APPLICATION OF LIDAR DEMS TO THE MODELLING OF SURFACE DRAINAGE PATTERNS IN HUMAN MODIFIED LANDSCAPES

A Thesis
Presented to
The Faculty of Graduate Studies
of
The University of Guelph

by
KIMBERLY DHUN

In partial fulfillment of requirements
for the degree of
Master of Science
September, 2011

© Kimberly Dhun, 2011
ABSTRACT

APPLICATION OF LIDAR DEMS TO THE MODELLING OF SURFACE DRAINAGE PATTERNS IN HUMAN MODIFIED LANDSCAPES

Kimberly Dhun
University of Guelph, 2011
Advisor: Professor John Lindsay

Anthropogenic infrastructure such as roads, ditches and culverts have strong impacts on hydrological processes, particularly surface drainage patterns. Despite this, these structures are often not present in the digital elevation models (DEMs) used to provide surface drainage data to hydrological models, owing to the coarse spatial resolution of many available DEMs. Modelling drainage patterns in human-modified landscapes requires very accurate, high-resolution DEM data to capture these features. Light Detection And Ranging (LiDAR) is a remote sensing technique that is used for producing DEMs with fine resolutions that can represent anthropogenic landscapes features such as human modifications on the landscape such as roadside ditches. In these data, roads act as a barrier to flow and are treated as dams, where on the ground culverts and bridges exist. While possible to locate and manually enforce flow across these roads, there is currently no automated technique to identify these locations and perform flow enforcement. This research improves the modelling of surface drainage pathways in rural anthropogenic altered landscapes by utilizing a novel algorithm that identifies ditches and culverts in LiDAR DEMs and enforces flow through these features by way of breaching. This breaching algorithm was tested on LiDAR datasets for two rural test sites in Southern Ontario. These analyses showed that the technique is an effective tool for efficiently incorporating ditches and culverts into the hydrological analysis of a landscape that has both a gradient associated with it, as well as a lack of densely forested areas. The algorithm produced more accurate representations of both overland flow when compared to outputs that excluded these anthropogenic features all together.
ACKNOWLEDGEMENTS

This research project would not have been possible without the support of many people. First and foremost I wish to express my deepest gratitude to my advisor, Dr. John Lindsay who supported me throughout my thesis. Without your encouragement and patience during this time, this thesis would not have been completed. It has been a pleasure working with you papa bear.

Thanks to Rashaad Bahamjee and Vanessa Stretch for identifying culverts in the field, creating my graphics, going on bike rides and last but not least always being there when I was stuck. Without your help I would probably still be verifying my reference data, trying to draw a laser beam, out of shape and hiding Jasleen in my back pocket. Never forget to “pan dat baby!” while ‘bringing it’.

I would like to thank Jared Cunningham and Steve Elgie for helping me with my data. Jared thanks for your input regarding my data analysis procedures and your patience with my ArcGIS questions. Steve thank you for interpolating my study sites with humor. Both of you have saved my time and sanity, both of which were exponentially dwindling during the last couple months.

Special thanks to Kerry Scutten. I would like to express my gratitude for editing my entire thesis. TWICE!! You are one of a kind. Thank you for your patience during that painful yet enlightening process. “I love Kerry! I love Kerry! I love Kerry!”

Thanks to Colleen Fuss for being intense. My thesis would not look as good as it does. I take back my jokes about the floor plan.
Finally I wish to express my love and gratitude to my family, particularly my daughter Jasleen Dhun. For a five-year-old you are such an understanding, loving daughter. It was quite difficult at times but we have reached the end of this journey. Thank-you for singing “never give up” and for being so patient when mummy studied. I love you very much and dedicate all my hard work to you, my precious lady.
# TABLE OF CONTENTS

Acknowledgements........................................................................................................... i

Table of Contents ................................................................................................................... iii

List of Tables ........................................................................................................................... vi

List of Figures .......................................................................................................................... vii

1.0 Introduction....................................................................................................................... 1

   1.1 Background...................................................................................................................... 1

   1.2 Research Objectives ..................................................................................................... 4

   1.3 Thesis Outline ............................................................................................................... 5

2.0 Literature Review............................................................................................................... 6

   2.1 Digital Elevation Models – Data acquisition and structures ........................................ 6

   2.2 Factors Determining the Quality of a DEM ................................................................. 9

   2.3 Light Detection and Ranging ........................................................................................ 13

   2.4 Applying DEMs for Hydrological Modelling ............................................................... 16

   2.5 Flow-Routing Algorithms ........................................................................................... 20

      2.5.1 Single Flow Direction Algorithm (SFD) ............................................................... 20

      2.5.2 Multiple Flow Direction Algorithms (MFD) ....................................................... 21

   2.6 Comparisons and Limitations of algorithms .............................................................. 23

   2.7 Hydrologic conditioning DEMs ................................................................................... 25
2.7.1 Drainage/ Flow Enforcement ......................................................... 26

2.8 Modelling flow in human-modified landscapes .................................. 30

2.8.1 Incorporating ditches and culverts through auxiliary data .................. 32

2.9 LiDAR DEMs .................................................................................... 35

3.0 Breaching Algorithm ........................................................................ 38

3.1 Introduction ...................................................................................... 38

3.2 Background ...................................................................................... 39

3.3 Breaching Algorithm Rationale .......................................................... 46

3.3.1 Algorithm Description .................................................................. 47

3.4.0 Methods ....................................................................................... 51

3.4.1 Study Areas .................................................................................. 51

3.4.2 Data collection ............................................................................. 54

3.4.3 Data Processing ........................................................................... 61

3.4.4 Accuracy Assessment ................................................................... 64

3.4.45 Ditch Gradient Calculation ......................................................... 65

3.5 Results .............................................................................................. 66

3.6 Discussion ......................................................................................... 74

3.7 Conclusion ......................................................................................... 78

4.0 Conclusion ........................................................................................ 80
LIST OF TABLES

Table 2.1. Factors Determining DEM Quality ................................................................. 9

Table 2.2. Comparison of LiDAR and Photogrammetric DEMs (Mouton, 2005; Murphy et al, 2008; Barber and Shortridge, 2005) ................................................................. 16

Table 2.3. Primary attributes derived from a DEM (Wilson and Gallant, 2000) .......... 18

Table 2.4. Secondary attributes derived from a DEM (Wilson and Gallant, 2000) .... 19

Table 3.1. User defined parameters used by the novel-breaching algorithm ............ 50

Table 3.2. LiDAR acquisition parameters for PEC Site ..................................................... 57

Table 3.3. Flight mission dates for the PEC site (Aero-Photo, 2009) .............................. 58

Table 3.4. Accuracy assessment for LiDAR point cloud at the PEC site ..................... 59

Table 3.5. Breakdown of depressions for each site ......................................................... 69

Table 3.6. Accuracy assessments for both the PEC and Rondeau Basin sites ............ 70

Table 3.7. Statistics associated with EOC frequency and road slope ......................... 72
LIST OF FIGURES

Figure 2.1. Typical LiDAR system ................................................................. 14

Figure 2.2. Flow allocation in D-infinity following Tarboton, (1997) ....................... 22

Figure 2.3. Runoff interaction with roads (Jones et al., 2000) .................................. 32

Figure 2.4 REA enforces flow on elevated roads. Runoff on the upslope side of the road is re-routed whereas runoff on the downslope side not affected. (Duke et al., 2003)) ............................. 33

Figure 2.5: REA flow enforcement along roadside ditches. Runoff is re-routed on both the upslope and downslope sides of the road (Duke et al., 2003) ......................................................... 33

Figure 2.6 Flat cross sectional road profile in which runoff is not considerably affected (Duke et al., 2003) ......................................................................................................................... 34

Figure 2.7 Flow patterns in an irrigated landscape: (a) cross-flow scenario with an irrigation canal as a solid line (grid representation shown as gray squares) and a stream course as a dotted line (white squares); (b) flow direction representation of cross-flow pattern (Duke et al., 2006) ..................................................................................................................... 35

Figure 3.1. Flowchart of breaching algorithm .......................................................... 49

Figure 3.2 Rondeau Basin Study Site ..................................................................... 52

Figure 3.3 Prince Edward County Study Site ......................................................... 53

Figure 3.4 a1, b1) The original DEM prior to pre-processing; a2, b2) the filled DEM; a3, b3) The breached DEM ................................................................................................................................. 66

Figure 3.5. Effects of flow- accumulations on a1, b1) The orginal DEM; a2, b2) The filled DEM; a3, b3)The breached DEM ......................................................................................................................... 68
Figure 3.6 PEC road slope relative to EOC’s ................................................................. 71

Figure 3.7 Rondeau road slope relative to EOC’s ......................................................... 71

Figure 3.8 Ditch enforcement along roads in both Rondeau Bay area and PEC sites .... 73
1.0 INTRODUCTION

1.1 Background

Many human activities can have a substantial impact on the hydrological functioning of drainage basins. For example, replacing the natural vegetation within a landscape with impermeable surfaces, such as asphalt and concrete, can reduce surface and soil moisture storage, increase runoff, and reduce percolation to ground water (Marsalek, 2007). Surface drainage patterns are also often modified through the introduction of anthropogenic conduits of flow, such as ditches and culverts. Ditches are used to channel and redirect surface runoff and culverts are used to channel water from one side of an embankment to the other. Both of these features are commonly associated with road construction. Wemple et al. (1996) showed that with the presence of ditches and culverts, drainage density increased from 21% to 50%. This increase in the density of drainage channels has been found to lead to increased peak flow magnitude and shortened time-to-peak (Liu and Wang, 2008). Despite the significant role ditches and culverts have on surface flow, these features are seldom accounted for when modelling drainage patterns in human-modified landscapes. This common situation is often the result of the inability of source data to represent ditches and culverts, owing to their coarse resolution and the inability of existing techniques for simulating surface drainage to incorporate these features (Duke et al., 2006).

Drainage patterns are usually mapped using flow-routing algorithms applied to digital elevation models (DEMs), which are digital representations of the Earth’s surface. Flow-routing algorithms assume that topography is a dominant control on the redistribution of surface runoff (O’ Callaghan and Mark, 1984; Band and Wood 1988;
To accurately model surface flow, the DEM must capture all features that influence drainage patterns (Jenson and Domingue, 1988).

Light Detection And Ranging (LiDAR) is an advanced topographic mapping technology that offers many advantages over traditional data acquisition methods. The fine spatial resolution and low per-point data collection costs associated with these datasets are superior for modelling surface drainage patterns when compared to traditional methods. LiDAR DEMs are capable of representing linear anthropogenic features such as ditches and culverts, and as a result, they form more accurate representations of the Earth’s surface when compared to traditional data collection techniques. Many authors have recognized the potential for LiDAR data in modelling accurate drainage patterns (Charloton et al., 2003; French, 2003; Shortbridge and Barber, 2005; Poppenga et al., 2009). However, there are many challenges associated with using LiDAR data for hydrological applications, especially when considering preprocessing methods.

Most DEMs contain depressions; these depressions can be a single cell or set of neighbouring cells that are not linked to an outlet (Florinsky, 2002; Tianqi et al., 2003). Most depressions present in the DEM surface are spurious and occur for various reasons, such as a rounding of elevation values in a regular grid (Aguilar et al., 2005), during the interpolation process (Chaplot et al., 2006; Lindsay and Creed, 2005), errors in the source data (Tribe, 1992; Reiger, 1998) or because of the limited horizontal and vertical resolution of the elevation data used to model the complex terrain (Wolock and Price, 1994; Thompson et al., 2001; Lindsay and Creed, 2005). Depressions interrupt overland flowpaths and alter flow direction. Due to the negative effects depressions have on
drainage pattern simulation, analysts usually identify and remove all depressions before trying to extract relevant hydrological terrain attributes from the DEM (Jenson and Dominque, 1988).

Many algorithms exist that ensure that each grid cell in the model has at least one downslope neighbour (Band, 1986b; Carrara, 1986; Jenson and Domingue, 1988). While these algorithms remove depressions, they do not differentiate between an artifact and an actual depression. For example, in human modified landscapes, ditches and road underpasses, such as culverts and bridges, significantly influence overland flow. Ditches are local linear anthropogenic features that channelize flow and culverts are linear structures, usually installed under roadways to facilitate overland flow across embankments. While in reality water may be diverted under a road bank through a culvert or flow under a bridge, these road underpasses are not captured in the LiDAR data, resulting in artificial damming. This is due to the fact that DEMs cannot represent the Earth’s surface three dimensionally (3D). The main objective of depression removal algorithms is to create a hydrologically connected surface, meaning that the algorithm alters the DEM in such a way that all cells are connected to an outlet by simply removing depressions. However, while the algorithms remove depressions that cause inaccurate drainage pattern simulations, they also remove anthropogenic linear features that greatly affect landscape drainage patterns, which are already embedded in LiDAR DEMs.

Murphy et al. (2008) stated that LiDAR DEMs capture anthropogenic features that affect overland flow by blocking and channeling drainage patterns. MacMillan (2003) illustrated that LiDAR DEMs show erroneous drainage pattern simulations due to their inability to represent the Earth’s surface. There has been limited research addressing the
potential of LiDAR DEMs to model surface drainage patterns in human-modified landscapes. Barber and Shortridge (2005) showed that flow related derivatives were incorrect if bridges and graded roadbeds were present in the DEM. They noted that without identifying road underpasses, LiDAR DEMs were “hydrologically challenged” (Barber and Shortridge, 2005). Karlin et al. (2010) referred to LiDAR DEMs as “blessings in disguise” because of the ability of the DEMs to capture detailed information of the terrain coupled with the curse of having to manually filter anthropogenic features such as road underpasses. Poppenga et al. (2009) emphasized the need for the development of a novel method for identifying these anthropogenic obstructions. They stated that if culverts and ditches were not identified, runoff would flow over the obstruction in the wrong location or drain the flow in the opposite direction (Poppenga et al., 2009).

LiDAR DEMs have the ability to substantially improve drainage pattern modelling, especially in human-modified landscapes where LiDAR DEMs are capable of representing linear anthropogenic features. As LiDAR data becomes more accessible, the use of this fine-resolution data will play an important role in accurately modelling hydrological processes. Traditional preprocessing techniques do not work as desired on this fine-resolution data, so there is a strong need to develop LiDAR specific preprocessing and flow-enforcement techniques.

1.2 Research Objectives

The overall aim of this research is to improve the modelling of surface drainage pathways in landscapes that have been heavily altered by anthropogenic infrastructure by
developing and testing novel techniques that use the topographic information contained in LiDAR DEMs. To achieve this goal, the following specific objectives have been identified:

1. To create an algorithm for enforcing DEM modelled flow patterns along roadside ditches;

2. To develop a predictive model for identifying potential culvert/bridge locations in a LiDAR DEM as well as a technique for breaching the artificial dams created by the embankments at these sites;

3. To evaluate the ability of the model to accurately predict road underpasses and to correctly enforce flow along road embankments.

