This is to certify that the
thesis entitled

USING DIGITAL ELEVATION DATA TO PREDICT
SLOPES OF COASTAL SAND DUNES IN
BERRIEN COUNTY, MICHIGAN

presented by

Juliegh R. Bookout

has been accepted towards fulfillment
of the requirements for the

M.A. degree in Geography

Major Professor's Signature

12/11/2006

Date

USING DIGITAL ELEVATION DATA TO PREDICT SLOPES OF COASTAL SAND DUNES IN BERRIEN COUNTY, MICHIGAN

By

Juliegh R. Bookout

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF ARTS

Department of Geography

2006

ABSTRACT

USING DIGITAL ELEVATION DATA TO PREDICT SLOPES OF COASTAL SAND DUNES IN BERRIEN COUNTY, MICHIGAN

By

Juliegh R. Bookout

In 1989 the state of Michigan amended the Sand Dune Protection and Management Act to include the designation of *critical dunes*. Some of the most spectacular dunes along the Lake Michigan shore are also the most vulnerable; therefore, the distinction of critical dunes is given to these areas in an effort to mitigate the negative impacts of change. The Michigan Department of Environmental Quality is the agency charged with regulating all site alterations in these areas. The MDEQ's decision to approve or deny a property owner a permit is based primarily on the slope of the dune, which is collected in the field. Using a digital elevation model and land cover data, I believe that MDEQ agents could quickly and efficiently gather the slope measures they need to complete the permit process. In the past, accuracy issues have hindered the widespread use of DEMS in management; however, in recent years the accuracy of the underlying elevation data has significantly improved with the development of lidar. I propose that lidar used in conjunction with calculated slope and land cover, in a linear regression model, could enable users to predict true slope more accurately than a model that uses NED as its source data and those same variables.

This thesis is dedicated to my son Jackson who made the final days of my journey more beautiful than all of the ones that came before.

ACKNOWLEDGEMENTS

First and foremost I would like to thank my advisor, Dr. Alan Arbogast, who roped me into grad school, nudged me towards studying sand dunes, and then kept my type A personality in check. I would like to thank my committee members, Dr. Bruce Pigozzi and Dr. Ashton Shortridge, for their patience, guidance, and invaluable knowledge. A special thanks to the many people from MSU's RS&GIS including Jessica Moy, for providing me with the means to carry out my research, Jeff Schlueter, for assisting me in the field, and Justin Booth and Sarah Acmoody, for continuing to offer outstanding technical support long after they were being paid to do so. I would also like to express my gratitude to the Michigan Department of Environmental Quality, for allowing me to use DEQ data for this study, and Matt Warner, for his time and expertise.

Although the skills and knowledge of many people took me a long way, I could not have done any of this without the love and support of my family, especially my husband Eric. From the moment I applied to graduate school he was willing to do whatever was necessary to help me accomplish my goal, and never once let me believe that finishing my project was impossible. Finally, I want to thank God for giving me the courage and ability to push ahead even when I thought I couldn't do it.

TABLE OF CONTENTS

LIST OF FIGURES

Images in this thesis are presented in color

LIST OF TABLES

ABBREVIATIONS

ALACE: Airborne Lidar Assessment of Coastal Erosion

ALS: Airborne Laser Scanning System

ATM: Airborne Topographic Mapper

yrs BP: Years Before Present

Cal. yrs BP: Calendar Years Before Present

CZM: Coastal Zone Management

CZMA: Coastal Zone Management Act

DEM: Digital Elevation Model

dGPS: Differential Global Positioning System

DTM: Digital Terrain Model

EROS: Earth Resources Observation Systems

FGDC: Federal Geographic Data Committee

GIS: Geographic Information System

GPS: Global Positioning System

IDW: Inverse Distance Weighting

INS: Inertial Navigation System

Lidar: Light Detection and Ranging

LULC: Land Use Land Cover

MAE: Mean Absolute Error

MDEQ: Michigan Department of Environmental Quality

MDNR: Michigan Department of Natural Resources

MICGI: Michigan Center for Geographic Information

MSU RS&GIS: Michigan State University Remote Sensing & Geographic Information Systems Research and Outreach Services

NAD83: North American Datum of 1983

NASA: National Aeronautics and Space Administration

NAPP: National Aerial Photography Program

NED: National Elevation Data

NGS: National Geodetic Survey

NLCD: National Land Cover Data

NOAA: National Oceanic and Atmospheric Administration

RMSE: Root Mean Square Error

SAR: Synthetic Aperture Radar

SHOALS: Scanning Hydrographic Operational Airborne Lidar Surveying System

SRTM: Shuttle Radar Topography Mission

USACE: United States Army Corps of Engineers

USGS: United States Geological Survey

UTM: Universal Transverse Mercator

I. INTRODUCTION

Coastal sand dunes are common along the southeastern shore of Lake Michigan. These sand dunes collectively form the largest complex of freshwater dunes in the world and are highly valued for the scenic, recreational, ecological, and economic opportunities that they provide (Buckler 1979; Lichter 1995; Arbogast and Loope 1999; VanOort et al. 2001; Arbogast et al. 2002). For this reason, the State of Michigan has designated nearly 480 km of the shoreline *critical dunes* (covering nearly 32,000 hectares) in an effort to mitigate the negative impacts of human-induced change and prevent irreversible damage to the ecosystem (MDEQ Atlas of Critical Dunes 1989; Bernd-Cohen and Gordon 1999).

The significance of the sand dune ecosystem is evident by the presence of several rare plant and animal species, including the Pitcher's thistle (*Cirisium Pitcheri*), a threatened plant species, and the Piping Plover, a federally endangered bird species (Lake Michigan Federation 1999). From a recreational standpoint, residents, visitors, and the state and local economy benefit from the many national, state, and county parks that occur within Michigan's coastal dunes (Lake Michigan Federation 1999; Arbogast et al. 2002). In addition to the revenue generated from tourism, the dunes are also highly valued by industry as a source of foundry sand (Lake Michigan Federation 1999).

