values that are caused by the 'half pipe' are evident on the tails of the distribution. If these features are removed then the histogram appears as in Figure 5.8.

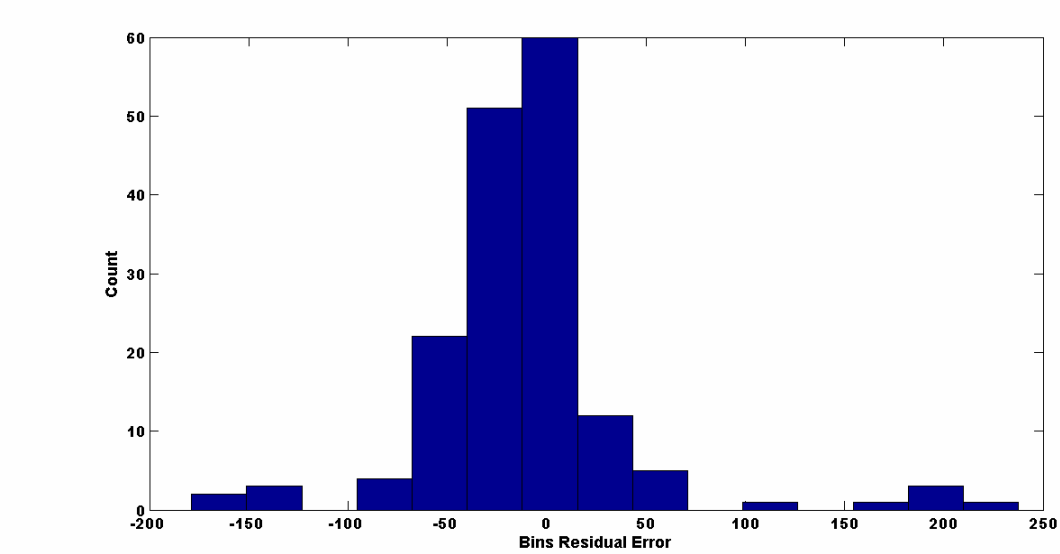


Figure 5.7
Histogram of Vertical Errors on Trail 1 prior to removal of blunders.

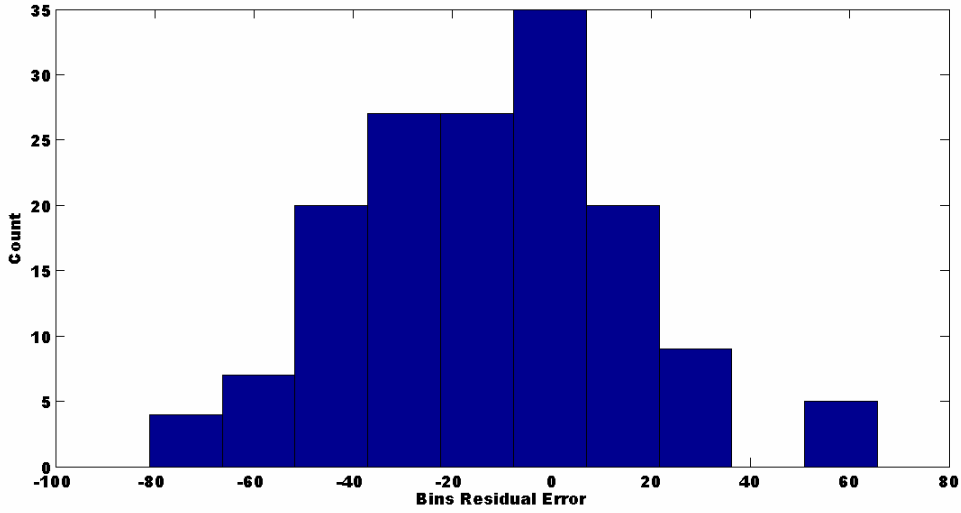


Figure 5.8
Histogram of vertical errors on Trail 1 after removal of blunders

The first test on the data considered whether observed histograms represented normal distributions. If results are not normally distributed then it can be concluded that the residual values are not a result of a random process. Since the initial horizontal and vertical errors due to the hardware components are a result of random processes, it follows that the subsequent errors due to slope will also distribute randomly. An Anderson-Darling test for normality was done to ensure the data was normally distributed (Anderson and Darling, 1952). The test was coded in Matlab and made available on the World Wide Web by Trujillo-Ortiz et al. [2007] and was performed at the 95% confidence level. This test concluded that the second histogram, after the residuals that resulted from the 'half-pipe' feature were removed was normally distributed. The updated statistics from this data are included in Table 5.3 and it is still evident that errors have increased significantly on the sloped terrain.

Table 5.3
Summary statistics for Trail 1 after removal of blunders

Statistic	Value (cm)
Mean	-13
StDev	27
Max	66
Min	-81

Prior to calculating the errors based on slope it is prudent to determine whether the residual errors that are being produced are correlated with the slope of the terrain. In order to perform this analysis a slope map of the terrain must be generated. A slope map will display the *x* and *y* positions of points and their associated slope value.

Slopes can be calculated with existing functionality within Matlab. The calculation must be performed on regularly spaced data, therefore, prior to performing

slope calculations a gridding of the LiDAR DSM points must take place. Gridding is a process which calculates the elevation value on regularly spaced points. Since actual observations do not fall upon the regularly spaced locations an interpolation must be done from original, irregularly spaced observations. The grid that was created for calculating slope was created at one-metre intervals in both the *x* and *y* directions. The default gridding method in Matlab creates a TIN surface with linear interpolations between data points. The grid point elevations are then generated from the surface of the TIN model at the specified regularly spaced grid points.

Once the grid has been created a slope is calculated on a single point in the grid by analyzing its neighbouring eight points. The following formula is employed in Matlab to calculate the slope and can be found in Wilson and Gallant [2000].

$$slope = \arctan\left(\left[\frac{dz}{dx}\right]^2 + \left[\frac{dz}{dy}\right]^2\right)^{1/2} \quad , \tag{5.1}$$

where:

$$\left[\frac{dz}{dx}\right] = (d-f) \Big/ (2*cellsize), \tag{5.2}$$

$$\left[\frac{dz}{dx}\right] = (b-h) \Big/ (2*cellsize\ ), \tag{5.3}$$

Where the letters, *a, b, c, d, e, f, g, h*, and *i* represent the following cells in a 3x3 grid surrounding the cell (*e*) in which the slope is being defined as follows,

a	b	c
d	e	f
g	h	i

and the cell-size is the spacing between successive points in the *x* and *y* direction which is chosen to be two in this case. Figure 5.8 displays the slope map for the section of the Martock LiDAR dataset containing each of the trails.

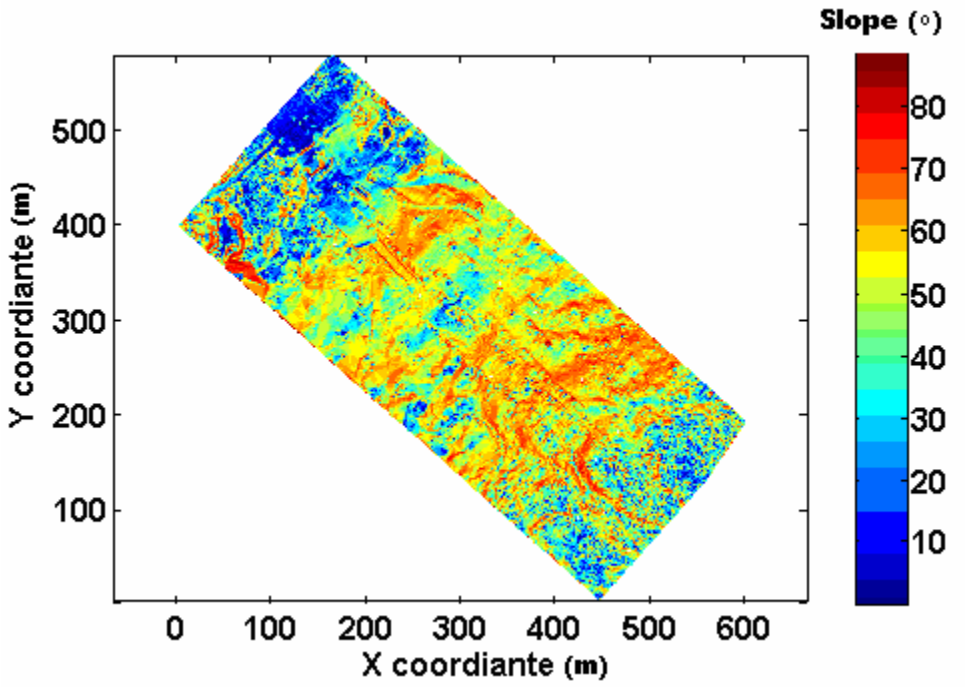


Figure 5.9
Slope map of Ski Martock trails. Colours represent the magnitude of slope at each grid cell with the colour bar being represented in degrees.

The determination of slope at each of the GPS validation point requires creating a surface of slope values. A similar procedure used to calculate the vertical errors can be used to calculate slope at each validation point. The coordinates of each validation point are overlaid onto the slope surface and the slope at these points is extracted. Once determined the vertical errors are plotted against the slope to determine the correlation between the two variables. This plot can be seen in Figure 5.10.

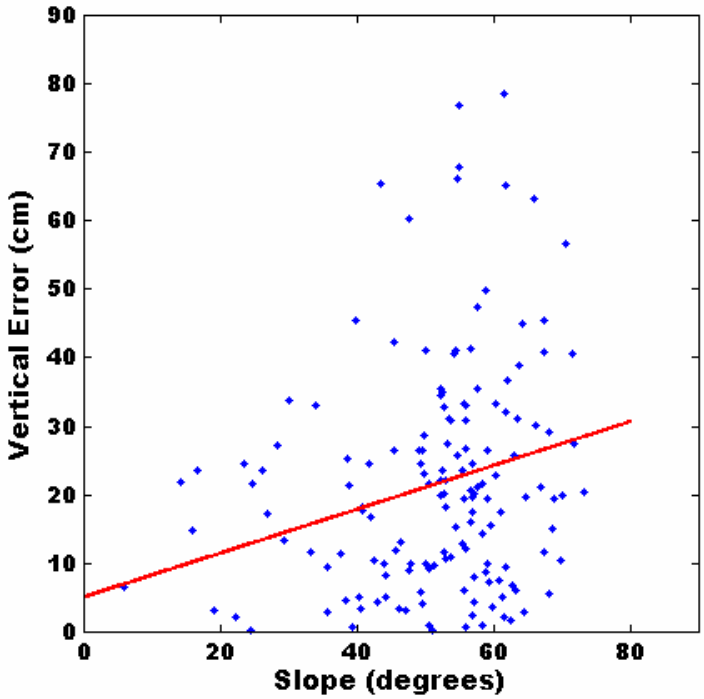


Figure 5.10
Trail 1 slope values plotted against vertical error. Blue dots represent the data point locations while the red line represents a linear line of best fit through the points.

Theoretically, an exponential curve is known to represent the phenomena. However a linear curve was the best fit to the experimental data, although little correlation actual existed. This could result from the fact that very little data was available in lower and higher slopes where the asymptotes of an exponential curve would be most evident. In addition, recall Figures 3.9 and 3.10 which diagram the error that would result in slope calculations from LiDAR observations. These figures indicate that the slopes could contain errors of over 20º making it difficult to match a theoretical model. Most importantly, the linear trend does imply that errors are increasing with slope as was postulated.