1.3 Thesis Outline

Chapter 2 summarizes the literature associated with techniques used for modelling drainage patterns in human-modified landscape. Chapter 3 explains the application of a novel automatic technique for extracting ditches and road underpasses on two study sites and summarizes the methods and data sources used for this novel technique. Chapter 3 will then quantify the performance of this automated technique when applied to various topographic conditions. Chapter 3 will then set out the limitations of the automated technique over a DEM and the type of landscapes that the automated technique is best suited for. The final chapter (Chapter 4) will then summarize the major findings and offer conclusions based on the methodology and results of this research.
2.0 LITERATURE REVIEW

2.1 Digital Elevation Models – Data acquisition and structures

A DEM can be defined as a digital representation of the Earth’s surface (Hengl and Reuter, 2008). The three main sources of DEM data are ground survey, topographic maps, and remote sensing (Wilson and Gallant, 2000). Ground surveyed data are acquired through the manual collection of surface-specific point elevations. Until recently, these datasets were mainly acquired using a theodolite, which is a surveying instrument used for measuring angles in the horizontal and vertical planes (Wainwright and Mulligan, 2004). However, advances in technology have created alternative sources of surveying instruments such as Global Positioning Systems (GPS), Electronic Distance Measuring instruments (EDMs) and digital theodolites of surveying data. These technologies have increased the accuracy of the elevation data while decreasing preprocessing time. Ground survey data are more suited to smaller catchments and are not often used for larger areas because of the extensive costs and labour associated with data collection (Gopi et al., 2007).

Elevation data are usually acquired from contour lines when an area is inaccessible or topographic maps are the only source of elevation data (Wilson and Gallant, 2000). DEMs are created from digitizing elevation data points from topographic maps (Carrara et al, 1997). The drawback of extracting contour data from topographic data is that undersampling of areas between contour lines can occur, resulting in a generalized representation of the terrain (Wilson and Gallant, 2000). However, despite this limitation, topographic maps have been a major source of elevation data for DEM generation (Moore et al., 1991; Wilson and Gallant, 2000).
Currently, various remote-sensing technologies have become the preferred method for collecting elevation data (Weschler, 2006; Campbell, 2007). In addition to the substantial decrease in data costs and increase in computing power, remote sensing can cover larger areas with less effort (Campbell 2007). Remotely sensed data are collected by airborne and satellite sensors that can derive two types of data: aerial photographs and radargrammetric images. Aerial photos are georeferenced fine-resolution photographs taken from an airborne optical sensor (Avery and Berlin, 1992). The sensor captures overlapping photographs of an area, which are then viewed through a stereoscope. The stereoscope allows the interpreter to view the aerial photos in 3D, thus extracting elevation data from a 3D point of view. Combined with ground control points (GCPs), photogrammetry has been a major contributor to DEM data (Avery and Berlin, 1992; Campbell, 2007). Radargrammetry is a technique used on Synthetic Aperture Radar (SAR) stereo data for extracting elevation data. Radargrammetry can be thought of as an extension to photogrammetry. Radar sensors offer some advantages over optical sensors. For example, radar sensors have their own source of energy and thus can be operated at night, they are not very sensitive to rain, and images from radar sensors provide fine resolution elevation data when compared with optical sensors (Madsen et al., 1993).

DEMs can be constructed by interpolating point-based data into any three data structures: regular grid, triangulated irregular networks (TIN), and contour lines. Regular grid, sometimes referred to as rasters, can be described as a tessellation using square tiles with point elevations at the center of the square (Lyon, 2003; Moore et al., 1991; Wilson and Gallant, 2000). These square grids are arranged in rows and columns, each grid representing elevation at the grid cell center. This data structure is most widely used.
because the coordinates of each data point are stored implicitly relative to the grid edge and because it is amenable to analysis that uses efficient image-processing type algorithms (Moore et al., 1991; Grayson and Blöschl, 2001). Triangulated Irregular Networks (TIN) use a set of continuous, non-overlapping, irregular-shaped triangles to represent the surface. This network of triangles adapts to the terrain variations because TIN cells vary in size according to topographic complexity, which in turn offers a convenient way of incorporating drainage lines (Wilson and Gallant, 2000). Finally, the contour data structure incorporates the ‘stream tube’ concept devised by Brakensiek and Onstad (1968). This structure segments the land into irregular shaped quadrilaterals based on an overlapping network of contour lines and flow lines (Wilson and Gallant, 2000). Each quadrilateral consists of two sides that are segments of contour lines and two sides that are flow lines, with the contours crossing at right angles. The term DEM is nearly synonymous with the regular-grid data format because it is so common.

Although the regular-grid data structure is by far the most commonly used format, many researchers argue that this structure is not well suited as a representation of terrain (Grayson and Blöschl, 2001; Maidment and Djokic 2000; Smith et al., 2007). Even though the scale of topographic variation is not constant in a landscape, a regular grid DEM represents topography using an invariant grid resolution. Thus, in some areas the grid resolution may be too coarse to properly represent the underlying topographic variation while in other areas the resolution may be overly fine, resulting in data redundancy (Grayson and Blöschl, 2001; Maidment and Djokic, 2000; Raafflaub and Collins, 2006).
Historically, digital data storage was a concern and the potential data redundancy of a fine-resolution DEM provided a major challenge for researchers. As storage issues have become less significant, there has been a persistent trend toward ever-finer resolutions in DEM data, in an attempt to represent the necessary detail of the underlying topography for particular applications (Wescheler, 2006; Duke et al., 2006). This, in part, explains the recent popularity of LiDAR datasets in research applications where fine topographic detail is needed (Barber and Shortbridge, 2006; Wescheler, 2006; Duke et al., 2006; Liu and Wang, 2008). Ultimately, high quality DEMs accurately represent the actual terrain, resulting in more reliable terrain derivatives.

2.2 Factors Determining the Quality of a DEM

Table 2.1 lists the many factors that determine the quality of a DEM, and ultimately the derived terrain attributes.

*Table 2.1. Factors Determining DEM Quality*

<table>
<thead>
<tr>
<th>Factors Determining DEM Quality</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data acquisition technique and elevation data source</td>
<td>Walker and Wilgoose, 1999; Endreny et al., 2000</td>
</tr>
<tr>
<td>Density and distribution of ground elevation points</td>
<td>Bostad and Stowe, 1994; Aguilar et al., 2005</td>
</tr>
<tr>
<td>The interpolation algorithms used to create the DEM</td>
<td>Chaplot et al., 2006; Lindsay and Creed, 2005</td>
</tr>
<tr>
<td>The data structure (regular grid, TIN, contours)</td>
<td>Aguilar et al., 2005</td>
</tr>
<tr>
<td>Terrain roughness</td>
<td>Quinn et al., 1990; Lindsay and Creed, 2005</td>
</tr>
<tr>
<td>The algorithms used to derive terrain attributes</td>
<td>Chaplot et al, 2006</td>
</tr>
<tr>
<td>The horizontal resolution and vertical precision of the elevation data</td>
<td>Wolock and Price, 1994; Zhang and Montgomery, 1994; Gyasi-Agyei et al., 1995; Thompson et al., 2001</td>
</tr>
</tbody>
</table>
Data acquisition techniques and data sources can affect DEM accuracy based on the sampling density of ground elevation points. For example, elevation data collected from a satellite tends to have more data points than data collected from field surveys (Endreny et al., 2000). Studies have shown that as sampling density increases there is a decrease in slope angle and an increase in specific catchment area (Sasowski et al., 1992, Bolstad and Stowe 1994, and Endreny et al., 2000).

The choice of the interpolation algorithm can have major implications on the quality of the DEM. Zimmerman et al. (1999) compared kringing and inverse distance weighing (IDW), arguing that kringing gave a better output because of the ability to account for the spatial structure of the data. Studies completed by Weber and Englund (1992) and Brus et al. (1996) argue that IDW and radial basis have the same or even better results than kringing.

Chaplot et al. (2006) compared five different interpolation algorithms and concluded that when surface area, terrain morphology and sampling density were discounted, there were few differences between the interpolation methods when the sampling densities were high. As sampling density decreased, interpolation results varied. This is due to the fact that as sampling density increases there is a decrease in the space between the known values.

Aguilar et al. (2005) studied the accuracy of five different interpolation techniques based on the data density and terrain roughness. They found that the greatest factor on the quality of interpolation was morphology, followed by sampling density and finally interpolation method. They found that interpolation by multi-radial basis functions were more accurate than that of the logarithmic function for mountainous areas, which
produced similar results in flatter, smoother areas. Thus, it is important to consider the
effect of sampling density on interpolation method. The analyst should choose their
interpolation method based on the sampling density while also considering their
computational limitations.

DEM accuracy is particularly important for hydrological applications because a
DEM should be able to capture all objects on the terrain that alter overland flow. The
horizontal resolution refers to the horizontal spacing of points in the elevation grid.
Maidment and Djokic (2000) stated that there must be consistency between the scale and
model of the physical process under consideration. In the case of a regular grid DEM,
horizontal resolution sets the effective scale of the data set. Many studies have found that
hydrological parameters are sensitive to horizontal resolution; with coarser DEMs often
demonstrating reduced slope and relatively sparse river networks (Band and Moore,
1995; Garbrecht and Martz, 1994; Lindsay and Evans, 2008; Raaflaub and Collins,
2006). Zhang and Montgomery (1994) showed that there was an increase in catchment
wetness and peak flow with coarser resolution. Wolock and Price (1994) showed that as
resolution decreased, specific catchment area increased, especially in small catchments.
Thieken et al. (1999) showed that coarse-resolution DEMs truncate flow lengths and
decrease drainage density. Therefore, coarse-resolution DEMs do not account for all
features in the landscape that alter overland flow. Stream patterns derived from coarse
resolution data rely heavily on large-scale topographic features resulting in generalized
flow patterns.

Gyasi-Agyei et al. (1995) found that accurate drainage patterns could be extracted
from a DEM only if the ratio of average elevation change per pixel (pixel drop) to the
vertical resolution was greater than unity. For example, areas of low relief should have finer vertical resolution than in areas with high relief. In addition to the above findings they found that other terrain attributes such as slope gradient and specific catchment area were not very sensitive to the decrease in vertical precision (Gyasi-Agyei et al., 1995). Terrain attributes are solely dependent on the elevation and therefore their accuracies are directly dependent on the DEMs representation of the terrain. Inaccurate representation of the Earth’s surface can result in various types of errors.

There are three types of errors that are the source of inaccurate elevation data: systematic error, gross errors, and random error (Hengl and Reuter, 2008). Systematic errors are those that reflect bias in the data collection method or sensitivity of an algorithm to compute parameters. The roughness of the landscape, the sampling density and the various DEM interpolations can potentially cause systematic errors, which can cause a follow-on effect on calculated hydrologic parameters (Lunetta et al., 1991; Raaflaub and Collins, 2006). Gross errors (sometimes called blunders or artifacts) are vertical errors in a DEM and can be caused by the mapping procedure or as a result of the interpolation procedure. For example, the input of incorrect elevation values or digitization errors can be classified as gross errors. An example of an interpolation error might be flat terraces caused by poor triangulation algorithms (Chaplot et al., 2006). Gross errors from source data are usually difficult to detect without reference data while gross errors from the interpolation process can be sorted out through visualization techniques such as relief shading and convolution techniques (Wechsler, 2006; Hengl and Reuter, 2008). Random errors result from noise signals and are usually detected prior to preprocessing (Hengl and Reuter, 2008).
All types of errors can have significant effects on calculated hydrologic parameters. The most robust flow routing algorithm or hydrological model will result in a poor output if the DEM used is of poor quality (Hengl and Reuter, 2008).

DEM accuracy is an important consideration in hydrological applications. The source, sampling density, horizontal resolution, vertical resolution and interpolation methods all affect DEM accuracy as well as the terrain attributes derived from the DEM (Wolock and Price, 1994; Walker and Wilgoose, 1999; Aguilar et al., 2005). Because terrain is a major control on overland flow, highly accurate elevation data is particularly important when modelling the spatial distribution of overland flow. From a hydrological standpoint, an accurate DEM is one that captures all of the features on the landscape that influence overland flow (Jenson and Dominque, 1988).

2.3 Light Detection and Ranging

LiDAR is an advanced topographic mapping technology used for collecting dense and highly accurate elevation values about the terrain (Hengl and Reuter, 2008). The integration of lasers, GPS, and Inertial Navigation Systems (INS) make up the LiDAR system. It is the combination of these three technologies that makes this technique far superior to traditional mapping technologies (Shan and Toth, 2008; Campbell, 2007). Figure 2.1 displays the LiDAR system.

To capture a LiDAR image, laser beams are emitted from the instrument sensor towards the terrain, where they are reflected off of the terrain back toward the LiDAR receiving sensor. The receiver accurately measures the travel time of the laser pulse from the start to the return. From the travel time and the speed of light the range measurement
can be calculated. The GPS records the position of each pulse while the INS records the laser orientation. Combining the information from the laser, GPS and INS, accurate x,y,
and z coordinates can be derived from each pulse (Shan and Toth, 2008; Campbell, 2007). The LiDAR system rapidly measures the underlying terrain at 10,000 – 150,000 pulses per second (Heritage and Large, 2009). Some LiDAR systems are capable of recording multiple returns from the same pulse. For example, a laser beam may hit the top of the tree canopy, while another part of the beam may travel further within the canopy and hit branches, or may hit the Earth’s surface, resulting in multiple returns with each part of the beam having its own x, y, and z coordinates (Heritage and Large, 2009).

LiDAR derived DEMs offer some advantages over DEMs created by more traditional approaches, such as photogrammetry. Table 2.2 lists a comparative analysis of these technologies. High point density, light independence, less data collection and processing time and digital format at data acquisition are just some of these advantages (Campbell, 2007; Hengl and Reuter, 2008; Mouton, 2005). LiDAR systems also provide a direct technology for 3D data collection. Raw LIDAR datasets are essentially a collection of spot elevations, also frequently referred to as point clouds (Barber and Shortridge, 2005; Mouton, 2005; Shan and Toth, 2008).