Beginning in the early 1900's, Lake Michigan sand dunes were targeted by local industries primarily for creating foundry casting molds. Not only were the physical and chemical properties of Lake Michigan dune sand ideal for this use, it was also widely available and convenient. Over the next few decades Michigan's automobile industry flourished and, in turn, the sand mining industry intensified. In the most extreme cases, entire dune systems were completely destroyed by overzealous mining practices (Buckler 1979; Lake Michigan Federation 1999; Albert 2006). In the early 1950s, Michigan residents began to take notice of the disappearing dunes. This culminated in an outpour of public concern that lead to the eventual passing of the *Sand Dunes Protection and Management Act* in 1976 (Act No. 222, Public Acts of 1976; Lake Michigan Federation 1999).

In the years that followed, concerns also grew over other developmental pressures on the dunes, such as recreation and construction. These new concerns, combined with the continued destruction of the dunes from mining, lead to amendments to the act in 1989, which resulted in the *critical dune* designation. The 1989 amendments also included the addition of a program that established standards and a permitting program for development in areas designated *critical dunes* (Michigan State Legislature Acts No. 146 and 147, Public Acts of 1989). More recently, Michigan reorganized its environmental acts dividing the 1976 Sand Dune Protection and Management Act into Part 353, Sand Dune Protection and

Management, and Part 637, Sand Dune Mining (Michigan State Legislature Act No. 451, Public Acts of 1994).

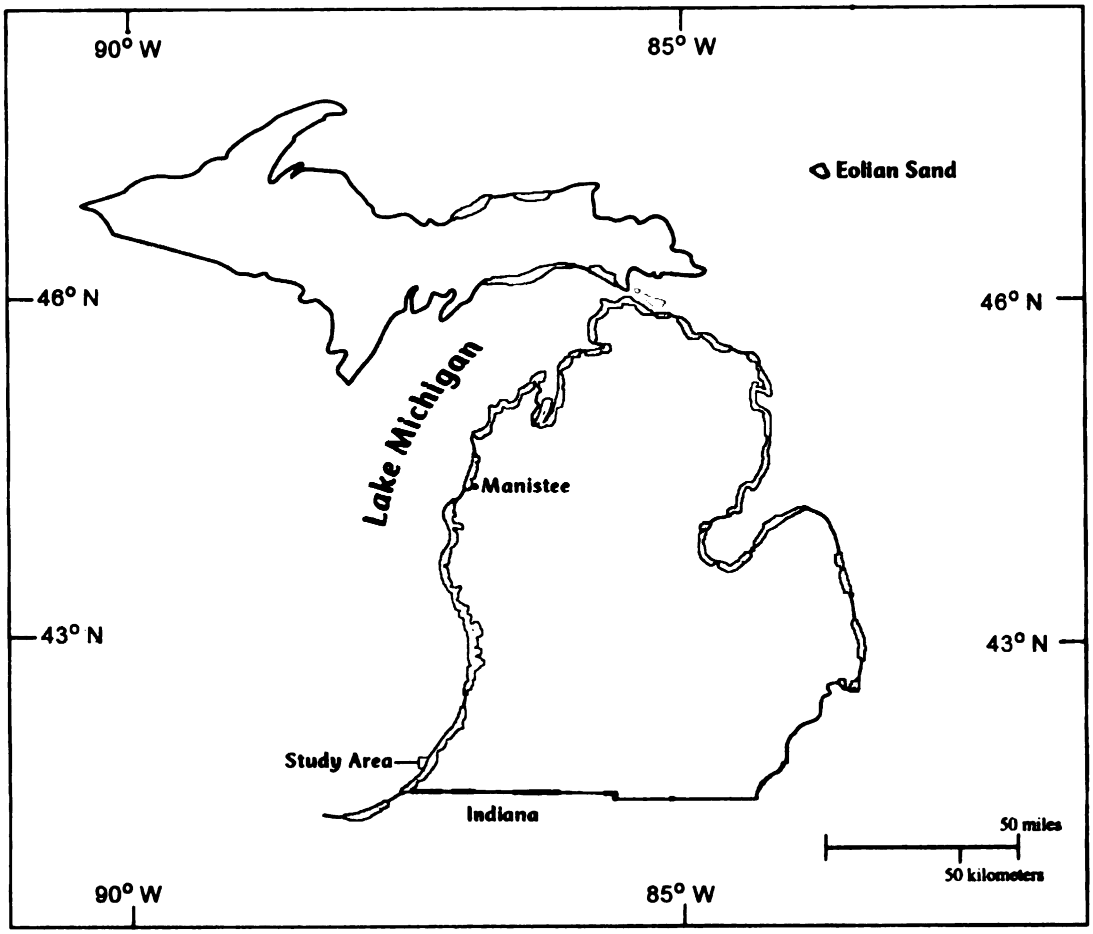


Figure 1-1 Map of the coastal dunes in lower Michigan and NW Indiana including the location of the study area (modified from Arbogast et al. 2002).

The administration and management of critical dune policies is divided between two branches of the Michigan Department of Environmental Quality (MDEQ). Sand dune mining (Act No. 451, Part 637, Public Acts of 1994) is regulated by the Office of Geological Survey, while all other matters concerning critical dunes, including the permitting process, are handled by the Land and Water Management

Division (Michigan State Legislature Act No. 451, Public Acts of 1994). The focus of this research is on residential and commercial site alteration, therefore, from hereafter I will be referring only to the Land and Water Management Division unless otherwise specified.

Problem Statement

The decision of the MDEQ to deny or reject an application for a permit to alter a site in critical dunes is based primarily on the slope of the dune that will be affected. This stipulation requires agents to conduct an onsite inspection, which can be very time consuming depending on the scope of the project. At present, the State of Michigan employs eleven agents to cover nearly forty counties. As such, these agents are constantly pressed for time. Further, the effectiveness of this program has never been evaluated. It is estimated, however, that of the hundreds of sites permitted every year, only about 10% of ever receive follow-up visits from permitting staff (Warner, personal comm., 2002).

The lack of accountability is troubling given the results of a number of studies assessing the impact of development on natural dune systems in other coastal locales (Gares 1990; Nordstrom 1994; Nordstrom et al. 2000; Brown and McLachlan 2002; Nordstrom et al. 2002). Gares (1990) studied dunes along a stretch of developed shore in New Jersey and concluded that residential development there is severely altering the naturally functioning dune system. Overall he found that the dunes had a gentler slope, lower elevation, and lower sediment transport than their undeveloped counterparts in the same area. He attributed these differences to less vegetation, the construction of sand fences, and the presence of houses, all of which are anthropogenic. Parking lots and roadways are impermeable surfaces that provide unobstructed pathways for entrained sand, while buildings may alter the local flow of wind as well as the location of accretion on beaches and dunes (Nordstrom 1994). With a host of threats facing coastal ecosystems and increasing developmental pressures increase, there is an urgent need for improved legislation and protection (Brown and McLachlan 2002).