After the correlation between the magnitude of slope and vertical error is established, the predicted error can be compared to the observed error. Recall that the predicted error is calculated at the horizontal location of each LiDAR observation. Therefore, in order to define the predicted error at the validation points a method similar to the calculation of the residuals is used. A surface of the predicted errors is created with a TIN model, as seen in Figure 5.1 and the values of the surface are extracted at the GPS reference points. As a consequence the predicted error is interpolated linearly between each LiDAR observation. Theoretically the error due to slope will not vary linearly between observations, and this is likely only a reasonable estimate on gradually changing slopes. Unfortunately, this is the sole method for determining the predicted error values at the GPS reference points. This is a consequence that must be accepted and considered in subsequent analysis

A comparison of the residual errors, predicted errors with slope considerations, and predicted errors based solely on vertical hardware errors consideration, are plotted in Figure 5.11. The residual errors appear in blue, the predicted error with slope included appears in red and the vertical errors from hardware components in green; all are shown at the 95% confidence level. Initial error estimates for the input to the slope based error predictions were at standard confidence to display a more realistic level of error. The systematic bias in the observed errors has also been removed by shifting the values by their mean, leaving only the random influences.

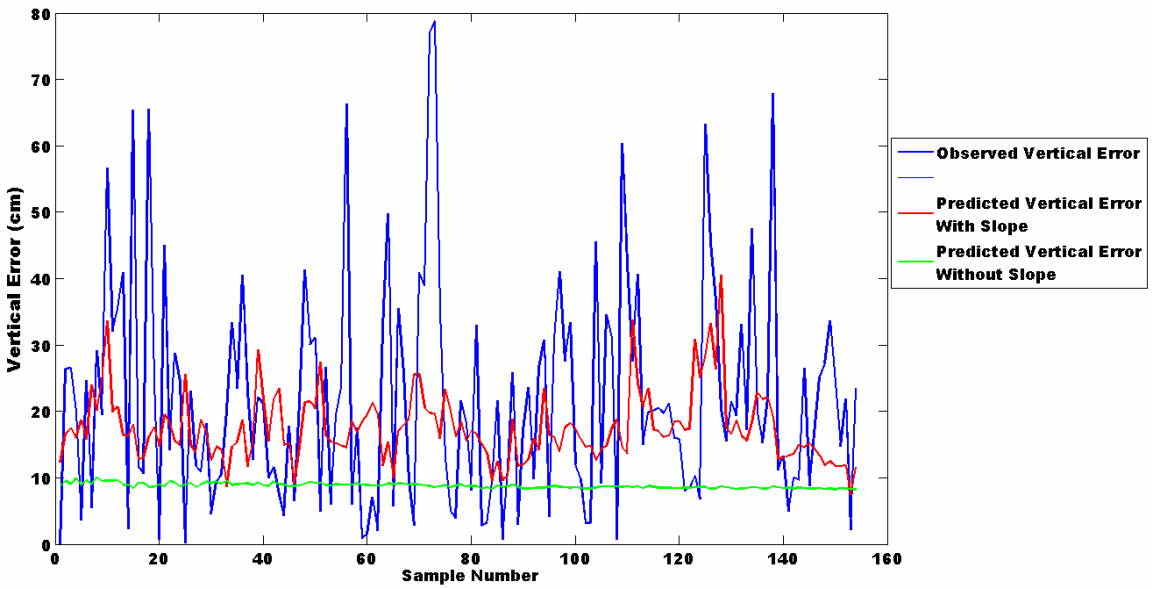


Figure 5.11
Comparison of observed and predicted errors on Trail 1. Observed errors are represented in blue, predicted vertical errors determined from the error prediction routine displayed in red and vertical error predicted from only hardware components shown in green.

The immediate point that can be noticed from the graph is that the vertical error predictions from only the hardware components show a consistent under-prediction of the observed error. In addition, there is no correlation between the predicted errors from hardware components and observed errors. The predicted error based on slope seems to follow the actual errors more closely. In general, the predicted errors based on slope also seem to under-predict the actual errors. Calculating the percentage of the residuals that fall within the predicted values reveals that 27% fall underneath the hardware error prediction while 48% fall underneath the predicted error based on slope. Considering that 'worst-case' scenario predictions were taken into account when designing the slope algorithm it seems unsuccessful when only 48% of the residuals fall underneath the error prediction.

This could result from the fact that the initial error predictions based on hardware errors were performed at standard confidence, indicating that the error ellipse contains only 68% of the observations. Had the error ellipsoid been scaled to the 95% confidence interval, the ellipse would have been larger and the resulting errors based on slope would have increased. In addition, and more importantly, there are several other errors that have not been taken into account, such as those due to beam incidence angle, atmospheric effects, and range based intensity effects. These factors would also increase the resulting errors horizontally and vertically. Subsequently, on highly sloped terrain this would cause a more pronounced increase in error due to slope. Another consideration is the fact that the actual error, and the predictions were linearly interpolated between the LiDAR points. This interpolation could cause a misrepresentation of the results if the terrain does not follow this linear trend.

5.1.4 Trail 2

Trail 2 contained 61 reference GPS points. The trail was a continuous slope and therefore most points were obtained in areas that are similar in the level of slope. Reference to Figure 5.9 shows that this area of the hill displays slopes around 50°. Table 5.4 displays the summary statistics for the dataset.

Table 5.4
Summary statistics for Trail 2

Statistic	Value (cm)
Mean	26
StDev	24
Max	83
Min	-27

Figure 5.12 displays the histogram of errors along Trail 2. The histogram follows a normal distribution according to the Anderson-Darling test indicating that the obtained errors were randomly distributed.

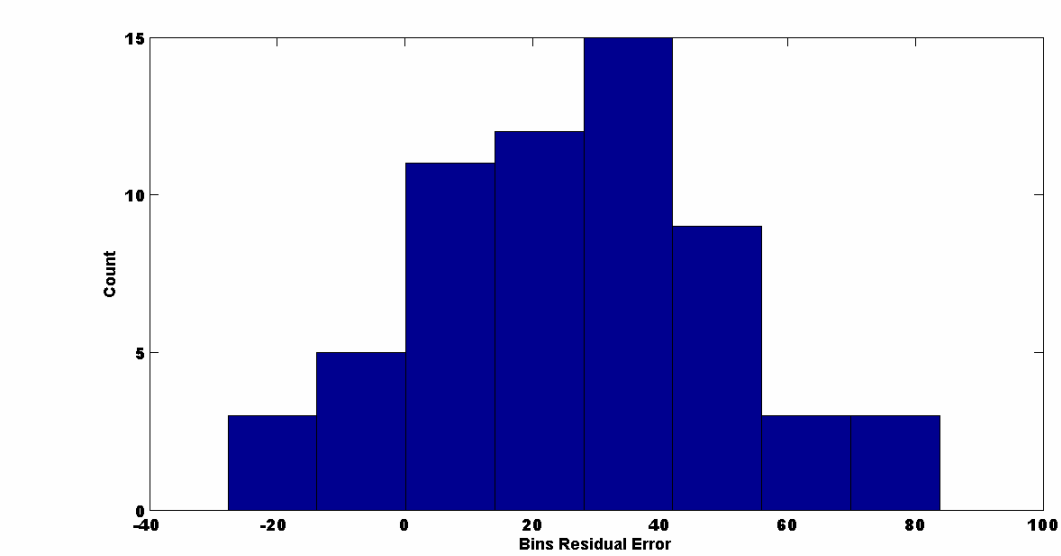


Figure 5.12
Histogram of vertical errors on Trail 2

The standard deviation between Trail 1 and Trail 2 agree well. The most striking difference is shown in the values of the mean which is different by 39 cm. Indicating that each of these trails contained separate systematic influences which have biased their vertical differences. This could be explained by an unaccounted terrain effect which had caused the systematic shift such as different levels of vegetation between the LiDAR survey and the validation survey. In addition, the trails were located in different portions of the swatch width in two separate strips of the survey. This could cause a slight systematic shift in the final point coordinates due to drift of the IMU observations or differences in GPS satellite constellation, accounting for part of the discrepancy.

The comparison of slope against the observed residuals results in Figure 5.13. The consistent nature of the slope in Trail 2 is easily identifiable in the figure as the slope values are contained within the range of 50 to 65°. This is unfortunate as it does not allow for an accurate trend line to be defined though the data series because there is no data over flat terrain or in the upper slopes. Nevertheless, the linear trend is similar to the trend identified on Trail 1 (Figure 5.10). The points seem to conform to the expectation of an increase in error with slope. Figure 5.14 shows a plot of the validation points on a shaded relief map of the ski hill. Similar to Figure 5.10 the colours indicate a range in centimetres the error fell within.

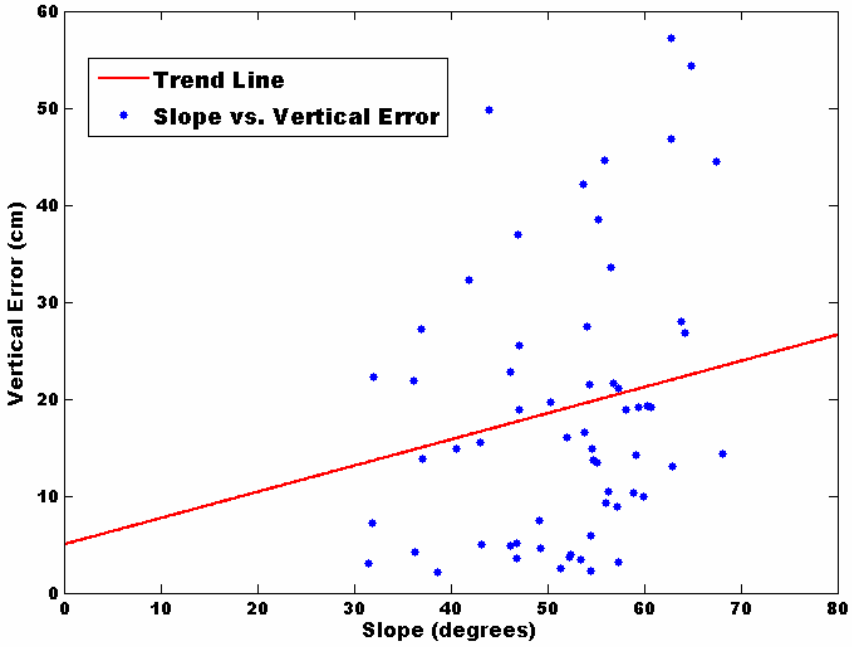


Figure 5.13
Trail 2 slope values plotted against vertical error. Blue dots represent the data point locations while the red line represents a linear line of best fit through the points.