LiDAR derived DEMs have a vertical accuracy of approximately 15 cm and a horizontal accuracy of 50 -100 cm (Campbell, 2007; OGMI, 2009) while DEMs generally used to model hydrological processes have a horizontal resolution of 10-30 m and a vertical resolution of 1-10 m. Schiess and Krogstad (2003) compared the differences of derived terrain attributes between a LiDAR DEM and a photogrammetric DEM by overlaying a 10m DEM on a 2 m LiDAR DEM. The LiDAR DEM provided considerably more topographic detail with incised streams, sharp ridges and abandoned roadbeds, while the photogrammetrically produced contours created a more generalized
Table 2.2: Comparison of LiDAR and Photogrammetric DEMs (Mouton, 2005; Murphy et al., 2008; Barber and Shortridge, 2005)

<table>
<thead>
<tr>
<th>LiDAR</th>
<th>Photogrammetry</th>
</tr>
</thead>
<tbody>
<tr>
<td>High point Density</td>
<td>Low point density</td>
</tr>
<tr>
<td>Light independent (can operate at night)</td>
<td>Light dependent</td>
</tr>
<tr>
<td>Less data and processing time</td>
<td>Time consuming</td>
</tr>
<tr>
<td>Already in digital format at data acquisition</td>
<td>Must be digitized and georeferenced</td>
</tr>
<tr>
<td>Direct 3D data collection</td>
<td>In 2D space</td>
</tr>
<tr>
<td>Vertical resolution 15cm</td>
<td>Vertical Resolution: 1 – 10m</td>
</tr>
<tr>
<td>Horizontal Resolution 15 – 100cm</td>
<td>Horizontal Resolution 10-30m</td>
</tr>
<tr>
<td>Able to map bare earth elevation</td>
<td>Usually interpolated from initial elevation point grid to a higher resolution</td>
</tr>
</tbody>
</table>

landscape. Most importantly, however, the photogrammetrical contour lines did not follow the stream channels. Murphy et al. (2008) modelled stream networks in a watershed using both photogrammetric and LiDAR DEMs. They found that channels derived from LiDAR DEMs had more complex morphology; the flow channels extended further upslope and total channel lengths were in accordance with field mapped networks when compared to photogrammetric DEMs. LiDAR is an advanced technology in that it provides accurate DEMs, which capture features that influence overland flow, making it an essential tool for reliable hydrological simulations.

2.4 Applying DEMs for Hydrological Modelling

DEM's are the primary source for deriving variables used by numerous hydrologic models (Tarboton et al., 1991; Moore et al., 1991; Tribe, 1992; Ludwig et al., 1996).
Moore et al. (1991) summarized the primary and secondary attributes derived from a DEM as well as the hydrologic significance of each attribute. They defined primary attributes as those that are calculated directly from the DEM (Table 2.3). Hydrological connectivity for some of the above attributes (catchment slope, catchment area and catchment length) must be known and therefore knowledge of flow directions is also required (Wu and Huang, 2008; Wilson et al., 2007). Secondary attributes are defined as those computed from two or more primary attributes and assist in describing spatial variability as a function of process (Table 2.4). For example, topographic wetness indices are secondary attributes that describe the spatial distribution of saturated zones for runoff generation and are a function of upslope-contributing area and slope gradient. Of those listed in Table 2.4, slope, flow length, upslope contributing area and watershed area are most significant to hydrologists (Band, 1993; Moore et al., 1991; Tarboton 1997; Wilson and Gallant, 2000). As stated above, to determine hydrologically significant primary attributes, overland flow routes must be known and are usually calculated using flow routing algorithms.
<table>
<thead>
<tr>
<th>Attributes</th>
<th>Definition</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Altitude</td>
<td>Elevation</td>
<td>Climate, vegetation and potential energy</td>
</tr>
<tr>
<td>Upslope height</td>
<td>Mean height of upslope area</td>
<td>Potential energy</td>
</tr>
<tr>
<td>Aspect</td>
<td>Slope azimuth</td>
<td>Solar isolation, evapotranspiration, flora and fauna distribution and abundance</td>
</tr>
<tr>
<td>Slope</td>
<td>Gradient</td>
<td>Overland and subsurface flow velocity and runoff rate, precipitation, vegetation, geomorphology, soil water content, land capability class</td>
</tr>
<tr>
<td>Upslope slope</td>
<td>Mean slope of upslope area</td>
<td>Runoff velocity</td>
</tr>
<tr>
<td>Dispersal slope</td>
<td>Mean slope of dispersal area</td>
<td>Rate of soil drainage</td>
</tr>
<tr>
<td>Catchment slope</td>
<td>Average slope over the catchment</td>
<td>Time of concentration</td>
</tr>
<tr>
<td>Upslope area</td>
<td>Catchment area above a short length of contour</td>
<td>Runoff volume, steady-state runoff rate</td>
</tr>
<tr>
<td>Catchment area</td>
<td>Area draining to catchment outlet</td>
<td>Runoff volume</td>
</tr>
<tr>
<td>Specific catchment area</td>
<td>Upslope area per unit width of contour</td>
<td>Runoff volume, steady-state runoff rate, soil characteristics, soil water content, geomorphology</td>
</tr>
<tr>
<td>Flow path length</td>
<td>Maximum distance of water flow to a point in the catchment</td>
<td>Erosion rates, sediment yield, time of concentration</td>
</tr>
<tr>
<td>Upslope length</td>
<td>Mean length of flowpaths to a point in the catchment</td>
<td>Flow acceleration, erosion rates</td>
</tr>
<tr>
<td>Dispersal length</td>
<td>Distance from a point in the catchment to the outlet</td>
<td>Impedance of soil drainage</td>
</tr>
<tr>
<td>Catchment length</td>
<td>Distance from the highest point to outlet</td>
<td>Overland flow attenuation</td>
</tr>
<tr>
<td>Profile curvature</td>
<td>Slope profile curvature</td>
<td>Flow acceleration, erosion/deposition rate, geomorphology</td>
</tr>
<tr>
<td>Plan curvature</td>
<td>Contour curvature</td>
<td>Converging/diverging flow, soil water content, soil characteristics</td>
</tr>
<tr>
<td>Tangential curvature</td>
<td>Plan curvature multiplied by slope</td>
<td>Provides alternative measure of local flow convergence and divergence</td>
</tr>
<tr>
<td>Elevation percentile</td>
<td>Proportion of cells in a user-defined circle lower that the center cell</td>
<td>Relative landscape position, flora and fauna distribution and abundance</td>
</tr>
</tbody>
</table>
Table 2.4. Secondary attributes derived from a DEM (Wilson and Gallant, 2000)

<table>
<thead>
<tr>
<th>Definition</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W_t = \ln \left[ \frac{A_s}{T \tan \beta} \right]$</td>
<td>This equation assumes steady state conditions and describes the spatial distribution and extent of zones of saturation (i.e., variable source areas) for runoff generation as a function of upslope contribution area, soil transmissivity, and slope gradient.</td>
</tr>
<tr>
<td>$W = \ln \left[ \frac{A_s}{\tan \beta} \right]$</td>
<td>This particular equation assumes steady state conditions and uniform soil properties (i.e., transmissivity is constant throughout the catchment and equal to unity). This pair of equations predicts zones of saturation where $A_s$ is large (typically in converging segments of landscapes), is small (at base of concave slopes where slope gradient is reduced, and $Tt$ is small (on shallow soils). These conditions are usually encountered along drainage paths and in zones of water concentration in landscapes.</td>
</tr>
<tr>
<td>$SPI = A_s \tan \beta$</td>
<td>This quasi-dynamic index substitutes effective drainage area for upslope contributing area and thereby overcomes limitations of steady-state assumption used in the first pair of equations.</td>
</tr>
<tr>
<td>$LS = (m + 1) \left[ \frac{A_s}{2213} \right]^m \left[ \frac{\sin \beta}{0.0896} \right]^n$</td>
<td>The sediment transport capacity index was derived from unit stream power theory and is equivalent to the length-slope factor in the Revised Universal Soil Loss Equation in certain circumstances. Another form of this equation is sometimes used to predict locations of net erosion and net deposition areas.</td>
</tr>
<tr>
<td>$CIT = A_s (\tan \beta)^2$</td>
<td>Variation of stream-power index sometimes used to predict the locations of headwaters of first-order streams (i.e., channel initiation).</td>
</tr>
</tbody>
</table>
2.5 Flow-Routing Algorithms

A flow-routing algorithm can be described as a method that determines the path and direction in which water from one cell is distributed to downslope cells (Grayson and Blöschl, 2001; Lyon, 2003; Maidment and Djokic, 2000). Flowpaths are usually created based on the difference in elevation between cells in which the flow originates and neighboring cells. There are two main types of flow algorithms: single flow direction algorithms (SFD) and multiple flow direction (MFD) algorithms, with many algorithms developed within these categories (O’Callaghan and Mark, 1984; Fairfield and Leymarie, 1991; Tarboton, 1997). Each flow algorithm uses unique methods to derive flow direction and other topographic attributes, therefore producing different results among algorithms. For example, all algorithms are capable of deriving terrain attributes such as upslope-contributing area, specific catchment and stream power index. However, because each algorithm has a unique method of deriving the same terrain attribute, they can produce varying outcomes of the same attributes. These differences can appear throughout the DEM and/or in certain areas of the same DEM (Wilson et al., 2007). Many flow algorithms have been implemented in hydrological models, the most common being the D8 and D-infinity flow algorithms, which are described in more detail below.

2.5.1 Single Flow Direction Algorithm (SFD)

A single flow direction algorithm is essentially the transfer of water from a pixel into one and only one neighbor, which has the lowest elevation (Lyons, 2003). The Deterministic 8 Neighbour (D8) or “steepest decent” algorithm was developed by O’Callaghan and Mark (1984), and is the most basic flow algorithm as it permits flow from one cell to the neighbour cell with the steepest downslope gradient (lowest elevation
value. The aspect (measured in degrees clockwise from north) marks the direction of steepest descent from each grid cell and is the direction from which water would flow from that grid cell. The calculation of steepest gradient is as follows:

Equation 1. Steepest Decent Formula

\[
S = \max_{i=1,8} \frac{z_g - z_i}{h \Phi (i)}
\]

Where, \( \Phi(i) \) = 1 for cardinal neighbors (grid cells where \( i = 2, 4, 6, \) and 8), \( \Phi(i) = \sqrt{2} \) for diagonal neighbors (to account for extra distance travelled for those cells) and \( z = \) elevation of a particular cell. This algorithm works best to simulate flow for rivers and streams in valleys. Upslope contributing area and specific catchment are easily derived from D8 considering that all flow from one pixel is routed to the steepest downslope pixel. By mapping the direction of overland flow drainage networks, watershed boundaries can be easily delineated. The disadvantages of D8 are that it generalizes flow direction from one grid cell to another and therefore cannot model divergent flow (Tribe, 1992; Wilson et al., 2007). The well-known expression of this limitation is the parallel flowpaths in either the cardinal or diagonal direction (Tribe, 1992). Although there are many limitations to the D8 algorithm, it is extremely useful in extracting river network maps, longitudinal profiles and basin boundaries (Jenson and Dominique, 1988; Mouton, 2005). It is also a primary component for other SFDs and MFDs.

2.5.2 Multiple Flow Direction Algorithms (MFD)

Unlike SFDs, MFDs handle divergent flow by partitioning the flow out of one cell into all lower neighbours (Hengl and Reuter, 2008). Tarboton (1997) developed the D-infinity MFD algorithm. In this algorithm, flow direction is determined in the direction of
steepest descent and is represented by a continuous quantity from 0 – 2 (Equation 2).

Figure 2.2 shows the procedures for this algorithm. Connecting the centres of all eight cells creates eight triangle facets from which the drainage direction vector is determined. The drainage direction vector is calculated using following Equation 2 below.

**Equation 2. Drainage direction vector for D-infinity**

\[ d1 = \frac{4 \alpha_2}{\pi}, \quad d2 = \frac{4 \alpha_1}{\pi} \]

The angles \( \alpha_1 \) and \( \alpha_2 \) are measured on a horizontal planar surface between the drainage direction vector and the vectors of the two pixels on either side of it (\( \alpha_1 + \alpha_2 = 45^\circ \)).

*Figure 2.2. Flow allocation in D-infinity following Tarboton, (1997)*
Each facet has a downslope vector, which may be at an angle within or outside the 45-degree angle range of the facet at the center point. If the slope vector angle is within the facet angle it represents steepest flow direction. If the slope vector angle is outside of the facet, the steepest flow direction associated with that facet is taken along the steepest edge. The drainage direction is associated with the pixel taken as the direction of the steepest slope vector from all eight facets. This method eliminates the bias generated in D8 and the over dispersion created by MFDs.

2.6 Comparisons and Limitations of algorithms

There are many comparative performance studies done on the flow algorithms listed above. Wolock and McCabe (1995) found that MFD algorithms produced smoother patterns of topographic wetness indices across a DEM, which is an indication of less abrupt variations in the magnitude of topographic wetness for adjacent cells. Desmet and Grovers (1996) compared the upslope contributing areas using both SFD and MFD algorithms for a small catchment. While the MFD algorithms produced distinctive spatial patterns, the SFD algorithms produced patterns different from each other and that of the MFD algorithms. They found that algorithms allowing flow to one or two downslope neighbours produced a stronger correlation with the main drainage lines.

Endreny and Woods (2003) determined a series of topographic attributes for two small fields in New Jersey and compared the spatial agreement of the simulated overland flowpaths with field data. The lowest agreements were between D8 (which constrained flow to a single neighbour) and FD8 (which directed flow into all of its neighbours). The D-infinity algorithm was one of the algorithms that produced the most realistic paths.
Wilson et al. (2007) also compared the same algorithms. They classified low flow cells as those near hilltops and ridgelines having a specific catchment area (SCA) of less than 10m$^2$ and stream channels as having an SCA greater than 5300 m$^2$. They found that D8 and Rho-8 produced many low flow cells in the wrong parts of the landscape. In addition, they concluded that the algorithms produced the largest variations in simulated flow patterns at high elevations, but that as elevation decreased, the algorithms behaved more like each other.

These studies effectively show that different algorithms produce different spatial patterns of overland flow in generalized landscapes. However, all of the SFD and MFD algorithms discussed above are solely dependent on altitude and does not account for other features that can significantly alter overland flow. For example, Duke et al. (2003) stated that roads, bridges and embankments, which have topographic expressions and thus alter overland flow, are not included in the hydrological models, which results in oversimplified simulations. Ludwig et al. (1996) stated that tillage furrows 2 cm in depth can significantly alter flow networks and erosion patterns. Few, if any, DEM-based flow-routing algorithms can resolve sub-meter features. All flow algorithms are exclusively dependent on the accuracy of the elevation data; in other words, these algorithms are solely reliant on the DEM data. The most robust flow algorithm would result in a poor output if a DEM were of low quality and resolution (Hengl and Reuter, 2008).

Before applying flow routing algorithms to a raster DEM, it is necessary that the spurious sinks be removed. That is, the DEM must be preprocessed so that there are continuous flowpaths. This is usually done using specific algorithms that adjust the elevation values of artifact depressions and flat areas.
2.7 Hydrologic conditioning DEMs

Depressions, sometimes referred to as sinks or pits, are ubiquitous in a DEM. Depressions are cells or groups of neighbouring cells that do not have an outlet; they are cells that are surrounded by higher elevation values, creating an area of internal drainage.