Research (Bernd-Cohen and Gordon 1999; Hershman et al. 1999) also shows that Michigan is not alone in respect to the problems faced managing development in the coastal zone. Hershman et al. (1999) report that all coastal mangers are overburdened with implementation tasks and, when combined with the political and legal pressures of managing a valuable resource, the focus is almost always on the current decision processes, not monitoring and evaluating past actions (Hershman et al. 1999). In order for a coastal program to change or improve, program managers must have the time and resources available to evaluate the state of coastal resources and make that information available to policy makers (Bernd-Cohen and Gordon 1999; Hershman et al. 1999).

At this time, a majority of states with a coastal zone management (CZM) program do incorporate digital technology to track permits, however, none of these states employs a database directly related to the extent to which resources are affected by permits or policies (Bernd-Cohen and Gordon 1999). In Michigan, the key to achieving a more objective assessment of the coastal zone may lie in integrating the use of digital and remotely sensed data, such as a digital elevation model (DEM), with the current sand dune protection and management program. Much like the current permitting process, gathering the data necessary to do that using traditional means, such as ground based surveys or photogrammetry, could become prohibitively time consuming and expensive. However, remote sensing can provide spatially dense quantitative data over regional scales. If sufficiently accurate this data, when used in a Geographic Information System (GIS), could be invaluable in determining and understanding patterns and magnitudes of dune development (Sallenger et al. 2003).

The National Elevation Dataset, or NED, is one digital data source that is presently available to all potential users on the United States Geological Survey (USGS) website (Gesch et al. 2002). It is a compilation of many data sources, deemed the "best available", including 7.5 minute, 15-minute, 2-arc-second, and 3-arc second DEMs (Smith and Sandwell 2003). The metadata that accompanies NED allows users to calculate accuracy statistics based on source DEM characteristics; however, the actual error in any one dataset may differ from what is stated depending on several factors including local terrain characteristics, the application for which it is being used, and any parameters besides elevation (Bolstad and Stowe 1994; Gesch et al. 2002; Hodgson et al. 2003). Although NED may be sufficient in some areas and for some applications, the recent shift has been toward models that require a more detailed and spatially explicit representation of process, which requires a better quality source for topographic data (French 2003; Barber and Shortridge 2004).

Light Detection and Ranging, or lidar, is a technology that can produce high-resolution elevation datasets; it is also the technology that may be able to better meet the needs of coastal zone managers. This technology can quickly gather spatially dense elevation data over a large area and is generally regarded as a more accurate, more efficient, and less expensive collection method for the creation of DEMs than alternative methods (Woolard and Colby 2002; Sallenger et al. 2003; Barber and Shortridge 2004).

The decision for any agency to incorporate remotely sensed data into their current practices and standards and decide between lidar, NED, and any other dataset for a particular application can depend on any number or combination of factors, including accuracy requirements, purpose, cost, and availability. In the case of the MDEQ, the accuracy of the terrain characteristic slope is of primary importance given that their permit decision is based on this measure and there could be legal consequences associated with an erroneous decision based on inaccurate measurements.

Although researchers have recommended using either a lidar or USGS Level 2 digital elevation model over other elevation products if an application requires accurate slope measures, investigations have been unable to provide evidence that either product is suitable for such uses, based on study area characteristics (Hodgson et al. 2003). While research (Barber and Shortridge 2005) has shown that slope is particularly sensitive to and greatly affected by resolution and, therefore, a high-resolution dataset like lidar may be a better choice, the researchers in that case also stipulate that this is more true in areas of low relief. This added condition begs the question, "What role do terrain attributes play in the measurement of elevation and its derivations?"

It has been previously established (Chang and Tsai 1991; Bolstad and Stowe 1994; Bowen and Waltermire 2002; Hodgson et al. 2003) that elevation, land cover, and slope influence remote collection methods, leading researchers to seek a better understanding of how in situ terrain attributes affect the quality of a digital elevation model and its derivatives. For example, researchers have determined that both elevation error and modeled slope error increase when slopes on the ground are steeper (Chang and Tsai 1991; Bolstad and Stowe 1994; Hodgson et al. 2003). Bolstad and Stowe (1994) also determined that, for a USGS Level 1 DEM, the largest errors were found in areas exhibiting the highest and lowest elevations for the study area. Further, error in elevation and slope measurements derived from both lidar and USGS DEMs increases to varying degrees based on land cover attributes, including forest canopy, stem obstruction, and understory vegetation (Bolstad and Stowe 1994; Hodgson et al. 2003).

Research Purpose

Given that errors are generally systematic and influenced by patterns of land cover, elevation, and gradient, it may be possible to determine actual slope if land cover, elevation, and calculated slope are known variables. The purpose of my research, therefore, is to determine:

1.) What terrain variables, or combination of variables, when used in a regression model, most closely predict true slope?
2.) What is the error in the predicted slope values?
3.) To what degree do the source, or collection method, and/or resolution of the elevation data influence the model?

By conducting this research and answering these questions, I am hoping to better understand how digital data can be used to aid management decisions that would typically require data collection in the field. Further, I hope to explore the relationships between digital data collection methods, data resolution, and the true slope values at any given point. Given the fairly recent emergence of lidar technology in the collection of elevation measures, a greater knowledge of its potential to model topography and its relationship with terrain characteristics that influence the elevation and derived slope values, could be beneficial to agencies like the MDEQ as well as other environmental applications.

II. LITERATURE REVIEW

The topic of this research concerns coastal sand dunes and their management, as well as remote sensing and terrain modeling, therefore, each of these areas will be addressed in this review of relevant literature. This chapter can be divided into two distinct parts. The first will cover coastal sand dune formation, focusing specifically on the origin of the dunes found in Michigan, and the history of sand dune protection and management from the national level down to the state level. The latter half of the chapter is devoted to digital elevation data including its collection, processing, and use, particularly as it relates to modeling applications.

Coastal Sand Dune Formation

Knowledge of coastal dunes and dune geomorphology is essential to understanding the spatial relationships between elevation, gradient, and vegetation on a coastal eolian landscape. Coastal dunes are located above the high water marks on sandy beaches of oceans and lakes, and can be found across all latitudes. Dunes range in size from small, transient features to very large, stabilized dune fields, and can exist in almost any climate (Nordstrom et al. 1990).