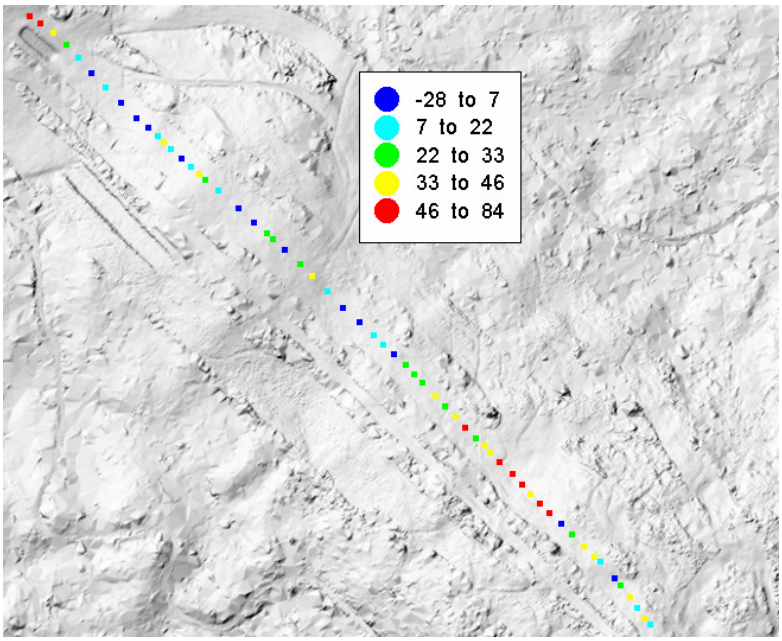


Figure 5.14
Plot of errors on Trail 2. Colors designate the range in centimetres the error falls within.

By plotting the observed vertical errors, predicted errors including only hardware components, and predicted errors including hardware and slope, Figure 5.15 is achieved. Similar to Figure 5.10 the hardware errors seem to drastically underestimate the observed errors. Similarly to the results observed for Trail 1 the predicted errors generated from the slope algorithm tend to follow the actual observed errors. Since initial error estimates were made at standard confidence this is as expected. This trend is more evident in Trail 2 than Trail 1 and is likely because the slope was consistent, thus resulting in reduced errors due to high variations in the terrain such as sharply changing slopes or surface features such as boulders. This situation is not as preferable as the worst-case scenario predictions the algorithm was designed for, however it is more acceptable than an optimistic estimation as it allows for greater assurance of the quality of the data. Most

likely, additional errors have affected the results causing some observed error values to spike above predicted values.

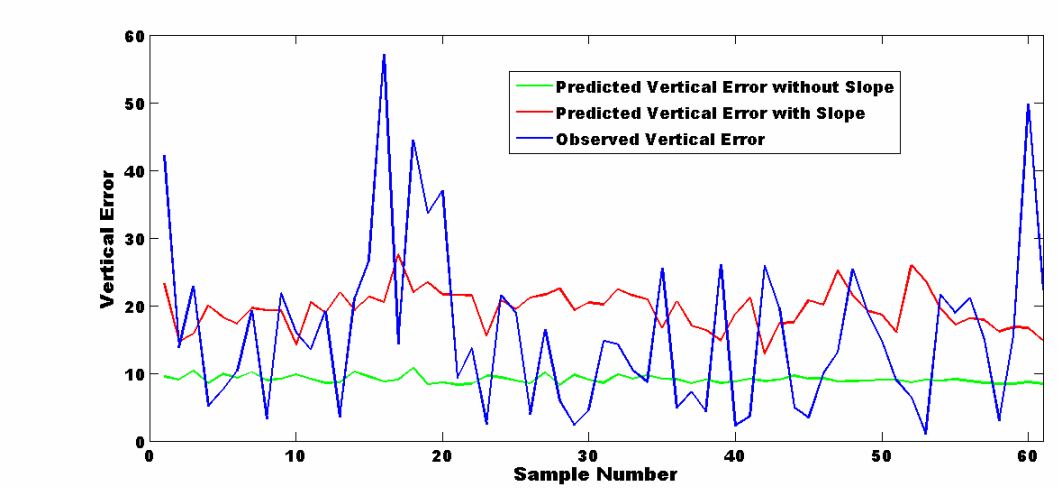


Figure 5.15
Comparison of observed and predicted errors on Trail 2. Observed errors are represented in blue, predicted vertical errors determined from the error prediction routine displayed in red and vertical error predicted from only hardware components shown in green.

5.1.5 Trail 3

The reference data for Trail 3 contained 61 GPS points. The slope within the trail was generally consistent throughout, although the lower half of the Trail does not cut the terrain perpendicular to the contours, creating a longer trail than Trail 1 or Trail 2 and a gentler slope. Figure 5.16 shows the histogram of the vertical errors obtained from the GPS observations and the LiDAR surface for Trail 3.

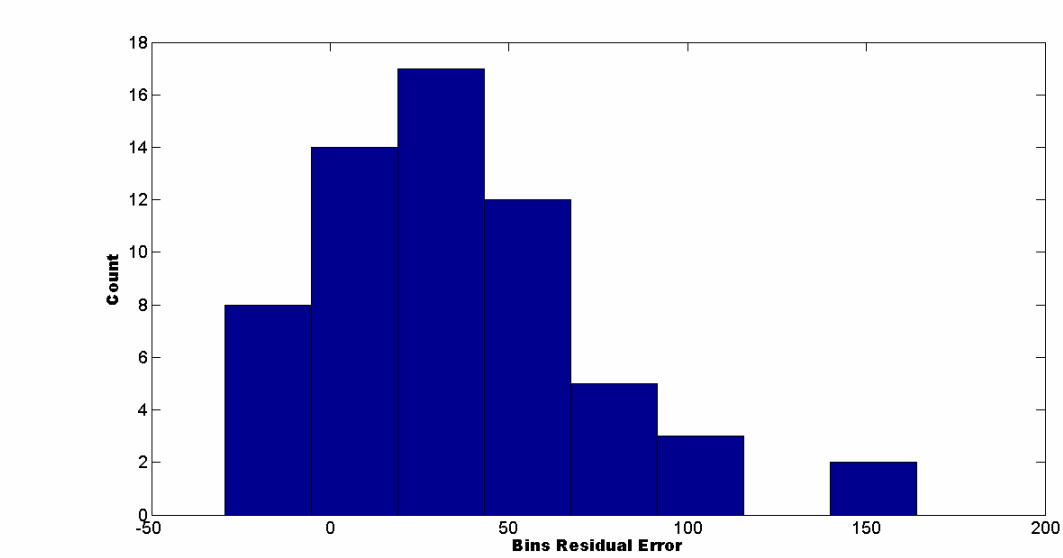


Figure 5.16
Histogram of vertical errors on Trail 2

The Anderson-Darling test fails at the 95% confidence interval indicating that the histogram is not normally distributed. The histogram appears to take on the shape of a normal distribution curve with a slight skew to the right. The two points on the far right of the distribution with errors in the range of 150 cm are both located in close proximity at the beginning of the trail. Perhaps an unaccounted terrain feature such as a boulder caused the discrepancy. The removal of these two points allows the histogram to pass the test for normality leading to the conclusion that the sample is normally distributed. However, without appropriate justification the points are left within the data set with the conclusion that the dataset is indeed normally distributed.

Table 5.5 shows the summary statistics for the data obtained from Trail 3.

Table 5.5
Summary statistics for Trail 3

Statistic	Value (cm)
Mean	36
StDev	39
Max	164
Min	-29

The mean and standard deviation display a similar trend as Trail 2, however are slightly larger. The results are slightly biased by the values at the tail of the histogram discussed above. Without these values the Trail 2 and Trail 3 statistics conform much closer as can be seen in Table 5.6.

Table 5.6
Summary statistics for Trail 3 with outliers removed

Statistic	Value (cm)
Mean	31
StDev	32
Max	109
Min	-29

The vertical errors show a similar positive systematic bias likely from similar nature of the terrain in the trails as each trail contains similar slope values without many terrain features or quickly changing slopes.

Plotting the slope along Trail 3 against the observed errors Figure 5.17 is produced.

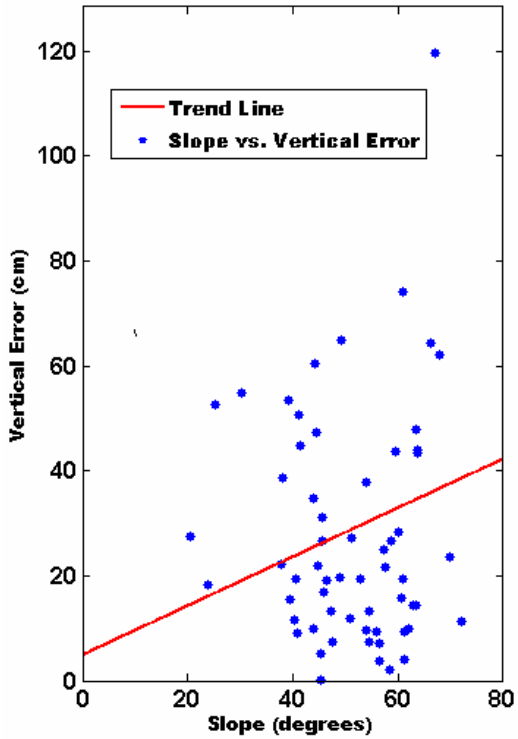


Figure 5.17
Trail 3 slope values plotted against vertical error. Blue dots represent the data point locations while the red line represents a linear line of best fit through the points.

The linear trend displayed in the graph is similar to the linear trend in the error analysis of Trail 1 and Trail 2 although the slope of the trend line is slightly steeper. This data displays the expected trend that residual errors will increase with the slope. Figure 5.18 displays the validation points and their associated errors on a shaded relief map of the Martock ski hill. Similar to Figures 5.6 and 5.14 the colours represent a range the error falls within.

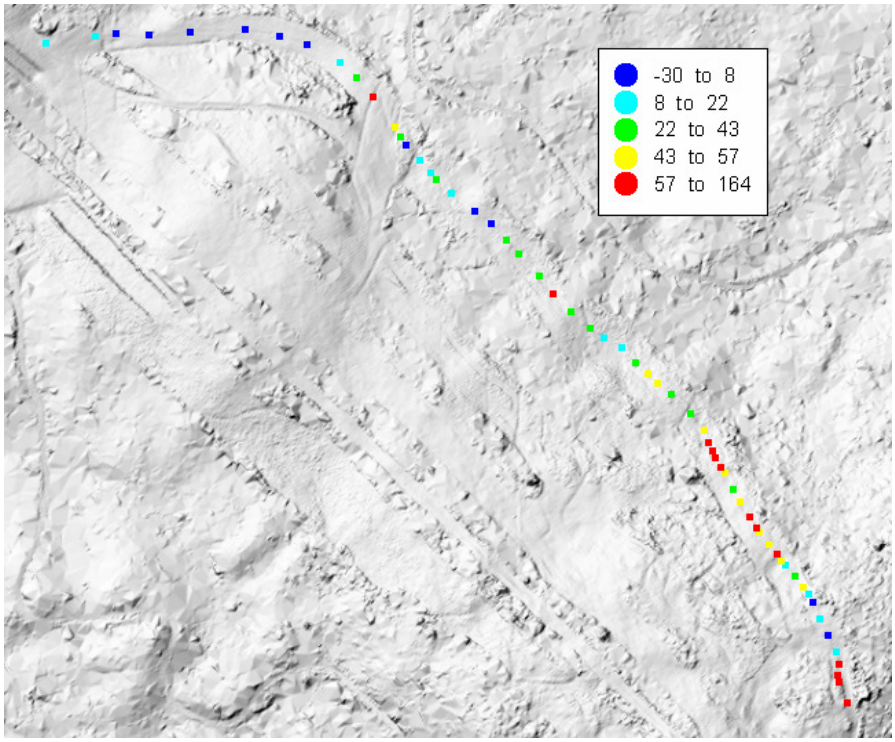


Figure 5.18
Plot of errors on Trail 2. Colors designate the range in centimetres the error falls within.