Topographic depressions in a DEM can represent real depressions in the landscape or artifacts. Artifact depressions are errors in the elevation data (Henger and Reuter, 2008). There are numerous sources for artifact depressions, mostly stemming from an error in data collection techniques or input errors (Reiger, 1998; Walker and Wilgoose, 1999; Aguilar et al., 2005). Artifacts can also be caused by interpolation during the DEM creation process (Aguilar et al., 2005; Chaplot et al., 2006), by the averaging or rounding of elevation values for each cell (Bolstad and Stowe 1994; Lindsay and Creed., 2005) or can be due to the limited vertical and horizontal resolution of the elevation data (Martz and Gerbercht, 1993; Wolock and Price, 1994; Thompson et al., 2001).

Real depressions are depressions in a DEM that represent actual topographic features. Although real depressions are not as common as artifact depressions, they do exist in some geomorphological landscapes such as glacial landscapes and karst, or in human-modified landscapes, from anthropogenic features such as ditches, detention basins and quarries (Mark, 1984; Zanbergen, 2010). Before using DEMs for hydrological applications, artifact depressions need to be removed.

Hydrologic conditioning of DEMs ensures the removal of all depressions and flat areas. Depressions and flat areas are quite problematic for hydrological modelling because they often artificially truncate flowpaths and alter flow direction (Thompson et al., 2001; Lindsay and Creed, 2005). Hydrologically sound DEMs are DEMS that are
depression-less, thereby allowing all cells to be connected to an outlet. Currently, there are many techniques used to hydrologically condition a DEM, with each technique having its own novel procedure for resolving depressions (O'Callaghan and Mark, 1984; Marks et al., 1984; Band, 1986; Jenson and Domingue, 1988; Martz and Gerbercht, 1993). It is important to keep in mind the manner in which these artifacts are resolved, as it determines the quality of the hydrological parameters extracted from a DEM (Band, 1986).

2.7.1 Drainage/ Flow Enforcement

Drainage enforcement refers to the procedure of creating a hydrologically sound DEM. That is, drainage enforcement techniques remove depressions in a DEM, thus ensuring that all cells are hydrologically connected. Drainage enforcement is usually done after DEMs are interpolated. There are many methods of removing pit cells and flat areas to create hydrologically conditioned DEMs. Reviews of some of these methods are listed below.

Smoothing is an early method developed by Mark and Aronson (1984) in which a moving average is used to eradicate small depressions. Many have criticized this method as it both alters elevation values and creates new sinks (Jenson and Dominique, 1988; Collins, 1975). In addition, smoothing generalizes the landscape by flattening the terrain and causing systematic errors to the DEM. Conversely, O'Callaghan and Mark (1984) found that one pass of smoothing can remove more than 90% of depressions. They suggested using smoothing and another method such as filling to remove the remaining depressions.
Breaching and filling are alternative flow enforcement techniques used to remove depressions in DEMs. While filling removes pits by raising the elevation value of the single cell, breaching lowers the cells adjacent to the pit. While both methods are utilized for hydrological conditioning of a DEM, filling has become the most popular flow enforcement technique because of its simplicity.

Jenson (1991) refers to filling as the preliminary step in the DEM conditioning phase. O’Callaghan and Mark (1984) created a filling algorithm that resolves single cell depressions. First, the pour points for all depressions are identified. Next, the flow direction is calculated from the pour point to the depression, resulting in the flow being re-directed from the sink through the pour point to an adjacent basin. Although this algorithm has been widely implemented, it has difficulty with complex depressions. The technique created by Jenson and Dominique (1988) raises a cell or multiple cells that have the lowest elevations when compared to surrounding cells, and these cells are then changed to the same value as the depression spill position. This method for filling depressions is quite similar to that of O’ Callaghan and Mark (1984), however it is capable of resolving complex depressions and flat areas within the DEM.

Wang and Liu, (2006) created a new method for detecting and filling sinks that involves concepts such as spill elevation and least-cost search to create a depression-less surface. The spill elevation refers to the minimum elevation cells that need to be raised to create a continuous path. The least-cost search starts at the outlet, linking the outlet to the interior cells using an upstream strategy. Ultimately, the cell with the lowest spill elevation value for the interior cell among the various paths is chosen. Similarly, Planchon and Darboux (2001) created a filling algorithm by first inundating the entire
DEM, and then iteratively draining the excess water from each cell. The entire DEM would then be scanned from eight directions to determine downslope path. Using a seed cell, the algorithm searches for an upstream tree by following dependence links and the excess water is removed from this network. Finally, the value of the highest pour point on the flowpath to an outlet is assigned to depressions, allowing for the water to be drained from these artifacts.

Of all the algorithms cited, the flow-enforcement algorithm of Jenson and Domingue (1988) is the most widely implemented with GIS and hydrological models. ArcGIS (ERSI, 2011) and GRASS (grass development Team, 2011) have it as their default flow enforcement method. However, Jenson and Dominque’s (1998) algorithm introduces systematic error, as it assumes that all pits are underestimation errors. The breaching method described below overcomes this disadvantage.

Breaching is best suited to situations where sinks are created by blocking flowpaths by overestimated elevation. Breaching is the lowering of elevation and is effective for DEMs that have low relief (Maidment and Djokic, 2000). Martz and Garbercht (1999) developed a breaching algorithm, which lowered the elevation pour points of the depression instead of raising the depression cell. Some authors have combined both filling and breaching techniques so that a user-defined criterion is set to identify whether cells are filled or breached. For example, Martz and Garbercht (1998) developed the constrained breaching algorithm in which the breach channel length was limited to two grid cells, if the depression was greater than two cells the depression was filled.

The impact reduction approach (IRA) developed by Lindsay and Creed (2005) chooses one method, filling or breaching, depending on which method resulted in the
least amount of modification of the DEM. The IRA is comprised of four steps, the first of which creates two DEMs, one for filling and one for breaching. The second step involves numbering the modified cells (NMD) and evaluating the mean absolute difference (MAD) for both the filling and the breaching methods. The third step evaluates the second step and chooses either filling or breaching, and the fourth step performs the depression removal method that modified the DEM the least. Lindsay and Creed (2005) stated that filling drastically impacted the DEM when compared to breaching, yet filling is the preferred and most studied method for depression removal. The reason for this is that it is relatively simple to directly alter the depression’s elevation value but methodologies for altering neighboring cells can easily become convoluted, resulting in severely altered terrain represented by the DEM.

Another common problem related to applying DEMs to model surface flow involves flow over flat areas. Flow over flat areas is problematic for modelling overland flow because there are no surrounding cells that have a lower elevation in which the flow can be routed. Flat areas are usually the second step when hydrologically conditioning a DEM; that is, they are resolved directly after depressions are either filled or breached. For example, the initial step of the Jenson and Dominque (1988) flow enforcement algorithm involves filling depressions, which is followed by resolving the problems created by flow over flat areas. To do this, Jenson and Dominque (1988) created an iterative procedure in which flat cells are assigned a single flow direction to a drainage network without actually changing the elevation values. In the first iteration, cells beside the outlet are drained. In the second iteration, cells next to the cell in the first step are drained. This procedure works backwards from the outlet to depict the course of flow.
Similarly, Martz and Gerbrecht (1998) created two algorithms for hydrological conditioning of a DEM. The first algorithm involves breaching all depressions in the DEM, while the second algorithm resolves flat areas. This resolution of flat areas involves making small changes to the elevation of flat cells in order to implement flow by creating a gradient. These changes are also done iteratively.

Techniques such as smoothing, filling and breaching are needed to create hydrologically sound DEMs. However, traditional drainage enforcement techniques such as the smoothing, filling and breaching algorithms discussed previously in this paper are not suited for incorporating anthropogenic linear flowpaths into hydrological analyses. The issues surrounding the modelling of flow in human-modified landscapes as well as the current techniques available for resolving these obstacles are discussed in more detail below.

### 2.8 Modelling flow in human-modified landscapes.

Human-modified landscapes adversely affect overland flow volumes and pathways. Anthropogenic structures, particularly roads, bridges, and embankments, alter overland flow patterns significantly and yet are not included in DEM based flow modelling, mainly because these features are not generally captured in coarse resolution DEMs (Duke et al., 2003; Duke et al., 2006). For this reason, DEM based flow routing in human modified landscapes have not been well documented, with many studies focusing on moderately unmodified landscapes (Wilson et al., 2007). When evaluating the spatial extent of overland flow in an area, one must consider anthropogenic features. For example, in the absence of roads, runoff tends to follow the gravitational flowpath from
higher to lower elevation, eventually reaching the stream network. However, when roads, ditches and road underpasses exist on the landscape, they tend to have adverse effects on the natural hydrology of a watershed. Jones et al. (2000) found that roads can alter the balance between the intensity of flood peak flows and stream network resistance to change. Jones et al. (2000) also found that roads can intensify and redistribute floods. LaMarche and Lettenmaier (2001) used GIS to model overland flow and discovered a 2 to 12% increase in peak flow due to roads. Wigmosta and Perkins (2001) found that some road networks in their study increased the contributing area of some channel segments and decreased the area draining in others. In addition, many authors have recognized that tillage furrows 2 cm or greater in depth can affect flow direction (Ludwig et al., 1996; Souchere et al., 1998; Cerdan et al., 2001; Takken et al., 2001).

Jones et al. (2000) classified four types of interactions between roads and runoff. These interactions are depicted in Figure 2.3 as a corridor, barrier, sink or source. The type of interaction between runoff and road networks depends on the location of the road along the hillslope. For example, roads commonly run parallel to main runoff networks and consequently act as a barrier to overland flow. Therefore, incorporating road features along with the anthropogenic conduits used to manage overland flow into hydrological models is important. Many studies have shown that the presence of road networks, ditches and culverts increase or decrease channel gradients and runoff velocities, leading to increased sedimentation and higher peak flow discharges (Montgomery, 1994; Croke and Mockler, 2001; Wemple and Jones, 2003). Two methods of incorporating anthropogenic features to accurately model flow include incorporating auxiliary data into
a DEM or into a LiDAR dataset in which these features are already defined and are discussed below.

![Figure 2.3 Runoff interaction with roads (Jones et al., 2000)](image)

2.8.1 Incorporating ditches and culverts through auxiliary data.

Coarse-resolution DEMs cannot capture ditches and road underpasses, and these features therefore need to be incorporated into the DEM through auxiliary data. Duke et al. (2003; 2006) incorporated such features by creating the Road Enforcement Algorithm (REA) and the Canal Enforcement Algorithm (CEA). The REA and CEA involve raising road elevations and lowering ditch elevations within the DEM to produce a more realistic flow direction.

2.8.1.1 Road Enforcement Algorithm (REA)

The REA was designed to include roads into the DEM allowing overland flow to divert to either side of a road independently by defining a “runoff collector network”. The runoff collector network is made up of linear depression features and areas upslope and next to raised roads. The runoff collector network significantly altered flow in three ways:
1) Elevated roads could form overland flow barriers and path sinks by directing flow parallel to and on the upslope side of the road (Figure 2.4);

![Figure 2.4](image)

*Figure 2.4 REA enforces flow on elevated roads. Runoff on the upslope side of the road is re-routed whereas runoff on the downslope side is not affected. (Duke et al., 2003)*

2) Ditches parallel to the road could create flow path corridors and sinks, thereby causing the runoff to flow either parallel or perpendicular to the road (Figure 2.5);

![Figure 2.5](image)

*Figure 2.5: REA flow enforcement along roadside ditches. Runoff is re-routed on both the upslope and downslope sides of the road (Duke et al., 2003)*
3) Roads with flat cross sectional profiles could cause minimum effects on the flow routes (Figure 2.6);

![Figure 2.6 Flat cross sectional road profile in which runoff is not considerably affected (Duke et al., 2003)](image)

Based on the above effects roads and ditches have on overland flow, the REA enforces simulated flow by converging surface flow toward linear depressions, thereby predicting road runoff locations.

2.8.1.2 Canal Enforcement Algorithm (CEA)

The purpose of the CEA is to incorporate culvert and irrigation canal networks into a DEM based on cross split patterns, which combine split flow channels and cross flow patterns. A split flow channel refers to the division of the channel into two parts. A cross flow pattern refers to routing water through culverts, which can be situated over or under drainage courses resulting in a diagonal flow direction (Figure 2.7).
Figure 2.7 Flow patterns in an irrigated landscape: (a) cross-flow scenario with an irrigation canal as a solid line (grid representation shown as gray squares) and a stream course as a dotted line (white squares); (b) flow direction representation of cross-flow pattern (Duke et al., 2006)

2.8.1.3 Rural Infrastructure Digital Elevation Model (RIDEM)

The Rural Infrastructure Digital Elevation Model (RIDEM) is essentially an amalgamation of the REA and CEA resulting in an improved representation of overland flow in human modified landscapes. The outputs from this model include improved flow direction matrices, improved watershed boundaries and dead drainage (Duke et al., 2006).

2.9 LiDAR DEMs.

Since resolution is the underlying issue for accurately modelling overland flow in human modified landscapes, LiDAR derived elevation data is thus considered extremely valuable. LiDAR systems are capable of acquiring points at sub-meter resolution, therefore the DEMs constructed from these LiDAR datasets will already contain anthropogenic features such as ditches and road underpasses. Although LiDAR DEMs are fine resolution, these data come with many challenges, particularly concerning the
application of preprocessing techniques used with DEMs acquired through traditional means. For example, depression filling is the method usually used to eliminate sinks within DEMs. However, when filling is applied to LiDAR data, sub-meter features such as linear anthropogenic features are removed in the LiDAR DEM. While in reality ditches are linear depressions that facilitate surface flow, in a DEM flow direction algorithms identify ditches as local sinks, obstructing simulated flowpaths. The main reason for this is that LiDAR systems might not be able to capture these features completely mainly because of their narrow morphologies. For example, some laser beams may hit the ditch while others may completely miss the ditch due to various reasons, such as vegetation that obscures the feature from the laser or the presence of water in the ditch. As a result, ditches are usually represented as discontinuous features in the LiDAR dataset. Culverts are linear structures, usually installed under roadways to facilitate overland flow across embankments. While in reality water may be diverted under a road bank through a culvert or under a bridge, these road underpasses are not captured in the LiDAR data, resulting in artificial damming.