Many studies have explored eolian sediment transport and deposition in relation to the coastal environment. Dune formation is a function of sediment grain size, the beach profile, and wind regime. In general, the development of coastal dunes is dependent upon 1) a source of sandy sediments, 2) a prevailing, onshore wind with a velocity sufficient to move those sediments, and 3) an area of land partially protected from wave action (Smith 1988; Nordstrom et al. 1990; Orme 1990; Sherman and Hotta 1990). Dunes most readily form on gently sloping shorelines under dissipative wave conditions where the foreshore dries out at low water. Ideal eolian sands are well-rounded, sand-sized (2 to .2 mm) quartz grains with a specific gravity in air of 2.65 g cm^{-3} (Nordstrom et al. 1990).

Eolian forces cause sediment movement when shear stress exceeds a threshold value, or the velocity required to entrain a particle of sand. Momentum is then transferred from the air to the sand grains and initiates movement by suspension, saltation, or creep. On beaches and coasts sediments are most often transported by saltation or creep because sand-sized particles are generally too large to be carried in true suspension (Sherman and Hotta 1990). Saltation, the dominate process of eolian sand movement, occurs when sand grains bounce or skip along the ground ($\leq$ 1 cm above the surface) by repeated lifting and deposition. As the saltating grains skip along the surface they dislodge smaller particles, which may then become saltating grains, and/or larger particles that slowly roll forward along the ground. The latter process, which accounts for a somewhat smaller portion of eolian sand movement, is called traction or creep (Figure 2-1; Gabler et al. 2004).

Forces that oppose the initial motion include a cohesive agent (moisture) and gravity. Once in motion, however, with sufficient velocity the sand can continue to move until it encounters either topographic obstructions or surface obstacles, when it is deposited and may accumulate in drifts. These drifts, or dunes, create a topographic barrier that further reduces the winds velocity. A dune will only grow as long as its size does not impede the wind's ability to feed sand to the dune (Gabler et al. 2004).

Sand dunes have three main features, a back slope, which is erosional and faces upwind, a crest, and a slipface or leeward slope (Figure 2-1). A dune will migrate as long as winds are strong enough to carry sand up the windward slope to the crest (Gabler et al. 2004). As sand is deposited over the crest, the angle of the lee slope continues to build until it reaches the steepest angle dry sand can hold without falling or the angle of repose (~ 30 – 35°). Once the angle of repose is met, sand will slide down the slip face causing the dune to advance downwind (Figure 2-1). A new slipface is created as the dune advances layer by layer. Small dunes (< 10 m high) can migrate up to 40 meters in a year while larger dunes move more slowly (Gabler et al. 2004). Sand will continue to move until the net transport rate decreases, which can be initiated by an increase in slope, decrease in wind velocity, and/or an encounter with vegetation (Sherman and Hotta 1990). Likewise, dune stability is enhanced by a reduction in the sediment supply or wind velocity, typically reflecting changes in the environment or an increase in vegetation (Gares 1990; Orme 1990).

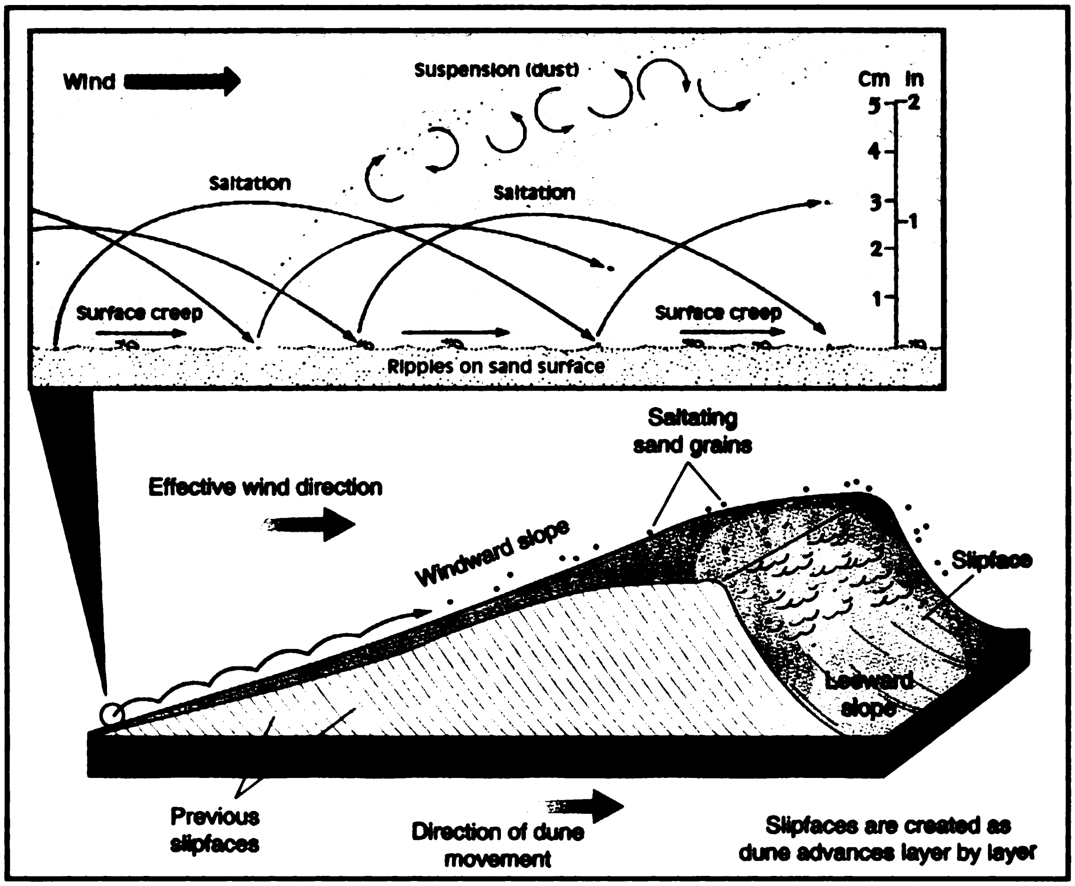


Figure 2-1 Wind movement by saltation and surface creep resulting from the impact of saltating grains. The wind moves sand up the windward slope of the dune toward the crest. The sand then slides down the steeper leeward side or slipface, causing the dune to advance (modified from Gabler et al. 2004).