Figure 5.19 displays the predicted vertical errors of hardware components, the predicted vertical error including slope as well as the observed residuals for Trail 3. With the exception of the exceptionally high residuals near the beginning of the plot, corresponding to sample numbers two and four, the predictions represent the actual error well. Since the trail contained a consistent slope, similar in value to the slope of Trail 2, the predicted and observed errors remained relatively consistent. A few observed errors increased significantly above the predicted values and this is most likely due to interpolation errors introduced in the representation of the terrain. Additionally, unaccounted for errors in the observed LiDAR points responsible for larger horizontal and vertical errors in the initial predictions would propagate more severely into the slope error predictions. These trends are consistent with the other trails and consistent

conclusions can be drawn, the most striking being that predictions based on only hardware components do not represent error quantities accurately.

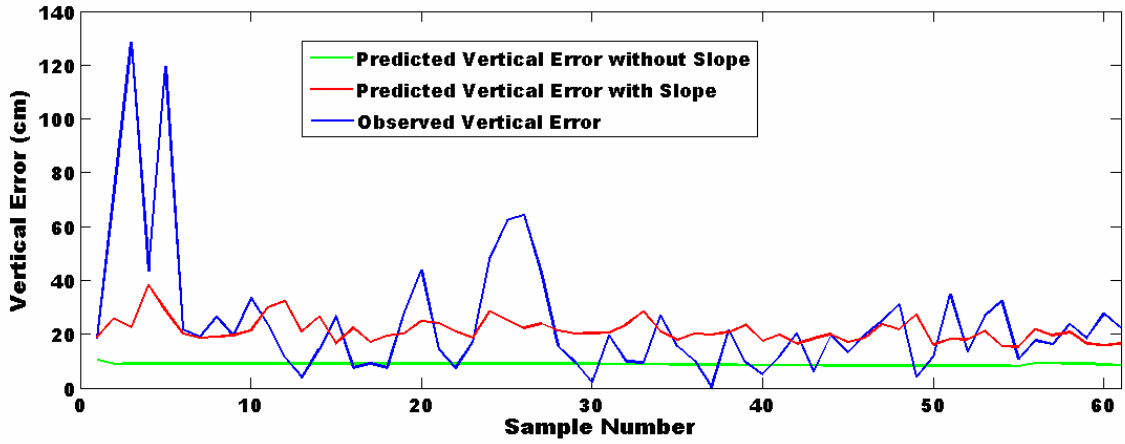


Figure 5.19
Comparison of observed and predicted errors on Trail 2. Observed errors are represented in blue, predicted vertical errors determined from the error prediction routine displayed in red and vertical error predicted from only hardware components shown in green.

5.2 Fredericton Test Site

To avoid the interpolation problems that were inherent in the first testing scheme a second test site was chosen. The site consisted of a sloped road near the campus of the University of New Brunswick in Fredericton, New Brunswick. A road was selected to reduce the effects of errors caused by the interpolation of the surface. A road surface is ideal since it can be assured there are no artefacts, such as boulders or mounds that deviate from the linear interpolation. Therefore, it can be concluded that the observed errors are due strictly to the effects of the condition of the terrain, atmospheric conditions and the errors in the system and are void of blunders due to interpolation.

The Fredericton survey was flown on May 20th, 2006. The entire survey consisted of 14 strips and required 75 minutes of flight time. Table 5.7 details the parameters used

for the survey. After processing, the points from last returns only were utilized in order to achieve the best representation of the ground surface.

Table 5.7
Survey parameters for Fredericton LiDAR survey

PRF (kHz)	Scan Frequency (Hz)	Scan Angle (degrees)	Flying Height (m a.g.l)	Flying Speed (knots)	Resolution (p/m^2)
50	33	18	1000	110	1.2

During the validation survey GPS points were collected over the road surface using post-processed kinematic (PPK) observation techniques. The base station used in the processing was a Canadian Active Control System monument (CACS) located within three kilometres of the site. The validation process is different from the process used at Ski Martock in that instead of forming a surface with the LiDAR observations and accumulating vertical errors between the LiDAR surface and GPS validation points, the GPS observations were observed sufficiently dense to allow for the creation of an entire reference surface. This was implemented by observing points at the road edge and near the centre and crown of the road. This observation technique includes the lowest and highest points on the road surface therefore the linear connections in a TIN representation surface will conform closely to the reality of the road surface. Subsequently, each LiDAR point could be subtracted from the reference surface, as opposed to GPS validation points being subtracted from the LiDAR surface. This increases the sample set allowing for more robust statistics. Figure 5.20 displays the TIN surface of the road and the LiDAR points that are included within it.

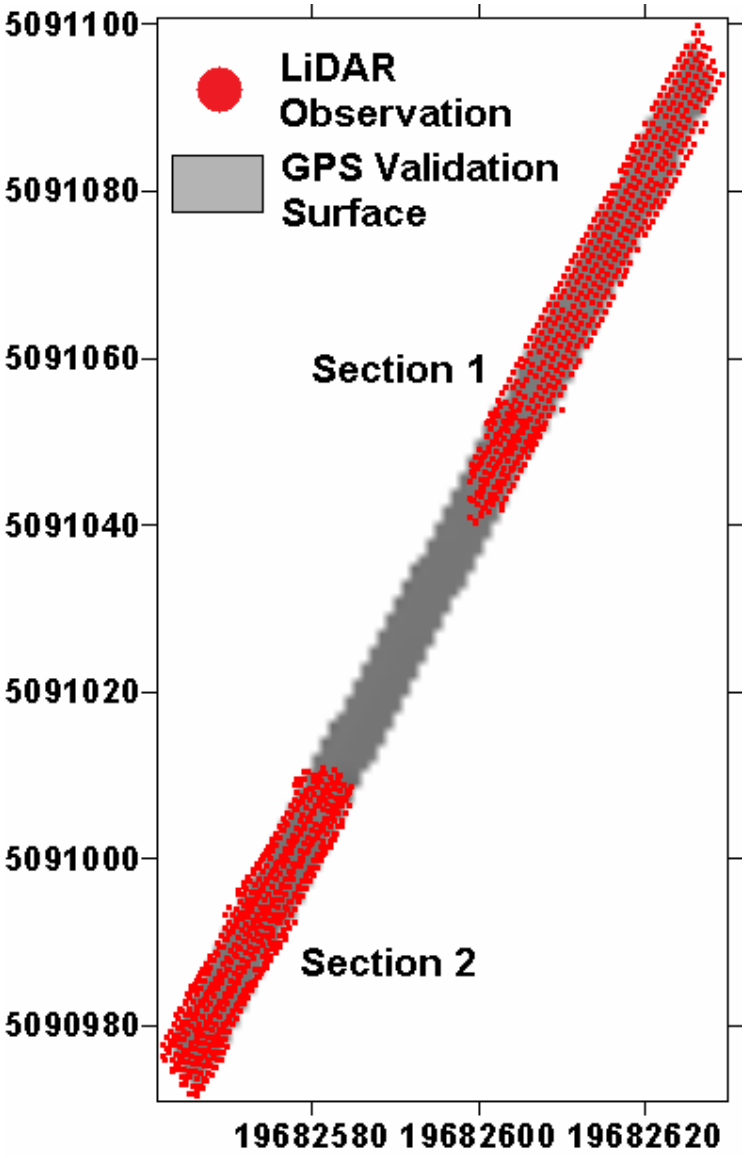


Figure 5.20
TIN representation of the road reference surface obtained from PPK GPS observations and LiDAR points that exist on the road surface. Two sections were acquired due to a loss of GPS lock during the GPS survey.

During the acquisition of the GPS reference data, there was a loss of lock to the satellite constellation. When performing a kinematic survey of this nature the 'on-the-fly' re-initialization must be performed in order to solve for the integer ambiguity after loss of lock to the satellite constellation. This is problematic for post-processed kinematic

surveys as during the data collection knowledge of these circumstances is not available. The realization that there had been a loss of lock came during the post-processing of the observations. Consequently, there is a gap in the GPS collection where re-initialization was occurring. Subsequently, the analysis was split into two sections, one on the northern part of the street where the slope is relatively less than the section at the southern part of the street where slopes are steeper. This divide is opportunistic since the relative difference in the slope values in these areas of the street allowing for additional comparisons to be made. A slope map of Windsor Street (the street used in the test) can be seen in Figure 5.21 and the variation of the slope from the north end of the street to the south end is evident. Slopes in the southern section generally ranged from approximately 8 to 12° while in the northern section ranged from 5 to 8°. Since each of the validation sections contained consistent slope values the plots of the error vs. slope included in the Ski Martock analysis section have been omitted as the lack of variation in the data would not provide for accurate interpretations. Analysis will hereafter be split into North and South sections.

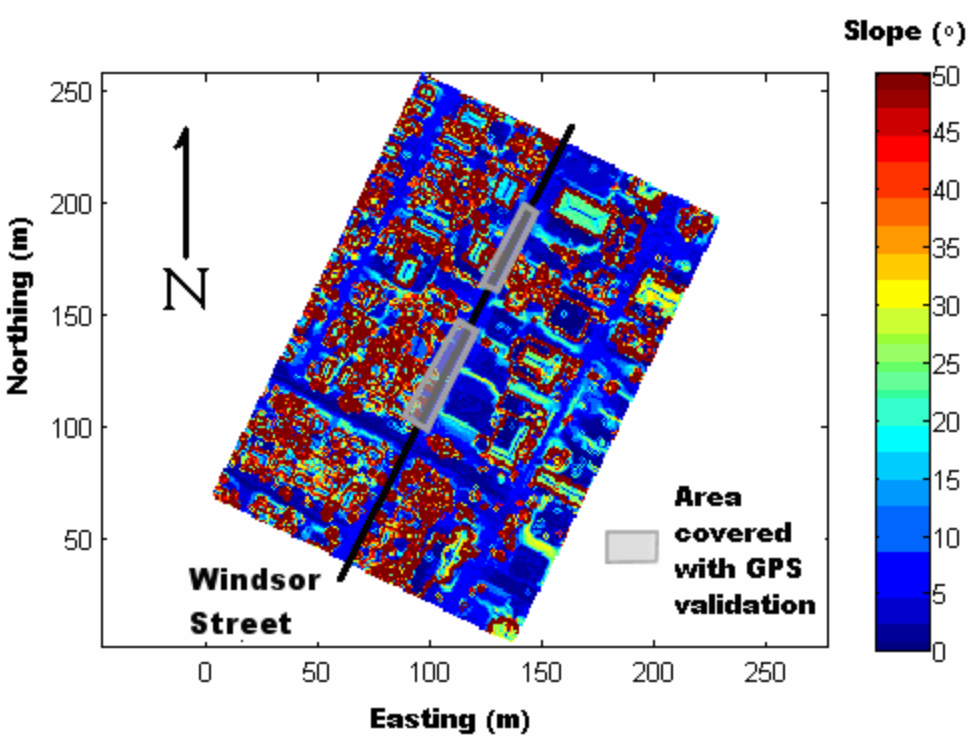


Figure 5.21
Slope Map of the Windsor Street Road. Black line indicates the approximate centreline of the road and grey areas show where the GPS validation data was collected in each of the sections.