Technicians usually correct these digital dams by manually scanning the DEMs and burning in ditches and road underpasses, but this process is both time-consuming and inefficient. This thesis will explain a novel automated flow enforcement technique for enforcing flowpaths through linear anthropogenic features. In particular, this new automated technique is capable of enforcing flow through ditches on either side of road embankments and enforcing flow through road underpasses. The novel technique presented in this thesis is expected to incorporate small-scale linear anthropogenic
features into DEM based hydrological models, thereby improving the accuracy of the simulated flowpaths in human modified landscapes.
3.0 BREACHING ALGORITHM

3.1 Introduction

DEM are commonly used to model surface flow paths and for many other related hydrological applications, such as stream network extraction (Tarboton, 1997) and watershed mapping (Jenson and Dominque, 1988; Costa-Cabral and Burges, 1994). Over the past several decades, considerable research has been carried out to improve the algorithms used in DEM-based flow path modelling (O’Callahan and Mark, 1984; Tribe 1992) and to improve the application of higher quality and finer-resolution DEM data (Mac Millian, 2003; Barber and Shortbridge, 2004; Murphy et al., 2008). The use of fine-resolution DEMs (i.e. < 5 m) presents new challenges for flowpath modelling. For example, fine-resolution DEMs, such as those created by LiDAR technologies, frequently contain road embankments in human-modified landscapes, which can create several issues in flow-routing algorithms.

Despite the fact that roads, and their associated ditches and culverts, have been shown to alter the surface hydrology of regions (Jones, 2000; LaMarche and Lettenmaier, 2001; Wigmosta and Perkins, 2001), most existing flow-routing algorithms are incapable of handling these topographic features. These existing algorithms cannot properly assess underpass sites such as culverts and bridges, which appear as dams in DEM data. Hydrologically significant terrain attributes derived from flowlines computed on fine-resolution DEMs can therefore be grossly incorrect because of the erratic behavior of flow-routing algorithms at such sites. The purpose of this study is to present and evaluate a new algorithm for addressing the issue of flowpath modelling using fine-resolution DEMs.
3.2 Background

Roads affect hydrology within a watershed in many ways, including constraining or diverting surface flowpaths, altering runoff velocities, and reducing low flows and increasing peak flows (Wemple et al., 1996; Jones et al., 2000; Wigmosta and Perkins, 2001). Nevertheless, roads and other embankment features are seldom incorporated into hydrological analysis largely due to the limited ability of commonly available DEM data to represent these features and the inability of most existing flow algorithms to accommodate the special cases created by the presence of embankment features. Currently, there are two approaches for representing the ditches, culverts and bridges that are associated with embankments. The first method involves the use of a specialized algorithm, which uses coarse-resolution DEMs and the associated ancillary data (vector data of the roads and drainage features) as primary inputs (Duke et al., 2006). The second method involves pre-processing fine-resolution DEMs that already contain embankments and canal features in such a way so that drainage in the image is consistent to the known effects of surface flow.

Duke et al. (2003, 2006) incorporated ditches and culverts by creating the Road Enforcement Algorithm (REA) and the Rural Infrastructure Digital Elevation Model (RIDEM). These algorithms were created for coarse-resolution DEM data because coarse-resolution DEMs tend to represent large-scale features and cannot represent small-scale features such as ditches and road underpasses.

The REA was designed to include roads in a DEM by defining a runoff collector network. The runoff collector network is the collection of linear depression features and areas next to and upslope of raised roads. The REA algorithm forced overland flow on
either side of the road by converging the flow towards depressions. The runoff collector network significantly altered flow in three ways:

1) Elevated roads could form overland flow barriers and path sinks by directing flow parallel to and on the upslope side of the road;
2) Ditches parallel to the road could create flow path corridors and sinks, thereby causing the runoff to flow parallel or perpendicular to the road;
3) Roads with flat cross sectional profile would cause minimum effect on the flow routes.

The REA identified road runoff interactions by manipulation of flow direction matrices along road embankments, which act as both obstructions and openings for flow paths.

Duke et al. (2006) combined the REA and the Canal Enforcement Algorithm (CEA) to create the RIDEM. While the REA incorporates road embankments, the CEA was designed to incorporate culvert and irrigation canal networks based on both cross-flow patterns and split-flow patterns. A split-flow channel refers to the division of a channel into two parts. A cross-flow pattern refers to the routing of water through culverts, which can be situated over or under drainage courses, resulting in a diagonal flow direction. Ancillary data is needed for the REA, the CEA, and therefore the RIDEM, to function. RIDEM results have shown promising improvements in modelled flow direction matrices, watershed boundaries and dead drainage.

Fine-resolution DEMs, such as those acquired through LiDAR, have the ability to acquire data at meter and sub-meter resolution and thus are able to represent small-scale features that affect overland flow. For example, hydrologically significant terrain attributes such as catchment slopes (Hill and Neary, 2005), watershed boundaries (Liu et
al., 2005) and wetlands depressions (Gritzner, 2006), have been derived from LiDAR DEMs, and the accuracy of all of these parameters were significantly improved when compared to data acquired from conventional techniques. What makes LiDAR far superior to traditional mapping technologies is the integration of laser scanners, Global Positioning Systems (GPS) and Inertial Navigation Systems (INS) (Hengl and Reuter, 2009). Laser scanners are uniquely capable of capturing topographic expressions more completely because of a particularly high point density (approximately 10,000-50,000 pulses per second) (Barber and Shortbridge, 2005; Poppenga et al., 2009). The GPS records the x, y, and z co-ordinates of each laser pulse, while the INS records the roll, pitch and yaw of the aircraft, which in turn records the orientation of each laser pulse. Due to the unique properties of these technologies, LiDAR has the potential of creating highly accurate, fine-resolution DEMs that represent both natural and anthropogenic features at various scales. These data are usually interpolated to a regular grid and then preprocessed to be made suitable for hydrological applications.

Despite the fact that ditches and road underpasses are represented in the LiDAR data, drainage patterns are by no means improved by the incorporation of these high-resolution data. Ditches and culverts are local linear anthropogenic features that are used to facilitate flowpaths, but flow routing in ditches and underpasses can be problematic when using LiDAR DEMs.

Murphy et al. (2008) stated that LiDAR DEMs capture anthropogenic features that affect overland flow by blocking and channeling flow. MacMillan (2003) also commented on the inaccuracy of altered flowpaths created using LiDAR models, highlighting the incapacity of the DEM to represent the Earth’s surface in 3D. There has
been limited research in addressing the potential of LiDAR DEMs to model surface
drainage patterns in human-modified landscapes. Barber and Shortridge (2005) showed
that flow-related derivatives were incorrect if culverts were situated along elevated roads.
They noted that without identifying culverts, LiDAR DEMs were “hydrologically
challenged” (Barber and Shortridge, 2005). Poppenga et al. (2009) emphasized the need
for the development of a novel automated method for identifying anthropogenic
obstructions. They stated that if culverts and ditches were not identified in the DEM,
runoff would flow over the obstruction in the wrong location or drain the flow in the
opposite direction (Poppenga et al. 2009). A digital dam is the best term to describe the
effect of road embankments and bridges on simulated flow. While in reality water would
pass under a road or bridge, in the DEM, water accumulates behind the road and bridge
as it would behind a dam, confining flow to a section within the watershed, which would
eventually spill over the road embankments. Thus, before using LiDAR DEMs for
hydrological applications, DEMs are usually preprocessed to remove surface flow
obstructions such as depressions and road embankments, thereby enforcing flow through
underpasses such as culverts and bridges.

Depressions in a DEM are ubiquitous and must be removed to allow for continuous
flowpaths. Depressions are a single cell or set of connected cells that are not connected to
an outlet, which create an area of internal drainage. Drainage enforcement techniques
remove depressions in a DEM, thereby ensuring that all cells are hydrologically
connected. Drainage enforcement refers to the procedure of creating hydrologically
sound DEMs.
Drainage enforcement is usually applied after DEMs are interpolated. Breaching and filling are drainage enforcement techniques that are used to remove depressions in DEMs. Filling methods assumes that all depression cells are underestimated and therefore removes pits (depressions) by raising the elevation values of the depressions. Conversely, breaching methods lowers the elevation values of the cells adjacent to the pit. While both methods are utilized for the hydrological conditioning of a DEM, filling has become the most popular flow enforcement technique by far mainly because the depression filling method provides the same solution regardless of the specific algorithm chosen. That is, no matter the filling algorithm used, the end result is the same; all depressions are removed by raising the elevation values of the cell. However several authors had noted that breaching is a more appropriate solution for depression removal (Reiger, 1998; Lindsay and Creed, 2005). For example Lindsay and Creed (2005) showed that breaching had a smaller impact on the spatial and statistical distributions of terrain attributes when compared to depression filling.

Unlike conventionally acquired coarse-resolution DEMs, which usually have few larger depressions, LiDAR derived DEMs contain a large number of small depressions. This is a result of the combination of LiDAR to capture small-scale surface roughness in which the landscape drainage is much disorganized and the fact that the error in the LiDAR DEM is speckle (white noise) (Kaiser et al., 2010; Zandbergen, 2010). A study conducted by Zandbergen (2010) showed that fine-resolution LiDAR DEMs contain a large portion of small, shallow depressions. The six meter LiDAR DEM that was used in Zandbergen (2010) study contained approximately 500 depressions per square kilometer. It was noted that this noise effect created by the small depressions might increase with
finer resolutions (less than six meters). Zandbergen (2010) also acknowledged the large number of artificial and sometimes deep depressions created by road underpasses. In Zandbergen’s (2010) study, most of the major road underpass depressions were removed; however, there were a substantial number of artifact depressions that were caused by minor roads. These road underpass depressions, while relatively small in number, caused challenges in error propagation because stochastic depression modelling predicted a high probability value for the presence of a depression. However this high value did not account for the systematic errors, which were the road underpasses in Zandbergen’s (2010) study. The current pre-processing algorithm used on DEMs acquired through traditional means cannot be applied to fine-resolution LiDAR DEMs.

Existing preprocessing techniques such as filling and breaching do not necessarily produce the desired results when applied to LiDAR DEMs. Filling is not appropriate for these data because depression filling tends to remove features such as ditches and road underpasses, creating digital ponds behind road and bridges that alter natural drainage patterns, especially in gently sloping landscapes. This is because filling algorithms assume that all depressions are caused by an underestimation of the elevations of grid cells in the DEM (Rieger, 1998) and thus raise the elevation values of ditches and road underpasses. However, the digital dams in the LiDAR data are caused by the higher elevation values of the road and bridges and, therefore, road and bridge elevation values are an overestimation for flow-routing. For this reason, breaching is a better approach for resolving digital dams in LiDAR DEMs.

Breaching has been deemed most appropriate for use with LiDAR images as it lowers the elevation values caused by the overestimation of grid cells. Unlike depression
filling, which only affects the depression, depression breaching affects the neighbouring cells of the depression (Rieger, 1998). Depression breaching lowers the surrounding cells of the depression, which then creates a continuous flowpath for surface flow. There are many solutions for each depression in a DEM when using depression breaching. Therefore, criteria are required by the depression breaching algorithm to provide the most suitable solution. For example, Martz and Gerbercht (1999) choose to use the minimum breach distance, which meant that they only considered a solution that was within a specified distance from the depression. Depressions near the edge of a grid are prone to edge-effects when using a breaching method. Any depressions near the edge of the grid have the potential to have edge-effects. This means that there are cases where the target cell may be beyond the edge of the grid and therefore the depression will not be resolved. However, a minimum distance criterion is not appropriate for solving digital dams in the LiDAR derived DEMs, as LiDAR DEMs contain substantially more and smaller depressions, when compared to traditionally acquired DEMs. There is a high probability that the distance between a depression cell and a solution cell is shorter than the length of road underpasses in the LiDAR DEM. In addition, ditches are local linear depressions used to channelize water to culverts. Artifact depressions are usually associated with ditches because ditches are long narrow features, which are usually flat in the LiDAR DEMs. Breaching is analogous to creating a trench through a dam and is therefore better suited for enforcing flow through road underpasses as well as narrow ditch channels.

LiDAR DEMs offer a novel and much more efficient alternative for incorporating ditches and road underpasses when compared to using coarse-resolution DEMs and manually integrating ancillary data. Duke et al. (2006) recognized the potential for using
LiDAR data to incorporate ditches and culverts, but they still needed to incorporate ancillary data, as this was a more feasible option due to the costs associated with LiDAR data at the time. However, as LiDAR datasets continue to decrease in price and increase in availability, they will play an important role in drainage mapping applications in the near future (Hengl and Reuter, 2009). The overall aim of this research is therefore to improve the modelling of surface drainage pathways in landscapes that have been heavily altered by anthropogenic infrastructure by developing and testing a novel technique that uses the topographic information contained in LiDAR DEMs.

3.3 Breaching Algorithm Rationale

There are two approaches for incorporating anthropogenic features such as ditches and road underpasses into hydrological analysis: 1) modify the flow-routing algorithm and 2) modify the flow enforcement algorithm. Duke et.al (2006) adopted the first approach, where the flow-routing algorithm was modified to account for these anthropogenic features. The second approach would be to modify a flow enforcement algorithm to create hydrologically sound DEMs. The latter approach alters the DEM so that all cells are connected to an outlet so that a standard flow routing algorithm can be applied to the DEM to simulate surface drainage patterns. The latter approach was taken for this work.

Although the results of manually processing ditches and road underpasses into hydrological models have been demonstrated to be satisfactory (Poppenga et al., 2009; Kaiser et al., 2010), this method is by no means efficient. Where geospatial datasets containing culvert and bridge locations are available, these features can be incorporated
in the DEM at the pre-processing stage; however, culvert locations are rarely mapped outside of urban areas. Sometimes, technicians embed stream vector data into the elevation data; however drainage data are usually inconsistent as they are usually acquired from coarse-resolution DEMs. As such, stream vector data tend to be of lower accuracy than those mapped from LiDAR DEMs, particularly for small headwater streams. Including these ancillary data into the DEM preprocessing step is then very questionable. In this study, a new breaching algorithm was developed to overcome the limitations of the current approaches by solving the problems related to flowpath modelling when using fine-resolution DEM data that contain roads and other anthropogenic embankments.

3.3.1 Algorithm Description

Removing depressions is necessary when using DEMs for hydrological applications. Enforcing flowpaths by removing depressions should be done using a technique that causes the least modification of the DEM (Lindsay and Creed, 2005). The breaching algorithm developed in this paper is a drainage enforcement technique used to incorporate road underpasses and roadside ditches into hydrological analyses. The preprocessing algorithm presented in this thesis is based on the breaching and least-cost pathway techniques.

Breaching is the process of lowering the elevation values of cells along a breach path to create a monotonically descending flowpath from a depression to the target cell (Martz and Garbercht, 1999). A target cell is defined as all cells lower than the depressions with a downslope neighbor. There are likely to be many distant downslope cells from a depression, and therefore several possible overland flowpaths; however, the
algorithm chooses the course that would have the least impact on the DEM. To do this, it performs a least-cost pathway analysis whereby the cost criteria to breach a pathway is based on the amount of material that would need to be removed along the pathway to make a continuous course.