Vegetation, if present, acts in three ways, specifically it 1) reduces wind velocity, 2) prevents wind from being directed against the land surface, and 3) secures sediments with its root network (Gabler et al. 2004). Sedimentation in the coastal environment is often limited by the density and growth rate of vegetation (Nordstrom et al. 1990). Vegetation serves to encourage deposition (trapping sand) and stabilize deposits by dissipating the energy of the wind and forcing it to deposits its load (Olson 1958c; Orme 1990; Robertson-Rintoul 1990; Sherman and Hotta 1990). Dunes in more developed areas often have sparse vegetation cover due to heavy pedestrian traffic and, therefore, little stability in the long term (Gares 1990).

Coastal Dunes of Lake Michigan

Coastal dunes are common along the eastern shore of Lake Michigan, with the most numerous being those that mantle lake terraces extending south from Manistee, MI and into Indiana (Figure 2-2). These sand dunes may collectively be the largest complex of freshwater dunes in the world (VanOort et al. 2001; Arbogast et al. 2002). The earliest published research on Lake Michigan sand dunes focused on ecology, specifically the vegetation found in and among the dunes in Indiana Dunes State Park along the southern shores of Lake Michigan (Cowles 1899). While studies of the plant species and the role of vegetation on the dunes continued (Olson 1958a, 1958c), a growing number of studies began to focus on dune geomorphology by concentrating on the processes, controls, and timing of dune development with an emphasis on lake-level oscillations (Dow 1937; Scott 1942; Olson 1958d; Buckler 1979; Thompson 1992; Lichter 1995; Thompson and Baedke 1995, 1997; Arbogast and Loope 1999; Baedke and Thompson 2000; Loope and Arbogast 2000; VanOort et al. 2001; Arbogast et al. 2002).

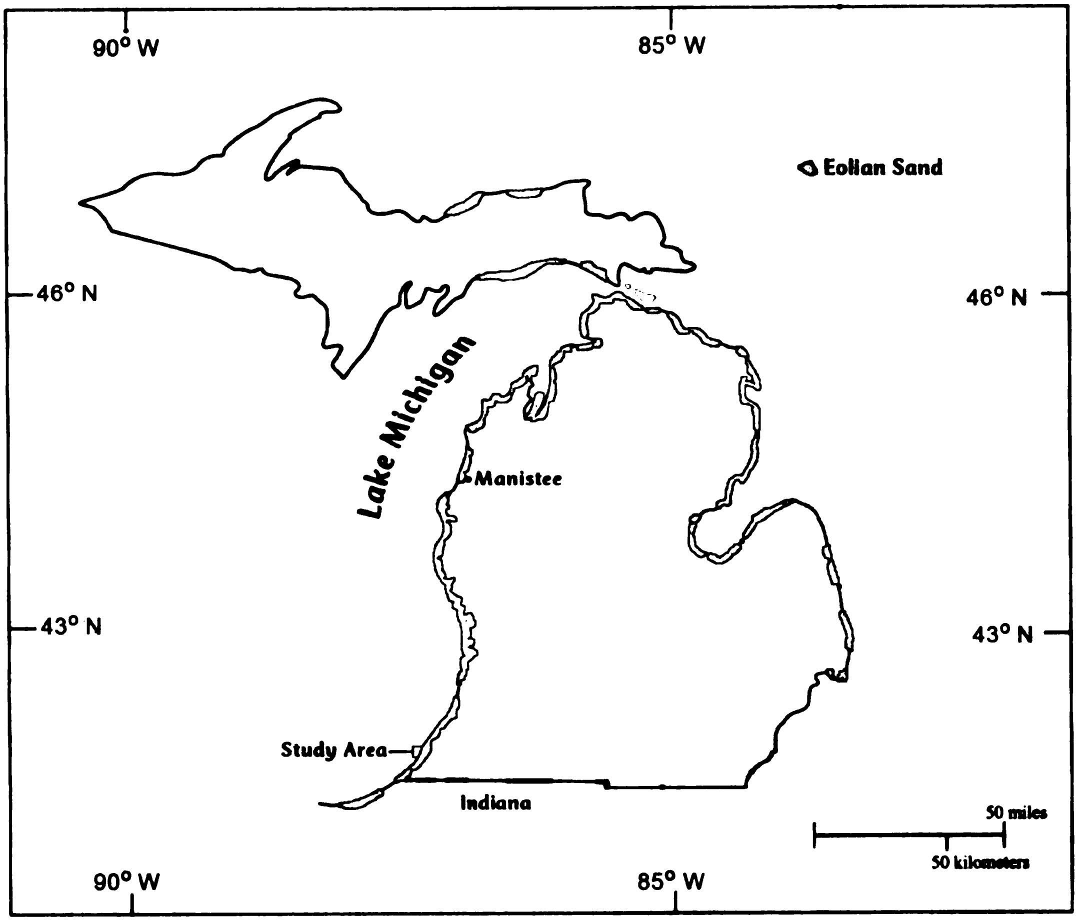


Figure 2-2 Map of the coastal dunes in Michigan and NW Indiana including the location of Manistee, which was noted in the text, and the study area (modified from Arbogast et al. 2002).

Both Dow (1937) and Scott (1942) recognized that fluctuations in lake level were the dominant control of the sand supply to the dunes on the eastern shore of Lake Michigan. Dow's (1937) work centered on what he called "perched" dunes at Sleeping Bear Point in Empire, MI. He used the term perched dunes to refer to eolian bedforms elevated above the present lake-level, often on high headlands, and overlaying non-eolian sediments. This includes dunes that mantle lacustrine sediments of post-glacial beaches as well as dunes found much higher above the lake on moraines. He found that when lake levels are high, waves undercut the bases of the bluffs exposing a fresh supply of sediments that can then be entrained by the wind and deposited at the top of the bluff.

Olson (1958d) expanded on the idea of lake-level oscillations as the dominant control of sediment supply, and recognized two basic dune forms on the southern shore of Lake Michigan: primary and secondary (Figure 2-3). This classification is loosely based on positional, morphological, stability, and age factors. Primary dunes (also called "incipient" or "embryo" dunes) are young and located near the lake, while secondary dunes, which include established foredunes, parabolic dunes, and blowouts, are positioned farther inland (Nordstrom et al. 1990).