5.2.1 North Validation Section

The northern validation area on Windsor Street contained a consistent slope just over 5°. There were 558 LiDAR points within the GPS validation surface available for testing. Table 5.8 displays the summary statistics for the determined vertical errors. The systematic shift is in the same direction as the shift on Trail 2 and Trail 3 of the Ski Martock data, but less in magnitude. This is reasonable as the slopes are not as steep indicating that a systematic bias due to the slope of the terrain would not be as prevalent. The standard deviation is also significantly less, likely as a result of the increase in the number of samples, the elimination of errors due to interpolation and the smooth terrain.

Table 5.8
Summary statistics for the vertical errors on northern section of Windsor Street

Statistic	Value (cm)
Mean	15
StDev	8
Max	35
Min	2

Figure 5.19 displays a histogram of vertical errors between the LiDAR points and the GPS reference surface. The histogram does not display a normal distribution due to a left skew. Inspecting the spatial distribution of the magnitude of vertical error values it becomes evident that a large portion of the data that caused the skew results from the edge of a scan line that only slightly overlapped the southern portion of the validation section. It is likely that this scan line had some separate systematic effect that caused a shift of the error values and the skew of the overall distribution. This could result from a systematic drift in the IMU or GPS equipment between the scan lines. Additionally, the error at swath edges is known to be at a maximum compared to errors at nadir. If the data from that particular swath is removed the distribution appears as Figure 5.20 and although still does not pass the test for normality appears to be approaching a normal distribution with the exception of a reduced skew.

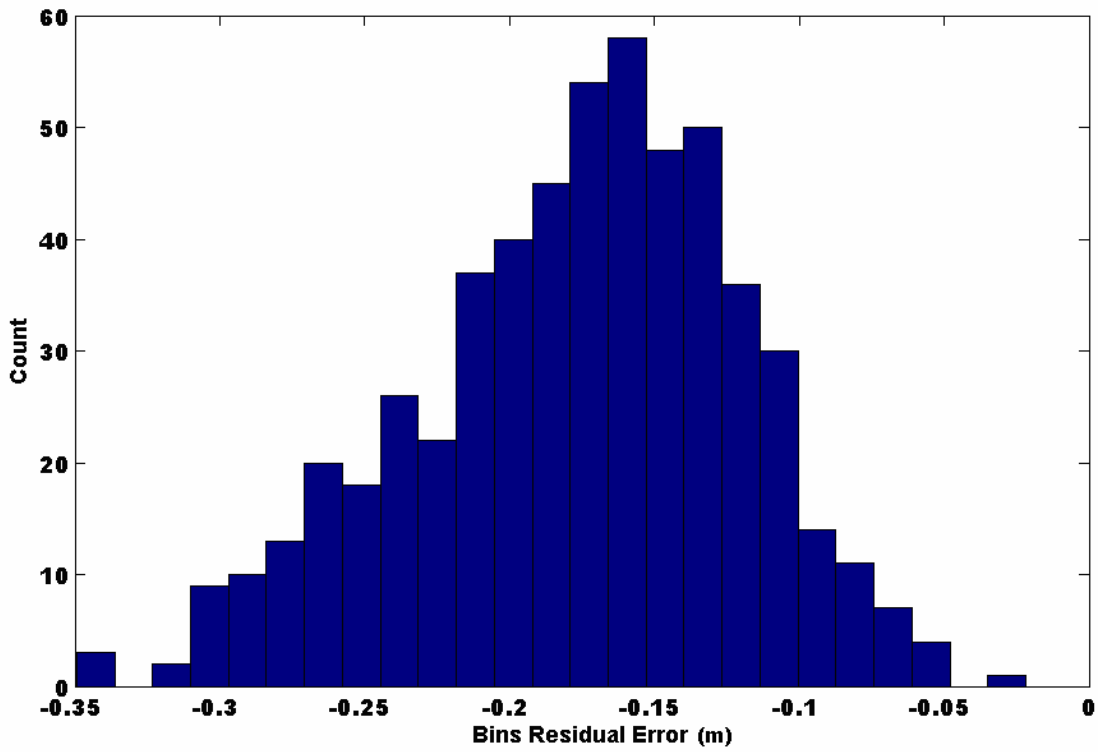


Figure 5.22
Histogram of vertical errors on northern section of Windsor Street

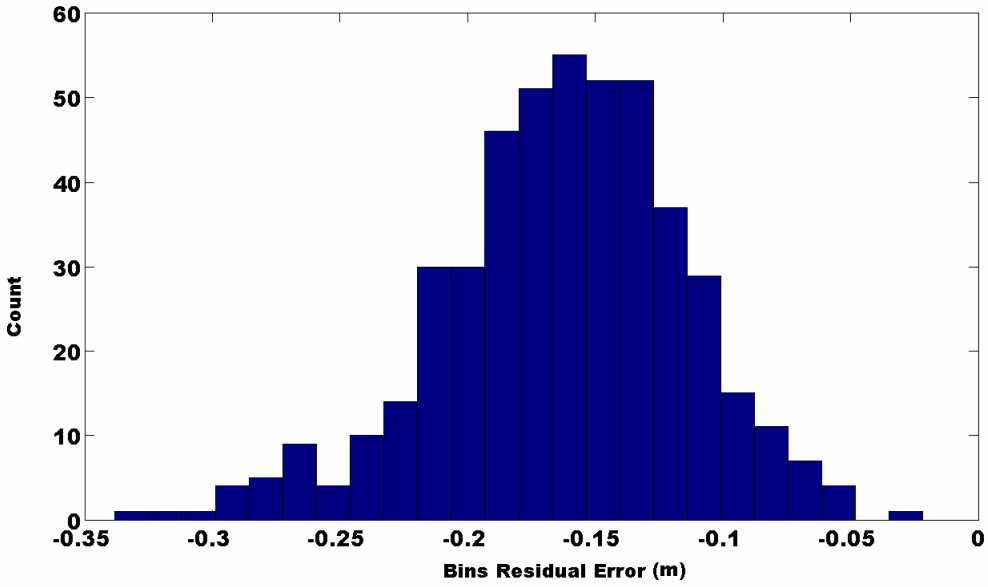


Figure 5.23

Histogram of vertical errors in the northern section of Windsor Street with values from one swath removed.

The residual errors causing the remaining skew are clustered in a single section of the validation surface that suffers from a removed GPS validation point due to large standard deviation. The resulting deviation from the normal distribution is a result of the validation surface being too sparse is this area and not representing the true terrain well. Therefore, it can be concluded the observations are from a normal distribution and have been randomly distributed. Based on this conclusion, both swaths are still kept for analysis as although the removed swath contained a slightly different systematic bias, the points still form normal distributions independently, indicating they are all the result of random processes.

Figure 5.24 shows a graph comparing the observed residual errors, the predicted error without the slope algorithm included and the predicted error when the slope algorithm is included. The predicted hardware errors are shown at the 95% confidence level and the predicted slope errors have been calculated with initial error ellipsoids also at the 95% confidence level. This was chosen in contrast to the Ski Martock data where initial error ellipses were at standard confidence. This was done to show the maximum values of the slope error prediction model.

Another explanation for the apparent systematic shift is that when collecting the Windsor Street data the entire road was perpendicular to flight lines. This causes all of the observations within the street to be gathered in a fraction of a second. This is in contrast to the data on the ski trails, which were parallel to flight lines and thus gathered over a longer period of flight time. The result of this is that all the data collected over

Windsor Street will be associated with only a single GPS airborne observation, and likely only a few IMU observations, as observations are sampled at 1 Hz and 200 Hz respectively. The random error in these observations will exhibit itself systematically within the points because they are all correlated with the sole GPS observation. A complete removal of all systematic error will also remove the error due to the GPS observations, which although is systematic within the points is random in a global sense. Therefore, a systematic bias equal to the magnitude of the vertical error in the GPS observations (5 cm) has been left in the data to represent this error component.

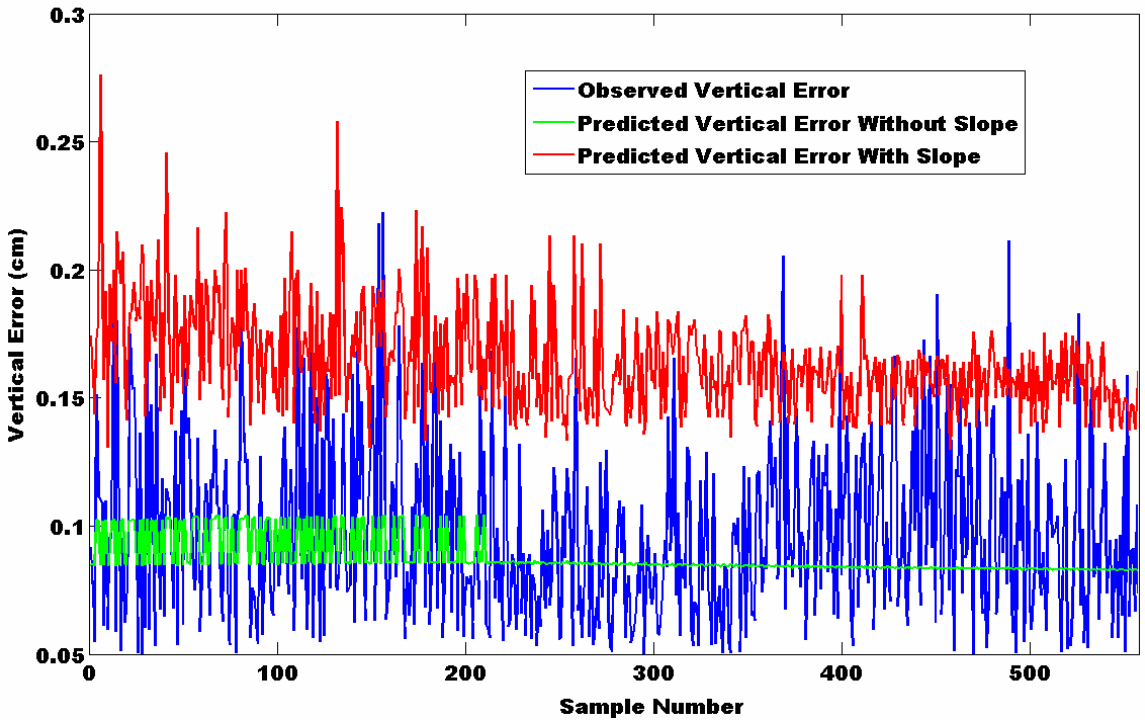


Figure 5.24
Comparison of observed and predicted errors on Trathe northern section of Windsor Street. Observed errors are represented in blue, predicted vertical errors determined from the error prediction routine displayed in red and vertical error predicted from only hardware components shown in green.

From Figure 5.24 it is seen that the observed residuals lie between the prediction of the hardware errors and the prediction that includes slope consideration. Rarely does the observed error reach the level of the predicted error with slope included. This is as the algorithm was designed, especially when the initial error estimates in the slope error algorithm are input at the 95% confidence level. Only very rarely will the actual surface be at this worst-case scenario condition for two adjacent points. However, as is evident in Figure 5.21 it will happen occasionally. Nevertheless, the true error does not increase drastically above the predicted values. The predictions are performing as expected since the observed errors are in general above the predictions based on hardware components only, however under the predictions based on slope. The percentage of residual values that are below the predicted hardware errors at 95% confidence is 49%, while 98% are under the predicted errors when slope is included.

5.2.2 South Validation Section

The southern validation section contained a consistently steeper slope than the northern section at approximately 12°. It contained 657 LiDAR points from two swaths. Each swath completely covered the validation area. Table 5.9 displays the summary statistics for the validation section. Results are similar to the northern validation section. The mean of the two sections are nearly identical, resulting from the similar slopes between the two sections. The standard deviation is also small resulting from the consistency in the terrain and slopes and the elimination of interpolation errors.