The least-cost pathway is a well-known GIS function and has been applied to many areas of study such as road planning (Atkinson et al., 2005) and corridor design for animal movement (LaRue and Nielsen, 2008). It involves creating a path between two non-adjacent cells based on a cost criterion by travelling backwards from the destination to the source cell. The target cells in this application of the least-cost pathway analysis technique are all cells not connected to an outlet (depression), while the destination cells are those that are lower than the target cell and also have a downslope neighbour. The cost criterion for this analysis is the amount of breaching that is needed to connect the depression cell with the distant downslope cell. The procedures for least-cost pathway analysis are as follows (Longley et al., 2001):

1) A cost value (the amount of material needed to be removed) is assigned to each grid cell.

2) An accumulated cost (the accumulated elevation each cell needs to be lowered to connect the depression and the distant downslope cell) is calculated.

3) The path with the least accumulated cost is chosen as the least-cost pathway.

Once the pathway with the least cost is identified, the cells are breached to allow for a continuous flow path. Figure 3.1 shows the flow chart for the novel algorithm developed in this thesis.
Figure 3.1. Flowchart of breaching algorithm

All together there are three components involved in breaching the algorithm, which are to identify depressions, find solutions, and breach the target and destination.
cells. The initial component identifies and orders all depressions in DEM. The second component requires finding a solution for the pit (least-cost pathway analysis). The third component involves breaching the cells along that path.

The solution for the majority of depressions is generally within a short distance around the depression cell. Performing a confined (5x5 cell) breach on the depression cell allows for these numerous depressions to be solved with a reduction in processing time. The breaching algorithm presented in this thesis solves all depressions regardless of their origin, as it does not differentiate between depressions originating from an embankment or from another obstruction. Using a combination of the least-cost pathway analysis and the breaching technique allows the algorithm to resolve depressions produced by linear anthropogenic features, such as roads and bridges, along apparent drainage courses.

Table 3.1 lists the inputs needed in the user-defined features for the novel breaching algorithm.

<table>
<thead>
<tr>
<th>Input Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input Raster</td>
<td>Original raster image</td>
</tr>
<tr>
<td>Output Raster</td>
<td>Preprocessed Image</td>
</tr>
<tr>
<td>Breach Length</td>
<td>Determines the distance from depression which the algorithm searches to find a solution (measured in grid cells)</td>
</tr>
</tbody>
</table>

The breach length is the most important parameter that requires the greatest consideration by the user because this parameter determines the distance the algorithm will search away from the depression. Increasing this value also substantially increases
processing time but has relatively little impact on the final depressionless DEM for reasonably large values. This breaching algorithm has been implemented in Whitebox Geospatial Analysis Tools GIS, which is an open-source software package for performing spatial analysis operations in environmental science. It can be found at: (http://www.uoguelph.ca/~hydrogeo/Whitebox/download.shtml). The following section describes the study area and methods for the applications of the novel algorithm used for this study.

3.4.0 Methods

3.4.1 Study Areas

The novel algorithm was tested on two sites in Southern Ontario, Rondeau Basin and Prince Edward County (PEC). The Rondeau Basin and Prince Edward County watershed terrain are similar to other agricultural landscapes in the area.

3.4.1.1 Rondeau Basin

The Rondeau Basin (42°17′N 81°52′W) is located in Southwestern Ontario along the coast of Lake Erie as seen in Figure 3.2. The headwaters of this watershed begin from the Blenheim moraine. The soils are considered to be well-drained gravelly loam with moderately high organic matter and high infiltration rates (Walker and Richards, 1994). Towards the centre of the watershed, the topography is gently undulating, made up of mostly silty clay loam. Closer to Lake Erie, the topography is level to gently sloping, with a soil texture that is sandy loam (Walker and Richards, 1994). The drainage system for the area has been established over 135 years, involving approximately 4,000 municipal drains, which include open ditches.
Deep gullies drain the landscape from the north, with the flow eventually joining the stream network in the southern part of the watershed. The major land use in this region is agricultural farming, with wheat and corn being the dominant crop type. Other vegetation includes hedgerows and woodlots separating farmland, as well as deciduous forest. There are some urban areas within and surrounding the watershed, such as Blenheim and Chatham-Kent, and some settlement along the edge of Lake Erie. Urban land use comprises less than 8% of the dataset.
3.4.1.2 Prince Edward County

Prince Edward County (44°00′N 77°15′W) is located in Southwestern Ontario. Most of the county is completely surrounded by Lake Ontario (Figure 3.3).

Figure 3.3 Prince Edward County Study Site

The topography of the Prince Edward County Site is undulating with a gentle rise in elevation from northeast to southwest. There are approximately 800 km$^2$ of shoreline with the northern and eastern shorelines containing rocky bluffs, while the western shores are comprised of inlets and sandbars (VQA, 2011). The soil type is characterized as reddish-brown clay to sandy loam overlaying the Trenton limestone plateau. The major land use practice is agricultural farming, with grapes being the dominant crop type (VQA, 2011). The prominent soil type allows for deep penetration and good drainage for
mature grape vines. Woodlots are abundant, and there are some hedgerows separating farmland. Small streams collect runoff that finds outlets in numerous ponds and bays.

3.4.2 Data collection.

3.4.2.1 Rondeau Bay dataset collection: Sensors and Equipment

Optimal Geomatics collected LiDAR data for the Rondeau Bay area in the Spring of 2008. These data were collected over two days (May 4\textsuperscript{th} and 5\textsuperscript{th}) in which weather conditions were optimal for LiDAR data acquisition. More specifically, the dataset was acquired on days where there was no rain, snow, fog, low clouds, or smoke.

A Cessna 208B Grand Caravan, N704MD aircraft performed the LiDAR survey (Optimal geomatics, 2008). Altogether, 77 flightlines were recorded, with 13 flown from an overlapping flight plane to ensure coverage for particular areas within the Rondeau Bay area. All the data were collected in four flights and all but two flightlines were used in the delivered database.

Mounted on the aircraft was an Optech Airborne Laser Terrain Mapper (ALTM) 3100. Optech 3100 ALTM offers area coverage of up to 50 km\(^2\)/hr at a rate of 100 KHz (100,000 measurements/sec) and a scan width of 50 degrees (Optech, 2006). The data produced has a vertical resolution of five to ten centimeters and a horizontal resolution of about 15 centimeters. The Optech ALTM 3100 is a multiple return system, recording up to four returns in addition to the intensity value for each pulse, and it has an operational altitude of up to 3,000 m. Applanix 510 IMU and Novatel GPS are the other technologies that completed the LiDAR system, which had a data collection rate of 10 Hz. Two Leica System 500 GPS base units were located within the Rondeau Bay area to support the precise positioning of the LiDAR sensor.
3.4.2.2 Rondeau Bay Post-Processing and final delivery format.

Optimal Geomatics carried out the following analyses on the LiDAR point cloud data. DASHMAP, a LiDAR processing software package from Optech, was used to process the point cloud, flight logs, raw air and ground GPS files. The airborne GPS data were processed from two base stations using POSGPS from Applanix, Inc. POSGPS combines the airborne GPS trajectories processed from multiple stations into a single solution and weighs the closest stations heavily. Combining the solutions provides redundancy, quality assurance (QA), quality control (QC), and gives the most accurate solution when flying long distances between base stations. These datasets were then used to process the inertial data that used the software PosProc form Applanix, Inc. The software produces a Smooth Best Estimate of Trajectory (SBET) by analyzing both the GPS trajectory and the information collected from the IMU. DASHmap uses the SBET to create the point cloud data in the LAS file format.

Each flightline was processed independently, as this provided the data analyst with the ability to ensure quality control of the overlap between lines. A Root-Mean Square Error (RMSE) value was computed for the LiDAR DEM that was created from the point cloud data. A RMSE could not have been reported for the point cloud data due to the nature of the area and the indefinite spot of each individual LiDAR point. However, the accuracy of the performance of the LiDAR system was documented in the quality control report. Quality control of the airborne GPS units was measured by comparing multiple solutions from the base stations, which had an agreement of less than five centimeters horizontally. The IMU sensor combined the post-processed GPS data with the raw inertial data and would not allow the IMU solution to be of a lower accuracy than the
provided input from the GPS solution. The altitude of the sensor used in this project was 1200 – 1500 m above the ground, allowing it to estimate a spot size of 34- 42 cm for the project. Each point was located within 17-21 cm of the center of the spot. The horizontal accuracy of the LiDAR system was 1/2000.

3.4.2.3 Prince Edward County (PEC) dataset collection: Sensors and Equipment

Aero-Photo Inc. collected the LiDAR data for the PEC area in May and December of 2009. These data were collected over seven nonconsecutive days in which weather conditions were optimal for LiDAR data acquisition (Aero-Photo, 2009). There were no weather conditions documented in the final report from Aero-Photo, Inc, however it was noted that there were no particular problems encountered during collection.

A Cessna 421 twin turbo piston aircraft and a piper Aztek twin piston aircraft performed the LiDAR survey. PEC was divided into 12 sections, and 14 flight missions were undertaken. Altogether, 170 flightlines with a side overlap of 30% were collected.

Mounted on the Cessna 421 aircraft was an Optech Airborne Laser Terrain Mapper (ALTM) 3100, while the Aztek twin piston was equipped with an Optech Gemini 167. While the LiDAR systems used to acquire the data have different specifications, both were set-up with the data acquisition parameters. Table 3.2 summarizes these specifications for both sensors.

Table 3.3 summarizes the 14 flight missions and their dates for the PEC site. Pos/AV/DG 510 with Sagem IMU and Trimble L1/L2 geodetic GPS are the other technologies that completed the LiDAR system used in PEC.
Table 3.2 LiDAR acquisition parameters for PEC Site.

<table>
<thead>
<tr>
<th>LiDAR Acquisition Parameters</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(PRF) Pulse Repetition Frequency (KHz)</td>
<td>100</td>
</tr>
<tr>
<td>Scan Frequency (Hz)</td>
<td>35</td>
</tr>
<tr>
<td>Scan Angle (+/-)</td>
<td>18</td>
</tr>
<tr>
<td>Field of View (FOV) (degrees)</td>
<td>36</td>
</tr>
<tr>
<td>Flight Height AGL (feet)</td>
<td>6234</td>
</tr>
<tr>
<td>Flight Speed (kts)</td>
<td>120</td>
</tr>
<tr>
<td>Point per square meter</td>
<td>1.3</td>
</tr>
<tr>
<td>Foot print (meters)</td>
<td>0.48</td>
</tr>
<tr>
<td>Swath width (meters)</td>
<td>1235</td>
</tr>
<tr>
<td>Swath Overlap (%)</td>
<td>30</td>
</tr>
<tr>
<td>Swath Overlap (meters)</td>
<td>370</td>
</tr>
<tr>
<td>Flight line spacing</td>
<td>864</td>
</tr>
<tr>
<td>Grid across flight path (meters)</td>
<td>0.8</td>
</tr>
<tr>
<td>Grid along flight path (meters)</td>
<td>0.8</td>
</tr>
</tbody>
</table>

Two Leica SR530 L1/L2 geodetic GPS receivers installed at the Kingston airport and in the Municipality of Picton were tied to the Natural Resources Canada (NRCAN) point KINGS of the Canadian Active Control System and to Madoc and Belleville of the Leica SmartNet continuous GPS receiver Stations. There were a total of ten control points for the area.
Table 3.3 Flight mission dates for the PEC site (Aero-Photo, 2009)

<table>
<thead>
<tr>
<th>Aircrafts</th>
<th>Dates</th>
<th>Number of Flight Missions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cessna 421 twin turbo piston aircraft</td>
<td>May 18, 2009</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>May 19, 2009</td>
<td>2</td>
</tr>
<tr>
<td>Aztek twin piston aircraft</td>
<td>May 18&lt;sup&gt;th&lt;/sup&gt; 2009,</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>May 19&lt;sup&gt;th&lt;/sup&gt; 2009</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>May 20&lt;sup&gt;th&lt;/sup&gt; 2009</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>May 21&lt;sup&gt;st&lt;/sup&gt; 2009</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>December 2&lt;sup&gt;nd&lt;/sup&gt; 2009</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>December 6&lt;sup&gt;th&lt;/sup&gt; 2009</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>December 18&lt;sup&gt;th&lt;/sup&gt; 2009</td>
<td>1</td>
</tr>
</tbody>
</table>

3.4.2.4 PEC Post-Processing, and final delivery format.

The airborne GPS data were processed from the base stations using POSGPS from Applanix, Inc. Similar to the pre-processing techniques used with the Rondeau Bay data, POSGPS combined the airborne GPS trajectories processed from multiple stations into a single solution and weighed the closest stations the most heavily. These trajectories were used to process the inertial data, using the software POSProc form Applanix, Inc. This software produces a smooth best estimate of trajectory (SBET) by analyzing both the GPS trajectory and the information collected from the IMU. DASHmap uses the SBET to
create the point cloud data in the LAS file format. Table 3.4 is an accuracy assessment of the LiDAR point cloud after comparing it to 1552 measured ground points.

Table 3.4. Accuracy assessment for LiDAR point cloud at the PEC site

<table>
<thead>
<tr>
<th>Statistics for LiDAR point cloud</th>
<th>Meters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum</td>
<td>-0.2494</td>
</tr>
<tr>
<td>Maximum</td>
<td>0.2092</td>
</tr>
<tr>
<td>Average</td>
<td>0.0146</td>
</tr>
<tr>
<td>Root mean square</td>
<td>0.1182</td>
</tr>
<tr>
<td>67%</td>
<td>0.1274</td>
</tr>
<tr>
<td>95%</td>
<td>0.2076</td>
</tr>
</tbody>
</table>

3.4.2.5 The Ontario Road Network Dataset

The Ontario Road Network (ORN) was used to locate roads in each study site so that each road underpass predicted by the algorithm was extracted for the analysis. The Ontario Ministry of Natural Resources initially created the ORN in 2004, which has since been updated on a monthly basis. The ORN is a province-wide database covering more than 235,000 kilometers of municipal roads, provincial highways and resource access roads across Ontario. Attributes such as street names and road numbers, address information, road classifications, surface types and speed limits are all contained in the ORN. The ORN has a horizontal resolution of 10 meters or less and is in vector format. Total road lengths for the Rondeau Bay area and the PEC were of approximately 315 km and 236 km respectively.
3.4.2.6 Reference data

The South Western Ontario Orthophotography Project (SWOOP) 2006 imagery was used to digitize road underpasses for the Rondeau Basin and the 2008 Digital Raster Acquisition Project for the East (DRAPE) imagery was used for the PEC site. SWOOP images were projected in UTM, zone 17 Datum NAD 83, and had a resolution of 0.30 m. DRAPE images were also projected in UTM, zone 18, Datum NAD 83 with a resolution of 0.2 m.

Colour, shape, and association (Avery and Berlin, 1992) were the factors of recognition used to identify culverts. For example, some culverts are made of galvanized steel, and have ends that usually protrude out of the side of the road. The ends of the culverts are usually easy to identify in aerial photographs because of the reflective quality of steel. Other culverts are made of concrete and have a distinctive shape associated with them. Where off-terrain objects obscured culverts, association was used, meaning that where there was evidence of drainage patterns such as dendritic drainage patterns on one side of the road and continued on the other side, a culvert was assumed to have been present. Bridges were identified using a similar set of criteria.