In 1958, Olson (1958a, 1958b, 1958c, 1958d) published a series of four articles describing the development of Lake Michigan coastal dunes. While the first three articles focus on the role of vegetation in dune building and stability, in the fourth, Olson (1958d) describes the formation of foredunes on the shore of Lake Michigan. He states that the process is contingent upon periodic fluctuations of the lake level occurring roughly every 30 years. During periods of low or receding lake levels, the beach widens and new dune growth begins as colonizing plants intercept eolian sand. If these primary dunes become sufficient enough to withstand wave erosion, when average or high lake levels return they become a new foredune ridge. Secondary dune building, on the other hand, generally occurs when the lake level is high or winds are strong. At this time, the previously established foredune ridges are destabilized by wave and wind erosion and sediment removed from the ridge can then be returned to the slope faces of the secondary dunes (Olson 1958d).

The next significant study of the geologic characteristics of Lake Michigan coastal dunes was conducted by Buckler (1979) in response to the State of Michigan *Sand Dune Protection and Management Act* (Michigan State Legislature Act No. 222, Public Acts of 1976). He inventoried and classified nine dune forms found along the Lake Michigan shore including parabolic dunes, linear dune ridges, dune terraces, dune platforms, domol dunes, complex dune fields, dune flats, marginal sand aprons, and interdune lowlands. The classification scheme proposed by Buckler (1979) combined several dune attributes including dune form, relative relief, orientation, arrangement, and the stratigraphic relationship of the dune to the underlying sediments.

The well-developed dune fields on the southeastern shore of Lake Michigan are large parabolic dunes (> 60 m) that may be separated from the lake by small foredunes (Figure 2-3; VanOort et al. 2001; Arbogast et al. 2002). Buckler (1979) described parabolic dunes as non-elevated, with a parallel orientation and high relief, often occurring in groups and overlapping. Widths, he noted, can be more than 400 m with lengths exceeding 900 m. On its windward side, a parabolic dune has a concave shape with a gentle slope that becomes steeper closer to the crest. The convex landward side has a steeper slope that descends abruptly from the crest at the angle of repose. Blowouts, or unvegetated, depressions created by deflation, may exist between the tails of the dune, which are often stabilized by vegetation (Buckler 2001; Gabler et al. 2004). Two generations of parabolic dunes are frequently found adjacent to one another with the shoreward group being higher, more individually well-defined forms while the landward group is lower, more dispersed, and less distinctive. The sand dune formation, or assemblage of dunes, that separates the present coastal zone from inland activities is also given the term *barrier dune* by the Michigan Legislature (Buckler 1979) .

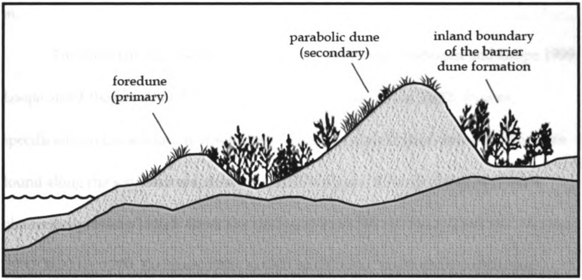


Figure 2-3 Cross section of a typical Lake Michigan dune field (modified from Wilson 2001).

In the 1990's several geomorphic studies of Lake Michigan dunes directed at lake level oscillations and beach ridge development were published (Thompson 1992; Thompson and Baedke 1995, 1997). Together, these studies focused on five beach ridge complexes around Lake Michigan, including sites at Toleston Beach, IN, Sleeping Bear Dunes National Lakeshore, MI, Wilderness State Park, MI, Manistique, MI, and Ridges Sanctuary, WI, and used vibracores collected from the bases of wetlands between the ridges to reconstruct four late Holocene (from 200 to 4,700 cal. yrs. BP) relative lake level curves. In 2000, Baedke and Thompson compared the radiocarbon dates of samples collected in all five complexes and determined that, when isostatic rebound is accounted for, all four curves show similar variations in lake-levels. Small beach ridges could be linked to ~ 33-year fluctuations of .5 - .6 m, while larger groups of ridges (4 - 6) corresponded to ~ 160-year fluctuations of .8 - .9 m.

The most current research on dune geomorphology (Arbogast and Loope 1999; Loope and Arbogast 2000; VanOort et al. 2001; Arbogast et al. 2002) focuses specifically on the evolution of some of the largest coastal dunes, many of which are found along the southeastern shore of Lake Michigan. It is significant because it challenges previous beliefs about the timing of dune development (Dorr and Eschman 1970; Buckler 1979; Eschman 1985) as well as the broad applicability of Olson's (1958d) foredune model by comparing dates obtained from buried soils to the historical lake-level curve. Arbogast and Loope (1999) first conducted a study of buried soils at three sites along the lakeshore to test the theory that all of the dunes along the western coast developed between 6,000 and 4,000 yrs BP during the Nipissing high stand of ancestral Lake Michigan (Dorr and Eschman 1970; Buckler

1979; Eschman 1985). The researchers found that although dune building at one of the sites began during the Nipissing transgression, dune formation at the other two sites began following the transgression (between 4,300 and 3,900 cal. yr BP and 3,300 and 2,900 cal. yr BP). They concluded that the sand supply to these lake-terrace dunes originated from bluff destabilization related to minor peaks in lake-level that occurred post-Nipissing.

A second study conducted by Loope and Arbogast (2000) aimed to distinguish between the origin of foredunes (~ 5 m thick) on the south and southeastern shore initially described by Olson (1958d) and that of larger dunes (~ 50 m thick) mantling lake terraces along the eastern shore, by comparing the dates of buried soils at 32 sites to a late Holocene lake-level curve (Baedke and Thompson 2000). The smaller foredunes occur close to the lake, are subject to direct wave action, and, as Olson (1958d) suggested, form as the lake level falls and the beach widens (the quasi-periodic ~ 30 yr regressions documented by Baedke and Thompson 2000). The larger dunes that sit higher on the landscape and are remote from direct wave action build as the lake-level rises and waves cut into bluffs and previously established dunes. Using radiocarbon dating they were able to correlate dominant periods of sand accumulation at a majority of the sites with ~150 yr lake highstands that have occurred over the past 1500 years, making the age of many of the larger dunes younger than previously thought (Loope and Arbogast 2000).