Table 5.9
Summary statistics for the vertical errors on sorthern section of Windsor Street

Statistic	Value (cm)

Mean	16
StDev	5
Max	28
Min	1

Figure 5.25 displays the histogram of the vertical errors.

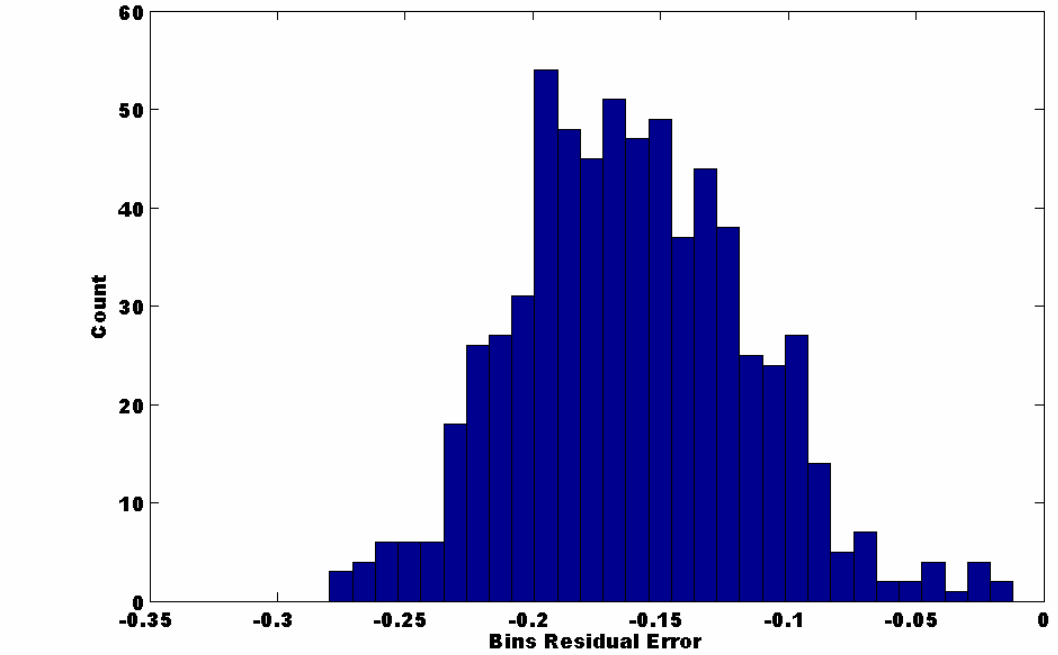


Figure 5.25
Histogram of vertical errors on southern section of Windsor Street

The values pass the test for normality at the 95% confidence level indicating that the samples were drawn from a normal distribution. Figure 5.26 displays the plot of the observed residual errors (blue) with the systematic error removed, the predicted vertical error from only hardware components (green) at the 95% confidence level and the predicted vertical error including slope with initial error estimates input at 95% confidence.

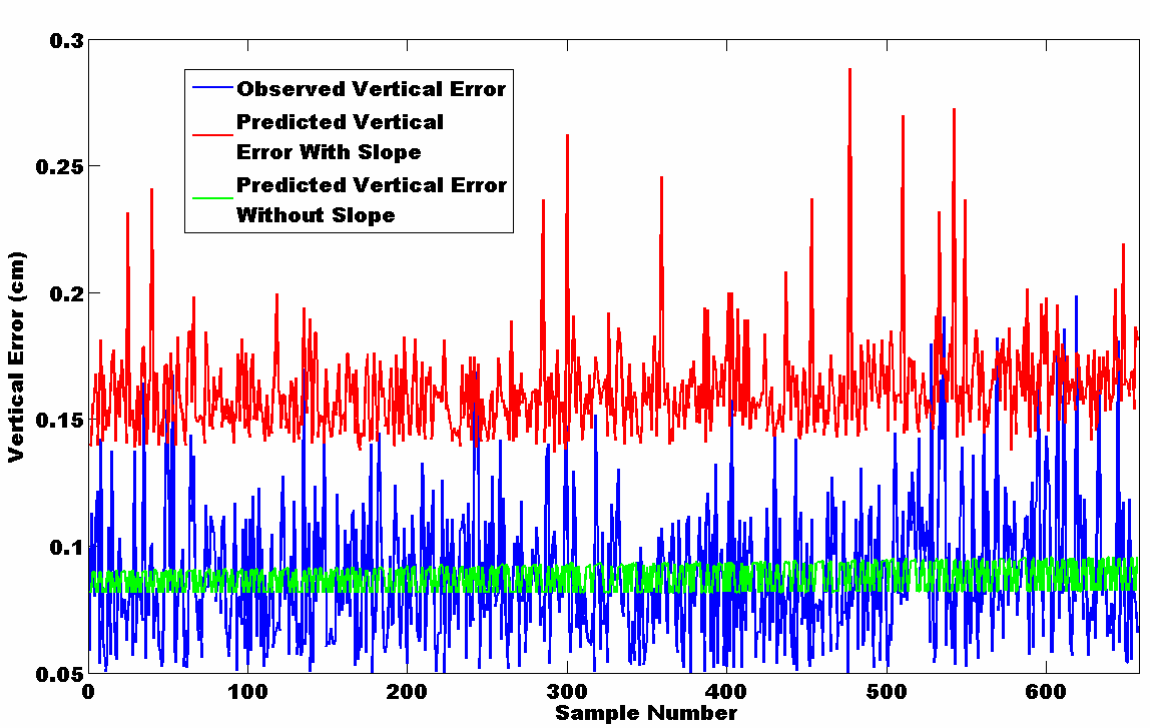


Figure 5.26
Comparison of observed and predicted errors on the southern section of Windsor Street. Observed errors are represented in blue, predicted vertical errors determined from the error prediction routine displayed in red and vertical error predicted from only hardware components shown in green.

The figure shows a similar trend to what was seen in the northern section. The observed errors are generally higher than the predicted hardware errors, with 58% falling underneath the predicted values, while 98% fell under the predictions which included additional slope based error. This is as is expected because the slope algorithm was designed to account for the worst-case scenario, which will occur in rare cases. The graph displays that the algorithm is reacting as is designed.

Chapter 6
Conclusions and Recommendations for Future Work

LiDAR is still a relatively new technology within the field of Remote Sensing and Geomatics. During its short history it has enabled significant changes to traditional DSMs of the past, especially pertaining to their density and accuracy. The quickly acquired, dense and accurate DSMs provided by LiDAR technology put the research focus on the many applications that benefit from these DSMs. With the large application base the technology quickly became a commercially viable method for acquiring DSMs, without research into inherent errors. Consequently, service providers and end-users are left confused about the capabilities of the system and the expected error that is contained within final results.

LiDAR manufacturers, such as Optech provide, error specifications in commercial literature, however, analysis seems to be overly simplistic and gathered under ideal conditions. Typically, these consist of reporting error levels at a standard confidence level, under smooth flight conditions, low GPS DOP predictions, and over isolated flat ground validation areas. Errors are also typically reported as single values, which do not reflect the dynamic nature of LiDAR errors that change with each laser pulse. Although this type of error reporting may be beneficial for commercial purposes, it does not leave service providers in a position to make informed judgements about the condition of the final data they are producing. The purpose of this thesis was to

investigate how these errors may change in sloped environments and provide an algorithm for predicting the maximum possible error in the observations due to slope.

Previous studies have shown that a correlation exists between the degree of terrain slope and the magnitude of errors observed on the slope. The relationship is positively correlated with errors increasing to a theoretical limit of infinity as slopes approach 90°. Previous studies only proved that this relationship existed by testing the level of error on different sloped areas. This was achieved by making observations of higher accuracy, through GPS or conventional surveying, and comparing the difference between LiDAR observations and those of higher accuracy. The information provided in these studies was valuable in that it indeed showed that error levels were increasing on sloped areas, and significantly above the level of errors viewed in the hardware components of the system. However, they made no attempt to quantify or predict the level of error based on slope. Therefore the conclusions are informative to those particular LiDAR surveys but provide little benefit to subsequent surveys or information on how to deal with these errors.

Prior to investigating the error based on slope, an analysis of the error due to the hardware components within the system was necessary to provide an initial estimate of error. These error quantities are the best documented of all of the LiDAR error sources, however, literature relating to their quantification is still being produced. Error estimates based on these components is developed by propagating the errors in each individual hardware component through the direct geo-referencing formula of LiDAR observations. This is facilitated by the General Law of Propagation of Variances (GLOPOV). To quantify the predicted error in the hardware components it is necessary to compile error estimates of individual hardware components observations which include, the GPS

relative positioning, Inertial Motion Unit attitude information, laser scanner and laser ranger. The appropriate values for the errors in each of these systems are open to dispute and therefore arguments were made to back the suggested error estimates which are listed in Table 2.1. Once propagated through the direct geo-referencing formula these errors combine to form three-dimensional error ellipsoids which serve as the initial error estimates for slope calculations.

Prior to performing the error estimates based on slope, the unique characteristics of LiDAR DSMs needed to be shown so that the differences compared to other DSM data collection methods could be appreciated. From this, the importance of error estimates based on slope in LiDAR DSMs can be easily seen. The characteristics of LiDAR DSMs that separate them from traditional DSMs were explained through various mission planning parameters of LiDAR surveys. From these parameters it can be seen how LiDAR surveys achieve their spectacular point spacing and how an operator can modify the way points are distributed on the ground. This distribution of points effects the determination of several quantities from the DSM, such as terrain slope.

Traditionally, the determination of properties of coarse DSMs didn't require rigorous error analysis of the individual points in the DSM because the interpolation error was great enough to render the error in the point observations insignificant. In addition, for a property such as slope, the final determination is greatly affected by point spacing and error propagation shows that errors in individual points only propagate significantly once the point spacing begins to reach levels of typical LiDAR surveys. This indicates that when utilizing LiDAR DSMs the way error is interpreted must change. Traditionally error could be viewed as unavoidable interpolation errors. As DSM models became finer

in newly available LiDAR surveys, the errors in the individual points become much more significant in the derived properties of the DSM. Therefore, the proposed error analysis in this thesis is significant to any subsequent analysis of LiDAR DSMs.

Following the discussions of the initial error estimations and the unique characteristics of LiDAR DSMs, the algorithm for determining slope could be developed. The algorithm needed to work on the 3D coordinates of a LiDAR DSM as well as have the initial error estimates of each individual point available. The algorithm operates by constructing a TIN surface from the LiDAR points and analyzing the edges of triangular faces that make up the surface. Each point in the surface is connected to several other points along edges of the triangular faces within the TIN. A slope can be calculated along these triangular edges.

With knowledge of the size and orientation of the initial error ellipsoid on each point along a triangular edge a two dimensional error ellipse can be calculated in a plane that contains the triangular edge and is perpendicular to the *x-y* plane. From this, an infinite number of realizations of the true terrain contained within the area of the two dimensional error ellipses are possible. However, there is one realization that will contain the maximum possible error at a particular point. This realization occurs when the slope line is tangent to both error ellipses. Once this realization is identified, the vertical difference between the point and the realization of the terrain can be calculated to determine error based on the slope between the two points. The resulting algorithm will produce results that show no additional error on flat terrain, and errors that increase exponentially as slopes increase which fit theoretical predictions.