Due to time constraints 80% of the Rondeau Bay area and 75% of Prince Edward County were surveyed through aerial imagery before field work was completed. These image-derived road underpass locations were then validated using ground truthing techniques. Samples of 440 underpasses identified from the imagery from the Rondeau site (67% of the total underpasses identified in the images) and of 290 underpasses from the PEC site (64% of the total underpasses identified in the images) were verified in the field in the Fall of 2010. Those road underpasses that were difficult to interpret in the
aerial imagery were first verified in the field. Next the underpasses that were apparent in the aerial imagery were verified. Finally the remaining area that was surveyed in the aerial imagery was verified in the field. There were culverts that existed in the field that were overlooked in the aerial photographs mainly because they were obscured by off-terrain objects such as trees and buildings. In total, 62 additional underpasses were identified in the field for the Rondeau site that were not identified by aerial photographs, and 163 additional underpasses were found in the field for the PEC. These culverts that were identified in the field were later added to those that were located by the imagery reference data and the remaining area that was not surveyed through aerial imagery was completed.

3.4.3 Data Processing

The Rondeau Bay area and PEC were selected as study sites in part because of the availability of the LiDAR data for these regions. Raw point cloud LiDAR data sets for Rondeau Bay were obtained from the Ontario Ministry of Agriculture, Food and Rural Affairs (OMAFRA); PEC data were obtained from the Ontario Ministry of Natural Resources, Canada. The processing techniques used for these datasets are described in more detail in the following sections.

3.4.3.1 Data interpolation and Sub-site aggregation

Whitebox Geospatial Analysis Tools (GAT) software was used to interpolate both the Rondeau Bay area and the PEC LiDAR point cloud data. Interpolation involves estimating unknown elevation values of a DEM based on known elevation values. Each site was interpolated using a nearest-neighbour interpolation routine. The nearest-neighbour interpolation routine assigns the value of an unknown point with the same
known elevation value of the closest sampled location (Akkala et al, 2010). The nearest-neighbour interpolation routine was deemed most appropriate for this research because of the high point density associated with LiDAR point cloud data (Akkala et al. 2010). The nearest neighbor interpolation routine is an efficient method for large datasets such as the datasets that were used in this study and therefore this simple interpolation approach is necessary when compared to other computational heavy interpolation routines. The nearest neighbor interpolation routine is a very efficient method for large datasets such as the datasets that were used in this study and therefore this simple interpolation approach is necessary when compared to other computational heavy interpolation routines. The LiDAR point cloud data were interpolated to one meter so that features such as ditches and road underpasses were apparent in the DEM. Off-terrain objects, such as buildings and trees, were removed from the LiDAR DEMs using an algorithm developed by Lindsay and Dhun (In Review). The resulting surface is what is commonly referred to as a bare-earth DEM. The individual tiles were then joined together to create multiple sub-site DEMs for each study area. Combining both sites, there were a total of 21 sub-sites, forming ten DEMs for the Rondeau Bay area and 11 DEMs for the PEC. Sub-sites were necessary so that the data were more manageable and so that urban areas were excluded.

3.4.3.2 Road underpass extraction

The novel breaching algorithm presented in this thesis is capable of resolving most of the depressions in a LiDAR DEM. However, for this analysis, depressions caused by road underpasses were the main focus. These specific breaches had to be extracted from the output DEM using both ArcGIS and Whitebox GAT.
ArcGIS was used to convert and clip both the ORN and the Observed Road Underpass (ORU) reference data to the raster format for each sub-site. This was done to ensure that all data were in a single format. Raster formats are efficient for extracting information when arrays are of the same cell size. The clipped and rasterized ORN and ORU data were then exported to Whitebox GAT for further analysis.

DEM differencing was performed in Whitebox GAT. Differencing is the process of subtracting the breached DEM image from the Original DEM on a pixel-by-pixel basis. As previously noted, LiDAR DEMs contain vast amounts of small depressions. For example, for this study an area of 20 km$^2$ can contain as many as 20,047,904 depressions. A threshold value of greater than or equal to 0.25 meters was applied to the differenced DEM, meaning that all breaches that were less than 0.25 m were regarded as blunders. This value was deemed appropriate because the LiDAR DEMs had a vertical precision of +/-0.25 meters. By applying this threshold value, the amount of breaches that needed to be processed were substantially reduced from anywhere between 80-92%. Road underpass breachlines (referred to as Predicted Road Underpasses (PRU) henceforth) were then extracted based on three criteria:

1) PRU had to be greater than or equal to 0.25cm;
2) PRU had to intersect the road data;
3) PRU had to be longer than or equal to the length of the road (usually greater than three pixels long).

The locations of the PRU were compared to the locations of the ORU to evaluate whether the PRU was successful or unsuccessful. A PRU was considered a success if a
PRU overlaid the ORU. If a PRU did not overlay an ORU was along the same flowpath, it was also considered a success.

3.4.4 Accuracy Assessment

To test the ability of the algorithm to extract road underpasses, errors of omission (EOO) and errors of commission (EOC) were calculated. EOO are exclusion errors. For this study, an EOO referred to an observed road underpass (i.e. it is known to exist in the landscape) that had not been breached in the DEM that had been processed using the algorithm. EOC are inclusion errors. An EOC was recorded in this study when the drainage enforcement algorithm breached across a road, implying a culvert or bridge location, but there was no evidence of the underpass in either the aerial imagery or the ground inspection. In these situations, it was usually the case that the algorithm should have enforced a downward slope along the roadside ditch that the depression cell occurred in, rather than breaching across the road. The producer’s accuracy and user’s accuracy were also used to test the accuracy of the algorithm to extract road underpasses.

The producer’s accuracy corresponds to the EOO, which refers to the probability that a road underpass on the ground is classified as a road underpass in the DEM. For example, a producer’s accuracy of 80% suggests that the algorithm identified 80% of all the road underpasses in the reference data. The user’s accuracy corresponds to the EOC, which refers to the probability a road underpass in the DEM represents a road underpass on the ground. For example a user’s accuracy of 70% suggests that 70% of road underpasses identified by the algorithm are road underpasses on the ground.
3.4.45 Ditch Gradient Calculation

The main purpose of roadside ditches is to collect road surface runoff and drain it away from the road. Ditches are designed based on the amount of runoff they receive. Therefore, the ditch gradient tends to follow the road gradient, where flat roads tend to be adjacent to gently sloping ditches, while steep roads tend to have steep ditches associated with them. Road gradient can therefore be used as a surrogate to calculate ditch gradient as ditch gradient is difficult to accurately calculate from a DEM. To test the hypothesis that there was an increase in erroneous breaches (i.e. EOC) across flat roads due to the gentler ditch channel gradients, an automated technique using road data and the DEM were used to calculate ditch slope. These datasets were not normally distributed; therefore both variables were transformed by the natural logarithm. This technique was applied because in flatter parts of the landscape, the algorithm must search further to find target cells for breaching. If the search distance is greater than the width of the road, there is the potential that a target cell may be located on the other side of the road, rather than along the same side of the road.

Linear regression analysis was performed between the dependent variable (frequency of EOC) and the independent variable (road slope) to test if the variability of frequency of EOC was a result of the road slope. Road gradients were divided into nine classes, the lowest class being 0.2 degrees with additional classes increasing by 0.2 increments to 2.70 degrees. 0.2 degree bin size was fine enough to capture the detail of the frequency distribution while being coarse enough that each bin is appropriately populated with data.
3.5 Results

A visual comparison of the effects of filling and breaching on a DEM (Figure 3.4) will first be discussed, followed by a quantitative analysis of the breaching algorithm.

*Figure 3.4 a1, b1) The original DEM prior to pre-processing; a2, b2) the filled DEM; a3, b3) The breached DEM*
Figure 3.4 a1 and 3.4 b1 shows the original DEM prior to pre-processing. LiDAR returns cannot distinguish roads and bridges from ground returns. The road acts as an obstruction to flow, creating a digital dam across the entire watershed. Figure 3.4 a2 and b2 illustrates the result of applying filling to the LiDAR DEM. The image depicts the severe damage caused by filling, particularly upstream of the road bank, losing the valuable detail needed for accurate flow modelling behind the apparent dam feature. Figure 3.4 a3 and b3 depicts the results of applying the breaching algorithm to the DEM. Notice that the fine detail in the stream channel is preserved along with a breach across the road to enforce flow throughout the watershed, both of which help to create realistic flowpaths. Figure 3.5 shows the results of the flow simulations on the original, filled, and breached DEMs. Figure 3.5 a1 and b1 illustrates the truncated flow-lines and altered flow directions that are due to the presence of depressions in the DEM. Figure 3.5 a2 and b2 shows the road impeding simulated drainage patterns from reaching an outlet by creating a digital dam behind the road which disconnects flow patterns and causes interior flooding within a section of the watershed. The confined water may eventually spill over the road in the wrong location or may exit through a different part of the watershed (MacMillan, 2003; Poppenga et al., 2009). Figure 3.5 a3 and b3 shows the flow patterns from the breached DEM. Notice the fully connected drainage network as a result of breaching the road. The simulated drainage patterns are such that they are consistent with actual surface flow. From this qualitative analysis it can be stated that although filling is the most commonly used pre-processing algorithm in hydrological applications, it is by no means appropriate for these LiDAR datasets.
Figure 3.5. Effects of flow-accumulations on a1, b1) The original DEM; a2, b2) The filled DEM; a3, b3) The breached DEM

Filling severely impacts the DEM by generalizing the terrain surface and consequentially the derived hydrologically significant terrain attributes. In addition,
Figure 3.4 a3 and b3 and Figure 3.5 a3 and b3 illustrates that breaching is more suited for road damming depressions. Table 3.5 presents the total depressions within each site along with those attributed to road underpasses.

*Table 3.5. Breakdown of depressions for each site*

<table>
<thead>
<tr>
<th>Sites</th>
<th>Total Depressions</th>
<th>Depressions due to Road underpasses</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rondeau</td>
<td>18,662,983</td>
<td>653</td>
<td>18,662,330</td>
</tr>
<tr>
<td>PEC</td>
<td>37,227,040</td>
<td>353</td>
<td>37,226,678</td>
</tr>
</tbody>
</table>

A majority of the depressions stem from the low vertical error of the LiDAR DEM. That is, LiDAR DEMs capture small scale surface variability more completely. Fine-resolution DEMs tend to have more small and shallow depressions when compared to traditionally acquired DEMs because of the greater detail associated with these finer resolution datasets. Although depressions due to the artificial damming of surface flowpaths account for a small portion (0.0035%) of the total depressions, they cause considerable damage in hydrological applications when left unattended. Table 3.6 summarizes the results of the accuracy assessment of the algorithm’s ability to extract road underpasses for the two sites. For the Rondeau site, the EOO and EOC were 17% and 58% respectively. The producer’s accuracy was 83% while the user’s accuracy was 42% for the Rondeau site. In other words, 83% of road underpasses were correctly classified while 42% of the PRU have a probability of actually representing an observed overpass. For the PEC site, the EOO and EOC were 22% and 66.5% respectively.
The producer’s accuracy was 78% while the user’s accuracy was 33.5% for the PEC site. This means that 78% of the culverts were correctly classified while 36% of the PRU have a probability of actually representing an observed overpass.

Table 3.6 Accuracy assessments for both the PEC and Rondeau Basin sites.

<table>
<thead>
<tr>
<th>Accuracy Measures</th>
<th>Prince Edward County (%)</th>
<th>Rondeau Basin (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Producer's Accuracy</td>
<td>78.5</td>
<td>83</td>
</tr>
<tr>
<td>User's Accuracy</td>
<td>33.5</td>
<td>42</td>
</tr>
<tr>
<td>Error of commission</td>
<td>66.5</td>
<td>58</td>
</tr>
<tr>
<td>Error of omission</td>
<td>21.5</td>
<td>17</td>
</tr>
</tbody>
</table>

There were fewer occurrences of EOOs when compared to EOCs for both sites, suggesting that the presented algorithm is a powerful tool for determining culvert location as it works well across a range of topographic settings. Based on the high frequency of the EOCs for both sites, the algorithm breaches roads liberally. The user’s accuracy scores were low for each site 33.5% and 42% for the Rondeau Bay area and PEC respectively. That is, for the Rondeau Bay area 33.5% of the PRU were apparent in the reference data while for the PEC, 42% of the road underpasses predicted by the algorithm were apparent in the reference data.

The results from the calculations of the road slope relative to the frequency of EOC are presented in Figures 3.6 and 3.7; illustrating that 80% of EOC occurrences had a slope of less than 0.40 degrees at the Rondeau site, and 66% at the PEC site.

The linear regression analysis in log-transformed variables revealed that the degree to which the algorithm displays the liberal behaviour when breaching roads is highly
dependent on the road gradient, with EOCs being more prevalent along flatter roads. This means that as the road slope decreases, the occurrences of EOCs were found to increase in frequency. Table 3.7 displays the $R^2$ values, $p$-values and slope constants for the Rondeau Bay area and PEC sites.

![Figure 3.6 PEC road slope relative to EOC's](image1)

![Figure 3.7 Rondeau road slope relative to EOC's](image2)
Table 3.7. Statistics associated with EOC frequency and road slope

<table>
<thead>
<tr>
<th>Site</th>
<th>R2</th>
<th>p-value</th>
<th>Significance level</th>
<th>Slope</th>
<th>Intercept</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rondeau Bay</td>
<td>0.90</td>
<td>0.001</td>
<td>0.95</td>
<td>-1.63</td>
<td>2.57</td>
</tr>
<tr>
<td>PEC</td>
<td>0.51</td>
<td>0.03</td>
<td>0.95</td>
<td>-0.88</td>
<td>3.15</td>
</tr>
</tbody>
</table>

Correctly identified culverts for both Rondeau and PEC were aligned with the reference data, meaning that the algorithm provided accurate positioning of the successfully identified culverts. In fact, the algorithm identified culverts in both sites that were overlooked when creating the reference data. Furthermore, the algorithm also enforced ditches along roads where these features were not completely captured by the LiDAR sensor and enforced flowpaths along stream networks.