Two studies of buried soils found in sand dunes south of Holland, MI were conducted by Van Oort et al. (2001) and Arbogast et al. (2002) and found that deposition of eolian sand began there during the Nipissing high stand (~ 5500 cal. yr BP) and lasted for ~ 3500 years. Arbogast et al. (2002) found that a dominant period of rapid dune building occurred between ~ 4000 and 2500 cal. years BP, and was only punctuated by brief periods of stability. Subsequently, the dunes were stable for ~2000 years before becoming active again between ~1000 and 500 cal. yr BP (similar to findings north of the study area). They hypothesize that the increase in sand supply as lake levels fell immediately following the Nipissing transgression most likely accounts for the onset of the initial growth period, while later growth (occurring ~ 3200, 2400, and 900 cal. years BP) is associated with intervals of higher lake levels and wave destabilization of bluffs.

National Coastal Zone Management

In the early part of the twentieth century, states began recognizing the value of the shoreline as both a natural resource and an economic commodity. In response to the rapid development of coastal resources during this time, the United States Congress enacted the Coastal Zone Management Act (CZMA). The CZMA had two primary objectives: 1) the protection of natural resources within the coastal zone and 2) the management of coastal development to minimize loss of life and property caused by improper development, including the destruction of natural protective features such as beaches and dunes (United States Legislature Coastal Zone Management Act of 1972, Sec. 303). Following the enactment of the CZMA, all coastal states (including Great Lakes states and island territories) were offered policy guidance, financial resources, and legal tools as incentives to upgrade their capacity for coastal zone management. By the early 1980's most state programs had been approved and implementation began thereafter (Figure 2-4). As of 1999, federal funding for CZM was approximately $50 million dollars, with a 50% state or local matching requirement for most programs (Hershman et al. 1999).

Under the CZMA, states determine the coastal zone boundary, key coastal problems, the policies and laws that will address them, and the state and local organizations to be involved in enforcement. The federal, state, and local governments are all given roles in CZM with considerable flexibility given to defining the extent of those roles, which has lead to many unique management programs around the country (Hershman et al. 1999). Although there are specific objectives of the CZMA, conflicting trends and policies are continually challenging the ability of states to protect natural resources. States are under pressure to balance the need for public access and recreation, the high economic value of coastal property, the protection of private property rights, and the preservation of natural resources (Bernd-Cohen and Gordon 1999).

Twelve of the twenty-nine states with CZM programs have amended their CZM program in the past two decades to increase the protection of beaches, dunes, bluffs, and rocky shores (Bernd-Cohen and Gordon 1999). For example, several states including Florida, Hawaii, Michigan, New Hampshire, New Jersey and North Carolina have expanded CZM jurisdiction landward and extended inland setback requirements. California, Maine, and New Jersey have all amended their CZM programs to further limit development on dunes and bluffs, while Connecticut, North Carolina, and South Carolina all added amendments to reduce the number of shoreline protection structures built (Bernd-Cohen and Gordon 1999).

In addition to improving their policies, states also use a wide range of tools to implement CZM. Land acquisition, coastal zoning, permit programs, research, and public education and awareness programs are common methods used in CZM programs (Bernd-Cohen and Gordon 1999; Hershman et al. 1999). Most states have also developed and/or incorporated digital technology to track permits; however no states employ a database on coastal statistics or resources affected by permits or policies (Bernd-Cohen and Gordon 1999).

Figure 2-4 Map of the twenty-nine states with CZM programs approved by the federal government. Illinois, the one state with approval pending, is currently seeking to reinstate their former program.

Sand Dune Protection & Management in Michigan

Nearly all of the studies mentioned earlier in this chapter (Cowles 1899; Buckler 1979; Lichter 1995; Arbogast and Loope 1999; Loope and Arbogast 2000; VanOort et al. 2001; Arbogast et al. 2002) noted the importance of Lake Michigan sand dunes for the unique scenic, recreational, ecological, and economic opportunities they provide. This section begins with a brief description of those qualities which increase the value of the dunes, to establish why they deserve and require legislative protection. It will end with a summary of the current management scheme in place to mitigate the effects of ongoing development, as well as some of the challenges of coastal zone management in Michigan.

Cowles (1899) was first to state the value of the dune habitat in the study of plant succession and ecology, due to the dynamic and often extreme physical conditions that exist there, including temperatures, sunlight, and ongoing wave and wind action. The importance of sand dunes as habitat is evident by the several rare plant and animal species found there, including the threatened plant species Pitcher's Thistle (*Cirisium Pitcheri*) and the Piping Plover, a federally endangered bird species that relies on the shoreline for nesting (Lake Michigan Federation 1999). In addition to the preservation of plant species, the vegetation found in and among the dunes is an important dune building agent that often dictates the area, shape, and height of a dune by trapping sand and altering surface roughness (Cowles 1899; Olson 1958c).

From a recreational standpoint the dunes contain many national, state, and county parks that are enjoyed year-round by visitors (Figure 2-5; Arbogast et al. 2002). For example, P.J. Hoffmaster State Park in Muskegon County, MI attracts nearly 500,000 visitors each year with miles of hiking and cross country ski trails. A larger park like the Sleeping Bear Dunes National Lakeshore in Leelanau County, MI draws an estimated 1 million visitors annually with a financial benefit estimated at more than $39 million since its creation in 1970 (Lake Michigan Federation 1999).

In addition to the revenue generated from tourism, the dunes are also highly valued as a source of foundry sand. Since the early 1900's Lake Michigan sand dunes have been intensively mined, and in some cases completely destroyed, primarily for use as foundry cores and molding sands in the automobile industry (Buckler 1979; Lake Michigan Federation 1999). Although severe restrictions had been placed on the mining of coastal dune areas in other states, under the control of local government, intensive mining continued on the shore of Lake Michigan with little regulation (Buckler 1979). It was not until the middle of the twentieth century that residents began to notice the disappearance of some of the highest and most magnificent dunes, which had once been local landmarks, and in the 1970's, public outcry finally lead to the passing of the *Sand Dunes Protection and Management Act* in 1976 (Lake Michigan Federation 1999).