Once the algorithm was developed it was tested on real datasets to determine if it is correctly predicting errors based on slope. To perform this task, two datasets were analyzed, that of a ski hill named Ski Martock in Windsor, Nova Scotia and a road in Fredericton, New Brunswick. Ski Martock was an ideal test site when attempting to predict how the algorithm would react on a glacial surface as both contained heterogeneous terrain, and the ski slopes were void of vegetation similar to glacial terrain. Validation data was obtained through RTK GPS observations throughout three sloped ski trails and a flat parking lot. Tests on these sites showed an improvement over error analysis based solely on hardware related errors which severely under predicted the observed error. The error predictions that were generated from the algorithm presented in Chapter 4 showed a better prediction of the error than the hardware predictions as they increased on the sloped surface.

Considering that the algorithm was designed to predict worst-case scenario error levels, the algorithm did not perform as pessimistically as would be expected on the ski hill data. This was likely caused by several other error components affecting the observed error values that haven't been considered in the error predictions which could include atmospheric effects, laser beam incidence angle causing smearing of the laser footprint, and intensity based range biases. Each of these errors would increase the total error in the individual points and would be propagated more severely in sloped environments. The situation is made more difficult by the resulting interpolation that was required in obtaining vertical errors and error predictions. Since control observations were not obtained directly on top of LiDAR points the resulting error estimates were obtained from linear interpolations between LiDAR points. If the terrain did not form a linear

connection then the error predictions will be incorrect. Also, because error predictions were made at each LiDAR point the error predictions were also linearly interpolated between points. The necessity for these interpolations would have negative effects on the compiled results but had to be accepted in the analysis.

To overcome the interpolation effects that were inherent in the first test site a second site was chosen along a sloped road (Windsor Street) near the University of New Brunswick Fredericton campus. The road surface was ideal because the assumption of a linear terrain interpolation is not going to adversely affect results. Therefore a coarse set of well placed validation points can be interpolated into a reference surface. From this reference surface each LiDAR point elevation can be subtracted from the reference surface to develop a sample set. This is advantageous because the sheer volume of LiDAR points that will exist within the validation surface allows for a significant increase in the sample size from which to generate statistics.

The statistics generated form the road surface showed improved results compared to the ski trails in that observed errors fell directly between predicted hardware errors and predicted errors from the algorithm designed in Chapter 4. This solution is expected since the predicted errors including the slope were designed as worst-case scenario predictions. The predictions were not severely pessimistic and seemed to contain the observed errors within a reasonable limit. These results were likely different from the ski hill results because there were little to no interpolation errors, the sample set was sufficiently large that outliers or blunders could be easily detected, and the surface was a smooth and consistent slope leaving little error due the pulse footprint encompassing quickly changing terrain or rogue features upon the terrain.

In addition, the strip of sloped terrain on Windsor street was observed perpendicular to the flight lines which may have had an affect of allowing random errors in the GPS and IMU units, which operate at lower frequencies than the laser scanner and laser ranger, to exhibit themselves systematically allowing the resulting observed error to appear better conditioned than the reality. If several data sets were obtained throughout the study area it is likely that the systematic portion of the error would change between locations due to differing random errors. Another benefit of observing sloped areas perpendicular to flight path is that there is a possibility the direction of the laser pulse will have a perpendicular intersection with the terrain. Schaer et al. [2007] describe this effect in a discussion of errors caused by beam incidence level and claim that this will reduce the associated error as it mimics nadir observations. Considering these factors it appears that the algorithm is working as is expected and is a better approximation of the inherent error in the observations than just the hardware components alone.

In summary the conclusions are as follows:

- The developed error prediction routine that includes slope better predicts the total error in LiDAR observations over sloped surfaces. This was demonstrated by the fact that existing error predictions which include only errors in the hardware components of the system generate extremely optimistic predictions on sloped terrain.
- Existing error models provided in the LiDAR industry are insufficient in describing the true nature of error in the observations. They do not take into account several sources of error including terrain effects.

- Slope based errors in LiDAR observations represent a source of error which will significantly and adversely affect any derived quantities from the DSM, such as slope. Care should be exercised when calculating these quantities and the vertical error in the LiDAR observations should be appropriately propagated through to these final quantities to ensure their quality is sufficient for the intended application.

6.1 Recommendations and Future Work

The total propagated error from all factors within LiDAR observations can be thought of as a puzzle that once solved will provide a clear picture of the magnitude and nature of error in a LiDAR system. The solution will give end-users and service providers a clear understanding of the system capabilities in a variety of settings allowing for a greater understanding of the uses and limitations of the DSMs they produce. Currently this thesis is attempting to describe the picture produced by the puzzle with only two pieces, represented by hardware errors and errors based on slope. Imagine attempting to put together a difficult puzzle with only two pieces, and then attempting to describe the picture the puzzle represents. One of the most important factors required for future research is obtaining the remaining puzzle pieces. For the total propagated solution of LiDAR observations this would include atmospheric effects, laser beam incidence angle, range based intensity biases, interaction of the pulse with terrain features all of which still require extensive modelling and verification. Once all the factors have been modelled, the appropriate combination of their effects will result in appropriate predictions of the total error in the system.

Hardware errors were modelled first because of the ease of propagating them through the General Law of Propagation of Variances and the existing available research on the Global Positioning System and Inertial Motion Unit errors. This enabled solutions to their relatively easy predictions. This thesis focussed on slope based errors because it was considered the second largest source of error in final observations and could be facilitated by initial predictions from hardware components. However, it is possible that an analysis of this type is premature as there are several other known, yet not modelled error sources that could be included into the initial predictions. On highly sloped environments these errors will propagate much more severely into final error estimates. The consistent underestimation of the results in the ski trails is likely a result of such factors, the road was not affected as much because of the consistent nature and orientation of the road which reduced the observed errors. Inclusion of the other error types would allow these factors to be appropriately taken into account and result in improved error estimations.

In addition to the total picture of the error propagation routine being available, appropriate test methods also need to be implemented. The testing of the ski trail data contained obvious biases that were due to the necessary interpolation of the terrain surface itself as well as the interpolation of the error predictions. It is an unlikely that the assumption that these surfaces take on linear connections between points is valid. However, in general, the assumption of linearity becomes increasingly valid with denser point spacing. Therefore, validation data should be collected with this in mind. If terrain is highly variable, heterogeneous, or with continuously changing slopes than point

spacing on validation data should be kept extremely low. If the terrain is consistent than the point spacing on validation data can be kept higher.

The validation on the road surface is ideal for two reasons; the first is that the assumption of linear connections on the terrain can be held true. Secondly, an entire reference surface can be created allowing the LiDAR points to form the sample set for statistical testing as opposed to the validation points. This is much more preferable as it drastically increases the size of the sample data allowing for more trustworthy statistics. Of course, when testing errors on consistent surfaces such as roads an adequate picture of the errors on less consistent variable surfaces is not given. This was seen in Chapter 5, where errors on the sloped terrain were much higher than those viewed on the road surface.

Just as error estimates derived from flat parking lot data cannot accurately extrapolated to sloped surfaces, error estimates derived from consistent sloped surfaces cannot be extrapolated to more variable surfaces. This indicates that validation data should ideally be taken from several surface and ground cover types within the survey area to gain a true understanding of the distribution of error within the dataset.

Hodgson et al. [2005] also developed an interesting tactic of using RTK surveying techniques to ‘stake out’ the position of LiDAR points within the dataset. This involves locating the existing LiDAR points on the ground using the real time corrections and then recording the elevation obtained from the GPS observations at the LiDAR point. Following this, the elevation obtained from the LiDAR point can be compared with that obtained by the GPS observations. This eliminates the need for interpolation as validation data is collected directly on the LiDAR point. The caveat of such a process is that

validation data cannot be collected until the LiDAR survey has been completely processed meaning a delay of at least several days to several weeks. Therefore, the survey site must be accessible after this period of time and there cannot be any significant changes to the structure of the terrain in this period. If these do not cause a substantial problems this seems to be an ideal method for collecting reference data, although it will not contain the available sample sizes available from creating reference surfaces as was done with the Windsor Street dataset. Nevertheless for areas where a reference surface is not possible this is an excellent alternative and would be the suggested method.

In summary the final recommendations for future work are as follows:

- The total propagated error in LiDAR observations has not yet been achieved. The error prediction routine will perform better once other errors sources have been properly documented and included. These sources include atmospheric affects, range based intensity biases and the interaction of the pulse with surface features such as vegetation.
- Existing quality control procedures are not ideal for testing errors on sloped surfaces. These procedures generally involve comparing validation points with LiDAR surfaces. On sloped surfaces this will incur significant errors from the necessity to interpolate the surface linearly between LiDAR points. Existing error analysis techniques need to be updated to consider utilizing reference validation surfaces or locating LiDAR points with RTK technology and observing validation data directly on the same horizontal location.

REFERENCES

Baltsavias, E.P., (1999). “Airborne laser scanning: existing systems and firms and other resources.” *ISPRS Journal of Photogrammetry and Remote Sensing* 54 (2-3), 164-198

Baltsavias, E. P. (1999). “Comparison between photogrammetry and laser scanning." *ISPRS Journal of Photogrammetry and Remote Sensing,* 54(2-3), 83-94.

Cramer, M., 1999. “Direct geocoding – is aerial triangulation obsolete?”, in Fritsch/Spiller (eds.): *Photogrammetric Week 1999*, Wichmann Verlag, Heidelberg, Germany, 59-70.

Cruickshank, J.M. and Harney, R.C. (1986). "Laser radar technology and applications” 3-5 June, 1986, Québec City, Canada. 219.

Filin, S., 2003. “Recovery of Systematic Biases in Laser Altimetry Data Using Natural Surfaces.” *Photogrammetric Engineering and Remote Sensing*. Vol. 69, No.11, pp. 1235-1242.

Filin, S. and Vosselman, G., 2003. “Adjustment of Airborne Laser Altimetry Strips.” *XXth ISPRS Congress*, 12-23 July 2004 Istanbul, Turkey, Commission 3, Working Group 3
FDA 2006

Fisher, P. F. (1991). "First experiments in viewshed uncertainty. The accuracy of the viewshed area." *Photogrammetric Engineering and Remote Sensing,* 57(10), 1321-1327.

Florinsky, I. V. (1998). "Accuracy of local topographic variables derived from digital elevation models." *International Journal of Geographical Information Science,* 12(1), 47-61.

Flood, M. (2004). “Vertical Accuracy Reporting for LiDAR data.” American Soceity for Photogrammetry and Remote Sensing. ASPR Lidar Committee. [On-line] January 5 2009 http://www.asprs.org/society/committees/lidar/Downloads/Vertical_Accuracy_Reporting_for_Lidar_Data.pdf

Glennie, C., 2007. “Rigorous 3D error analysis of kinematic scanning LiDAR systems.” *Journal of Applied Geodesy* 1, 147-157.

Goulden, T. and Hopkinson, C. (2008). “The forward propagation of integrated system component errors within airborne LiDAR data.” Under review with *Photogrammetric Engineering and Remote Sensing*.

Grejner-Brzezinska D. A. and Wang J. (1998). “Gravity Modeling for High-Accuracy GPS/INS Integration.” *Navigation*, Vol. 45, No. 3, pp.209-220, also in Proceedings of ION Annual Meeting, Denver, CO, June 1-3.