It was extremely difficult to determine the proportion of roadside-ditch cells for which the algorithm was able to correctly resolve the drainage pattern. Therefore, a qualitative analysis was also undertaken to increase understanding of the causes of the flow enforcement along road-side ditches. Figure 3.8 shows the before and after effects of flow enforcement through ditches caused by the breaching algorithm. Qualitatively speaking, the algorithm performed well in enforcing flow along ditches in the DEM. Figure 3.8 a – d depicts the flow accumulation before the breaching algorithm was applied, while 3.8 g - h shows the flow accumulation after application of the algorithm from various sites within the Rondeau Bay and PEC areas. Before the application of the algorithm, the DEMs showed undefined flowpaths; after the algorithm was applied to the DEM it is evident that the breached DEMs channelize flow along these features. The observations from this analysis will be discussed in the following section.
Figure 3.8 Ditch enforcement along roads in both Rondeau Bay area and PEC sites
3.6 Discussion

Ditches are narrow linear features that are used to channelize surface flow to a culvert. Because the LiDAR data are collected in a random manner, there are some instances where the LiDAR system could not capture the entire ditch. As a result, these narrow features are usually segmented in the LiDAR DEM instead of being continuous linear features. Where ditches are continuous they are usually flat, meaning that the elevation values associated with them usually do not have a downslope neighbor, resulting in artifact depressions in the DEM. Whatever the flow source, flow enforcement along ditches is a necessary step when hydrologically conditioning DEMs.

The breaching method outlined in this research managed to define ditches in such a way that they follow the shape of the road. This means that where roads were linear, ditches were linear, and where there was a bend in the road, ditches followed this shape. The novel algorithm transformed a noisily represented ditch in the DEM into a continuous flowpath that followed the original channel with high fidelity. This algorithm is well suited for enforcing flow along ditches because it lowers dam cells along these narrow linear features. Unlike the approach adopted by Duke et al. (2006), the approach in this study accommodated meander bends. That is, instead of enforcing a straight pathway between two points along a stream network, the algorithm enforced a flowpath that followed the channel morphology.

Manual incorporation of ditches and road underpasses seems to be the most labour intensive step when pre-processing LiDAR DEMs, although the time and cost associated with this step are dependent on both the size of the DEM, and the calibration of the
LiDAR system. Martin Flood (2007) defined six main steps for the process of manual editing of a DEM. These steps are outlined below.

1) Visualization of each tile to initially assess QA/QC of classification results;
2) Manual editing/cleanup to generate a bare-earth DEM;
3) Generation of model key points;
4) Review of model key points;
5) Creation of DEM contours, and
6) Final QA/QC assessment of contours.

Flood (2007) estimated the best, average and worst case scenario editing times range from 30, 55, and 220 minutes respectively for each tile (size of 1 km²). Worst-case scenarios are substantially influenced by the need to clarify ambiguous points from ancillary data, the removal of off-terrain objects, and additional final QA/QC times.

Although the breaching algorithm presented in this paper substantially decreased the pre-processing time (approximately 30 minutes for 20 km²) associated with the LiDAR-derived DEMs when compared to traditional techniques, EOOs and EOCs did still occur. Errors of omission were substantially less when compared to errors of commission for both sites.

The EOOs for the Rondeau site were 18%. Stated differently, this means that 18% of the ORUs were not identified by the algorithm. Upon further analysis, the characteristics of the omitted road underpasses were similar throughout the Rondeau Bay area, where most EOO’s were actually dredged channels. The terrain for this site has been altered to accommodate farming practices, making these channels particularly flat when compared to natural river channels. The intended purpose for dredged channels is
to collect runoff from farming fields by providing faster drainage and reducing the risk of flooding in the local area. By recalling that the criteria for defining a target cell is a cell that is lower than the depression with a downslope neighbor, it is clear that dredged channels do not fit these criteria and as a result do not breach the road properly because of the little to no slope associated with them.

The PEC site had an EOO value of 21%. Some EOOs occurred around residential areas in the site. This is expected because unlike agricultural landscapes, residential areas tend to have smaller gradients associated with them, and result in no target cell being associated with these types of landscapes. Unlike the Rondeau Bay area, it was difficult to find a commonality associated with EOOs. For the most part, the unidentified channels were natural but the algorithm could not recognize a target cell across the road. There were many situations where one side of the road contained off-terrain objects while the other contained dense vegetative canopy. For example, many farmers’ houses were situated on one side of the road while their wheat crops were situated on the other. EOOs occurred in these cases because most vegetation were not removed and as a result, the algorithm had difficulty finding a cell that was lower than the depression with a downslope neighbor due to the higher elevation values associated with the vegetation.

LiDAR systems have difficulty reaching the ground where vegetation cover is greater than 80% resulting in clusters of relatively high elevation points. These clusters cause a high cost associated with breaching the road in areas where road underpasses were surrounded by them. There may be some instances where the laser is able to reach the ground and other cases where the laser may only reach a branch. As a result, the affected area in the landscape would tend to have varying elevation values due to the
dense canopy. Other cases that caused EOOs occurred where vegetation cover such as woodlots may have obscured ditches or culvert openings, and were not completely removed during the ground filtering stage. In other words, ditches and road underpasses were not completely captured by the laser scanner because of the tree canopy, resulting in these errors of omission.

For the Rondeau site, the EOC was 62.5%. This value may seem high, but the EOCs occurred predominately on flat roads. Ditches along flat roads tend to have to be flat and as a result, a target cell for the ditch depression cell may lie across the road. The distance to solve a ditch depression cell correctly may be far from the depression, implying a need to search a much greater distance to locate a lower cell to potentially breach towards. Therefore, the most practical or efficient pathway for the algorithm may be to breach across the road. Of the 62.5% of EOCs that occurred, 80% were found to be present on flat roads.

For the road underpasses that were successfully predicted, the algorithm identified the culvert location, with the majority of predicted culverts aligning with observed culvert location. The landscapes in which successfully predicted culverts did not align with the observed culverts occurred at the outskirts of residential areas and around most woodlots in both sites. The algorithm also was able to breach culverts that were overlooked when creating the reference data. An important advantage that this novel algorithm has over traditional techniques is, unlike Duke et al. (2006), the algorithm presented in this paper operates without the need to explicitly locate road or ditch networks. That is, it performed well at locating road underpasses and did not need ancillary data. Furthermore, the algorithm enforced ditches and stream flowpaths that
would have been erased if filling were used to preprocess this data. Apart from breaching culverts, the algorithm performed well at enforcing both ditch and streams within the DEM. Past studies show that culverts and ditches greatly influence intra-watershed transport, the size and shape of the watershed, and runoff response time, so correct enforcement of ditch and stream flow is essential in a pre-processing algorithm (Duke et al., 2006; Jones et al., 2000; La Marche and Lettenmaier, 2001).

The breaching algorithm presented in this research correctly identified culverts within both sites, and most errors that did occur were caused due to the flat or gentle gradients in some parts of the study sites. As a result, this breaching algorithm is best suited for landscapes that do have downslope gradients and for areas.

This paper presented a new technique for automatically detecting culverts and ditches in a DEM. This new breaching algorithm does this by finding a depression, searching and selecting a path that would cause the least modification of the DEM, and then breaching that path. This technique is successful in mapping culverts in human-modified areas where there is little to no gradient associated with the landscape. The algorithm did not perform as well when enforcing flow in ditches as it did when finding road underpasses. This automated method can replace current manual methods that have been proven to be time-consuming and inefficient.

3.7 Conclusion

Ditches and road underpasses are hydrologically significant features in human-modified landscapes. In LiDAR DEMs, ditches are usually associated with artifact depressions because they are often long, narrow and flat features. Road underpass are
generally not captured in the LiDAR data, resulting in the loss of a water flowpath in the DEM that in reality would flow under the road through a culvert. An automated technique for incorporating ditches and road underpasses was presented in this paper with the following main findings:

1. The new breaching algorithm presented in this paper is a novel technique for incorporating ditches and road underpasses for hydrological applications. This algorithm incorporates ditches and road underpasses by finding all depressions in a DEM, searching and selecting a path that would cause the least modification of the DEM, and then breaching that path. This method has proven to be successful in landscapes where the majority of predicted road underpasses coincide with the observed underpass locations.

2. A qualitative analysis showed that flow enforcement along roadside ditches and major streams were captured, when they would otherwise have been obliterated if depression filling were applied to the DEM.

3. The high value of producer’s accuracy for both sites shows that the algorithm is a valuable tool for predicting culvert location in different landscapes.

4. Errors occurred when these flowpaths were heavily altered (such as with dredged channels) and on flat roads due to the lack of gradients. As a result, this algorithm is not suited for such landscapes.
4.0 CONCLUSION

Many landscapes have experienced substantial modification due to human activities, which has had direct impacts on the hydrological regimes within watersheds. Roads and other embankments block surface flow-paths, and ditches and other conduits assist in redirecting water. Ditches and road underpasses alter the natural drainage patterns of a landscape by altering the magnitude, timing and quality of runoff, yet these influences are often not accounted for when modelling surface drainage patterns in human-modified landscapes due to the limitations of the available data and flow-routing algorithms. This is due to the fact that coarse-resolution DEMs cannot capture these anthropogenic features and therefore cannot account for the impacts ditches and road underpasses have on the local hydrology and manual DEM correction can be impractical in many applications. Therefore, to avoid these problems, researchers have chosen to study relatively unmodified sites, or disregard the effects of anthropogenic features on the landscape altogether. LiDAR DEMs, however, are able to capture these anthropogenic features due to the fine spatial resolution of these images. Although LiDAR data is certainly capable of recognizing significant linear alterations in the landscape such as ditches and road underpasses, when using these datasets in hydrological models, these human-made landscape features act as local sinks, creating digital dams, which incorrectly represent the influence of these linear features.

This thesis has aimed to improve the modelling of surface drainage pathways in landscapes that have been heavily altered by anthropogenic infrastructure. This was done by developing and testing a novel automated technique that uses the topographic information contained in LiDAR DEMs to overcome the limitations of the current
methods. For example, the method employed by Duke et al. (2006) involved explicitly identifying roads and linear features (with the use of ancillary data) in the DEM prior to extracting them. Road underpass locations are often not available for most sites, and where the road underpass data is available, the data may not coincide perfectly with the road network as depicted in the LiDAR DEM. As a result, the method employed by Duke et al. (2006) which explicitly needed ancillary data to extract linear anthropogenic features is not always reliable. The algorithm presented in this thesis differs from Duke et al. (2006) as it is not dependent on ancillary data such as road networks and road underpasses, meaning that the breaching algorithm does need to explicitly know where a road is to breach the road at the underpass location.

Objective 1 of this thesis was to develop an algorithm for enforcing DEM modelled flow patterns along roadside ditches and drainage canals. Currently, human operators search through the DEMs and interpret culvert location through association, by using proximity to stream networks to deduce the location of anthropogenic features. Another method that is commonly used involves overlaying stream and road vector data onto the elevation data to find the intersection of roads with flowpaths, and then manually removing the elevation values associated with the roads. The above methods are all manual techniques, which involve human subjectivity. Manual LiDAR DEM preprocessing for hydrologic application is then highly dependent on the technician’s understanding of the data and familiarity with the landscapes. The algorithm presented in this thesis involves an automated technique used to incorporate human-made features. A qualitative analysis was undertaken in this work to assess how well the algorithm was able to enforce flow in ditches. While the algorithm performed well at enforcing flow
along steep roadside ditches, it did not define ditches along flat roads (having a slope value less than 0.60 degrees) as accurately. This was made evident as most EOCs were associated with flat roadside ditches.

Objective 2 of this research involved creating a predictive model for identifying potential road underpass locations in a LiDAR DEM as well as a technique for breaching the artificial dams created by the roads at these underpass sites. The algorithm first predicted and identified underpass locations based on a least-cost pathway analysis, and then breached the elevation data between the depression and the target cell. There were two criteria used to identify target cells:

1) The target cell must be lower than the depression and,

2) The target cell also must be connected to a downslope neighbour.

In cases where there is more than one target cell, the algorithm chooses a path that will modify the DEM the least. This means that the least-cost pathway is determined by the degree of enforcement (i.e. the amount of DEM breaching) that is needed to connect the depression cell to a distant downslope cell. Once the pathway with the least cost is identified, the cells are breached to allow for a continuous flow path, so that cells along the path will be lowered to create a monotonically descending flowpath from a depression to the target cell. This approach for resolving the artificial damming at road underpasses is quite advantageous because, unlike minimum distance breach algorithms that favor straight breach paths, this approach usually breaches along the poorly represented ditches. This means that the algorithm considers drainage channels that are noisily represented and then defines a continuous flowpath for these channels that follows the original channel. The algorithm does all the above without the need for ancillary data.
Two case studies were presented in Chapter 3 of this thesis. Rondeau Bay and Prince Edward County are agricultural landscapes located in southwestern Ontario where culverts and ditches are important contributors for routing overland flow. The algorithm defined in this research performed well in these landscapes. The culverts that were successfully delineated by the model coincided with the exact location of the observed road underpasses in the final output. In fact, the algorithm identified culverts in both sites that were manually overlooked when creating the reference data. Ditches seem to follow the shape of the road, bending as the road bends and linear along straight roads. Altogether, based on the results the algorithm was not as efficient at enforcing flow in ditches as it was at finding road underpasses. Ditches are narrow features that are usually flat. Although ditches are often apparent in the LiDAR DEMs, they are poorly represented in these images, which resulted in the difficulty of this algorithm to enforce flow along these flat features.

Although the algorithm has substantially decreased the preprocessing time (approximately 30 minutes for 20 km²) when compared to current and similar methods discussed throughout this paper (approximately 30 to 220 minutes per 1 km²) (Flood, 2007), some errors did occur in the preprocessing of datasets. Errors of both omission and commission occurred, which were due to the fact that the criteria for creating a breach path were not met by the test datasets, meaning that there were no cells that were lower than the depression with a downslope neighbour. These errors were found to occur on the outskirts of residential areas, were on flat roads, or had unnaturally deep, flat channels.

In its entirety, this novel automated breaching technique is an effective tool for efficiently incorporating ditches and culverts into the hydrological analysis of a landscape
that has both a gradient associated with it, as well as a lack of densely forested areas. The algorithm produced more accurate representations of both overland flow when compared to outputs that excluded these anthropogenic features all together. In addition, the algorithm also enforced flow within the major streams that would have been otherwise erased from the dataset if using the most widely implemented preprocessing technique, depression filling. This novel breaching algorithm has provided some of the technical solutions to the problems related to using LiDAR DEMs to model flow in human-modified landscapes. This automated technique for preprocessing LiDAR DEMs is substantially faster than manual incorporation of anthropogenic features, and is less subjective and provides more reproducible results. The algorithm presented in this thesis may be used as a way of guiding field observations when searching for culverts. Finally, the afore-mentioned attributes of this novel technique can provide a more realistic representation of surface drainage patterns, which can in turn help to more accurately predict potential water pollution source areas, predict potential soil erosion zones, and increase the accuracy of soil moisture modelling.
REFERENCES CITED


Southwestern Ontario Orthoimagery Project (SWOOP): Orthoimagery (2006)[County of Chatam/Kent]. First Base Solutions


Thompson, J., Bell, J., Butler, C. (2001). Digital elevation model resolution: effects on terrain attribute calculation and quantitative soil-landscape modeling *Geoderma* 100: 67 – 89