Hare, R., 2001. Error Budget Analysis for US Naval Oceanographic Office (NAVOCEANO) Hydrographic Survey Systems, a Final Report for the Naval Oceanographic Office, Prepared for the University of Southern Mississippi for Naval Oceanographic Office.

Heidenhain, 2005. Angle Encoders, [On-line] 24 February, 2008. http://www.heidenhain.com/

Heuvelink, G. B. M., Burrough, P. A., and Stein, A. (1989). "Propagation of errors in spatial modelling with GIS." *International Journal of Geographical Information Systems,* 3(4), 303-322.

Hodgson, M. E., Jensen, J. R., Schmidt, L., Schill, S., and Davis, B. (2003). "An evaluation of LIDAR- and IFSAR-derived digital elevation models in leaf-on conditions with USGS Level 1 and Level 2 DSMs." *Remote Sensing of the Environment,* 84 295-308.

Hodgson, M. E., and Bresnahan, P. (2004). "Accuracy of Airborne Lidar-Derived Elevation: Empirical Assessment and Error Budget." *Photogrammetric Engineering and Remote Sensing,* 70(3), 331-339.

Hodgson, M. E., Jensen, J., Raber, G., Tullis, J., Davis, B. A., Thompson, G., and Schuckman, K. (2005). "An evaluation of lidar-derived elevation and terrain slope in leaf-off conditions." *Photogrammetric Engineering and Remote Sensing,* 71, 817-823.

Hopkinson, C., and Demuth, M. N. (2006). "Using airborne lidar to assess the influence of glacier downwasting on water resources in the Canadian Rocky Mountains." *Canadian Journal of Remote Sensing,* (32) 212-222.

Huising, E. J., and Gomes Pereira, L. M. (1998). "Errors and accuracy estimates of laser data acquired by various laser scanning systems for topographic applications." *ISPRS Journal of Photogrammetry and Remote Sensing,* (53) 245-261.

Krabill, W. B., Abdalati, W., Frederick, E. B., Manizade, S. S., Martin, C. F., Sonntag, J. G., Swift, R. N., Thomas, R. H., and Yungel, J. G. (2002). "Aircraft laser altimetry measurement of elevation changes of the Greenland ice sheet; technique and accuracy assessment; Laser altimetry." *Journal of Geodynamics,* 34(3-4), 357-376.

Lee, H.K., 2004. Integration of GPS/Pseudolite/INS for high precision kinematic positioning and navigation, Sydney, NSW : School of Geomatic Engineering, University of New South Wales

Lemmens, M. J. P. M. (1997). "Accurate height information from airborne laser-altimetry." *Part 1 (of 4),* IEEE, Piscataway, NJ, USA, Singapore, Singapore, 423-426.

Maling, D. H. (1989). Measurements from Maps. Pergamon Press.

Marini, J.W. and Murray, C.W. Jr. (1973). "Correction of Laser Range Tracking Data for Atmospheric Refraction at Elevations Above 10 Degrees." Technical Report, Goddard Space Flight Center. NASA-TM-X-70555

Maune, D. F. (2007). Digital elevation model technologies and applications : the DSM users manual. American Society for Photogrammetry and Remote Sensing.

Mikhail, E. M. (1981). Analysis and adjustment of survey measurements. Van Nostrand Reinhold Co., New York.

Morin, K., 2002. *Calibration of Airborne Laser Scanners*, Masters dissertation, Faculty of Graduate Studies for the Department of Geomatics Engineering University of Calgary. UCGE Report No. 20179

Mostafa, M.M.R., 2001. "Airborne Image Georeferencing by GPS_Aided Inertial Systems: Concepts and Performance Analysis." *22nd Asian Conference on Remote Sensing*, 5-9 November 2001, Sinapore.

Mostafa, M., Hutton, J., Reid, B., 2001. "GPS/IMU Products – The Applanix Approach." *Photogrammetric Week 01, Proceedings of the 48th Photogrammetric Week*, Stuttgart, Institue for Photgrammetry, University of Stuttgart

Müller, R., Lehner, M., Müller, R.,Reinartz, P.,Schroeder, M., Vollmer, B., 2002. A Program for Direct Georeferencing of Airborne and Spaceborne Line Scanner Images, *International Soceity of Photogrammetry and Remote Sensing*, Commission I, WG I/5

Oksanen, J., and Sarjakoski, T. (2006). "Uncovering the statistical and spatial characteristics of fine toposcale DSM error." *International Journal of Geographical Information Science,* 20(4), 345-369.

Oksanen, J., and Sarjakoski, T. (2005). "Error propagation of DSM-based surface derivatives." *Computers and Geosciences,* 31(8), 1015-1027.

Optech© Inc., 2004. ALTM Specifications. Optech© Inc.

Reinshaw, 2006. RESR Angular Encoder System, Data Sheet L-9517-9128-04-A, [On-line] 24 February 2008. http://www.renishaw.com/UserFiles/acrobat/UKEnglish/L-9517-9128.pdf

Reutebuch, S. E., McGaughey, R. J., Andersen, H., and Carson, W. W. (2003). "Accuracy of a high-resolution lidar terrain model under a conifer forest canopy." *Canadian Journal of Remote Sensing,* 29 527-535.

Schaer, P., Skaloud, J., Landtwing, S., Legat, K., 2007. "Accuracy Estimation for laser Point Cloud Including Accuracy Estimation." Padua, Italy.

Schenk, T., 2001. Modeling and Analyzing Systematic Errors of Airborne Laser Scanners, *Technical Notes in Photogrammetry* No. 19, Department of Civil and Environmental Engineering and Geodetic Science, The Ohio State University, Columbus, Ohio, 40 p.

Schwarz, K.P., 1983. "Inertial Surveying and Geodesy." *Reviews of Geophysics and Space Physics*, Vol. 21, No. 4, pp. 878-890

Schwarz, K. P. and El-Sheimy, N., 2004. Mobile Mapping Systems - State of the Art and Future Trends, Invited Paper, TS SS 3 – Mobile Multi-sensor Systems, [On-line] February 24, 2008. http://www.isprs.org/istanbul2004/comm5/papers/652.pdf

Skaloud, J. and Lichti, D., 2006. "Rigorous Approach to Bore-sight Self-calibration in Airborne Laser Scanning." *ISPRS Journal of Photogrammetry and Remote Sensing* 61 (1), 47-59

Su, J., and Bork, E. (2006). "Influence of vegetation, slope, and lidar sampling angle on DSM accuracy." *Photogrammetric Engineering and Remote Sensing,* 72(1), 1265-1274.

TMS International Ltd., 2005. The Global Market for Airborne Lidar Systems and services. TMS International Ltd. Houston Texas. 158 pp.

Trujillo-Ortiz, A., R. Hernandez-Walls, K. Barba-Rojo and A. Castro-Perez. (2007). AnDartest : Anderson-Darling test for assessing normality of a sample data. A MATLAB file.[On-line] February 24 2008. http://www.mathworks.com/matlabcentral/fileexchange/loadFile.do?objectId=14807

Vaughn, C., Bufton, J., Krabill, W. and Rabine, D., 1996. "Georeferencing of airborne laser altimeter measurements." *International Journal of Remote Sensing*, 17(11), pp. 2185-2200.

Vosselman, G., and Maas, H. G., 2001. "Adjustment and filtering of raw laser altimetry measurements." *Proceedings of OEEPE Workshop on Airborne Laser Scanning and Interferometric SAR for Detailed Digital Elevation Models*, 01-03 March, Stockholm, Sweden, OEEPE Publication No. 40, pp.62-72

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

Wehr, A., Lohr, U., 1999. "Airborne Laser Scanning – an introduction and overview." *ISPRS Journal of Photogrammetry and Remote Sensing* 54, 68-82.

Wolf, P. R. (1997). Adjustment computations : statistics and least squares in surveying and GIS. John Wiley & Sons, New York.

Appendix A

Full Derivation of Tangent Calculation

Parametric Equation of a non-rotated ellipse at origin

$$x = a\cos(t)$$
$$y = b\sin(t) \tag{1}$$

Parametric Equation of a line beginning at arbitrary point p

$$x = p_x + m_x s$$
$$y = p_y + m_y s \tag{2}$$

Line must be tangent to the ellipse surface. Slope of this tangent line found by taking the first derivative of ellipse equation.

$$\frac{\partial x}{\partial t} = -a\sin(t)$$
$$\frac{\partial y}{\partial t} = b\cos(t) \tag{3}$$

The slope values can be substituted in fro m in (2) and x and y equations can be set equal to each other leaving only *t* and *s* to be solved, appearing as follows;

$$a\cos(t) = p_x - a\sin(t)s \tag{4}$$
$$b\sin(t) = p_y + b\cos(t)s \tag{5}$$

To integrate into a single expression (4) can be re-arranged and solved for *s* and substituted into (5) as follows

$$b\sin(t) = p_y + \frac{b\cos(t)\left(a\cos(t) - p_x\right)}{-\sin(t)} \tag{6}$$

From this, the equation in multiplied by the denominator of the second term on the right hand side to yield

$$-ab\sin^2(t) = -p_y a\sin(t) + ab\cos^2(t) - b\cos(t)\,p_x \tag{7}$$

This equation can be simplified to

$$-ab = -p_y a\sin(t) - b\cos(t)\,p_x \tag{8}$$

Isolating the *sin* terms to the right hand side and the *cos* terms to the left hand side and squaring both sides yields

$$a^2b^2 + 2p_x ab^2 \cos(t)\, p_x + p_x^2 b^2 \cos^2(t) = p_y^2 a^2 \sin^2(t) \tag{9}$$

And the right hand side can be modified to

$$RHS = p_y^2 a^2 - p_y^2 a^2 \cos^2(t) \tag{10}$$

Collecting like terms yields

$$a^2b^2 - p_y^2 a^2 - 2ab^2 p_x \cos(t) + \left(b^2 p_x^2 + a^2 p_y^2\right)\cos^2(t) = 0 \tag{11}$$

which can be thought of as a pseudo-quadratic, utilizing the quadratic equation with

$$l = b^2 p_x^2 + a^2 p_y^2 \tag{12}$$
$$m = 2ab^2 p_x \tag{13}$$
$$n = a^2b^2 - p_y^2 a^2 \tag{14}$$

results in

$$\cos(t) = \frac{-m \pm \sqrt{m^2 - 4ln}}{2l} \tag{15}$$

Yielding the two required solutions.

CURRICULUM VITAE

Candidate's full name: Tristan James Goulden

Universities Attended:
University of New Brunswick, BSc. Eng. GGE, 2005

Conference Presentations:

Goulden, T., C. Hopkinson (2008). The Forward Propagation of Integrated System Component Errors Within Airborne LiDAR Data. Presented at American Society of Photogrammetry and Remote Sensing annual conference, Portland, Oregon, April 28th to May 2nd, 2008.

Goulden, T., C. Hopkinson, P.Dare and M. Demuth (2008). Quantifying Uncertainty Resulting from Terrain Effects in Airborne LiDAR Observations. Presented at American Society of Photogrammetry and Remote Sensing annual conference, Portland, Oregon, April 28th to May 2nd, 2008.

Goulden, T., C. Hopkinson, P.Dare and M. Demuth (2008). Quantifying Uncertainty Resulting from Terrain Effects in Airborne LiDAR Observations. Presented at Canadian Geophysical Union annual conference, Banff, Alberta, May 11th to May 14nd, 2008.