Although Michigan's *Sand Dunes Protection and Management Act* was initially enacted to regulate the mining industry, subsequently, concerns grew over other developmental pressures on the dunes like recreation, silviculture, and residential and commercial construction (Holt 2002). In response to the continued destruction of the dunes from mining, and the added pressure of various other uses, the act was amended in 1989 to include these activities (Michigan State Legislature Acts No. 146 and 147, P.A. 1989). At that time the Michigan State Legislature found that:

(a) The critical dune areas of this state are a unique, irreplaceable, and fragile resource that provide significant recreational, economic, scientific, geological, scenic, botanical, educational, agricultural, and ecological benefits to the people of this state and to people from other states and countries who visit this resource.

(b) Local units of government should have the opportunity to exercise the primary role in protecting and managing critical dune areas in accordance with this part.

(c) The benefits derived from alteration, industrial, residential, commercial, agricultural, silvicultural, and the recreational use of critical dune areas shall occur only when the protection of the environment and the ecology of the critical dune area for the benefit of the present and future generations is assured (Michigan State Legislature Acts No. 146 and 147, Public Acts of 1989).

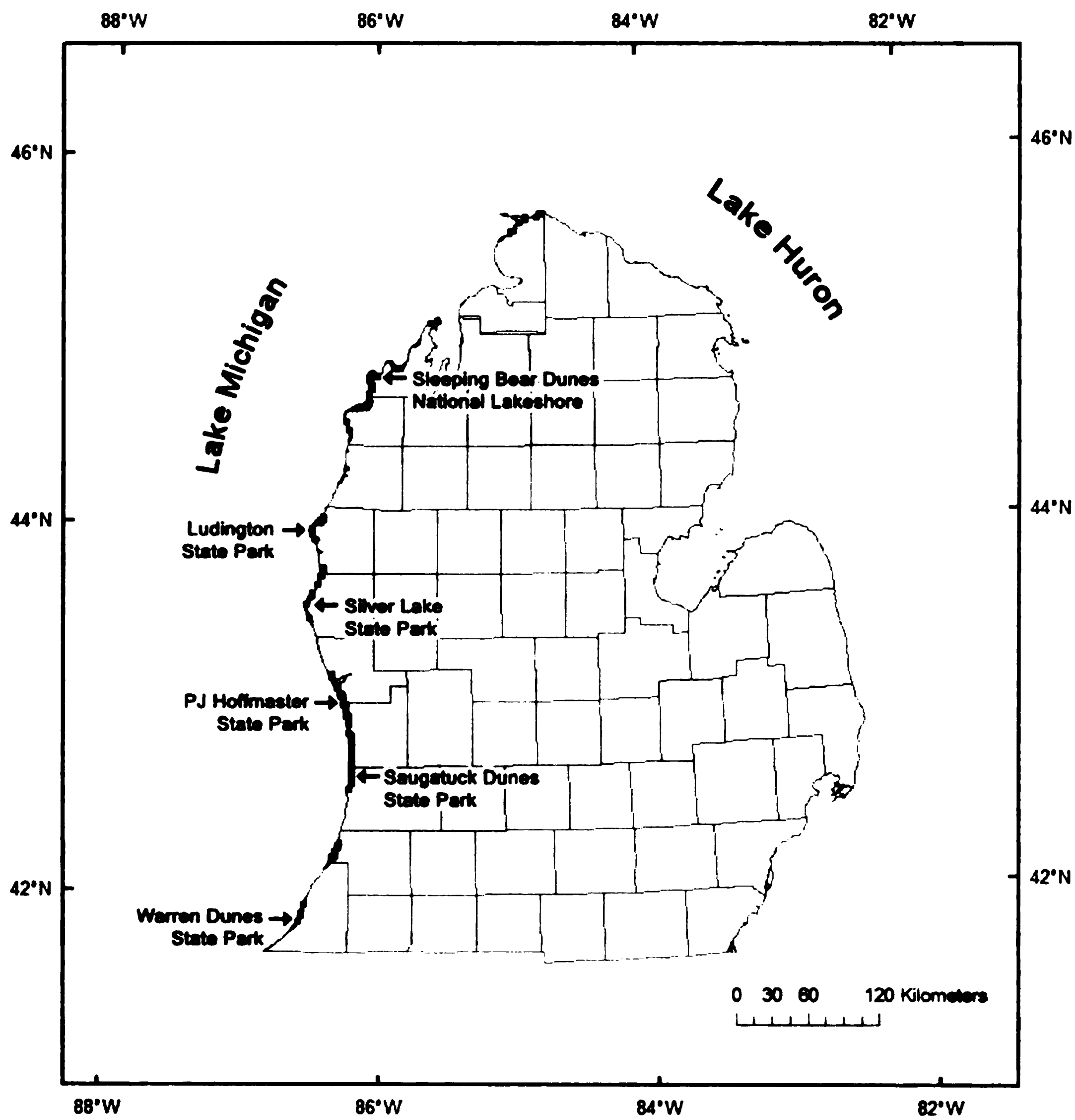


Figure 2-5 Map of the areas designated "critical dunes" in Lower Michigan.

At present Michigan has designated nearly 480 km of critical dunes (covering nearly 32,000 hectares), 400 km of natural preserves, and 500 km of high risk erosion areas along its shores (Figure 2-5; Bernd-Cohen and Gordon 1999). However, one of the biggest imperatives of this legislation is that a permit is required before any of the included activities can be carried out in an area designated as critical dunes (Holt 2002).

The Michigan Department of Environmental Quality is the state agency that administers all activities within critical dune areas. Within the MDEQ, management responsibilities are shared by two separate branches. Sand dune mining is regulated by the Office of Geological Survey, while all other matters concerning critical dunes are handled by the Land and Water Management Division (Michigan State Legislature Act No. 451, Public Acts of 1994). Various means are used to administer CZM including regulatory and planning tools (permits and restrictions), direct land management, restoration and acquisition, investment restrictions and incentives, education outreach programs, and mandated research (Bernd-Cohen and Gordon 1999). At the forefront of sand dune management is a permit program that prohibits any new development or modifications in critical dune areas on slopes measuring 33% or greater and on the first lakeward facing slope, and requires a variance for projects affecting slopes greater than 25%. Further, an environmental impact statement is required for all special use projects such as the building of condominiums or subdivisions (Holt 2002). With the exception of sand mining, the overall effectiveness of this program has never been evaluated.

Digital Elevation Data: Collection, Processing, & Use

Nordstrom et al. (1990) concluded that the management of dunes everywhere should not be based on subjective assessments, using words like 'fragile' and 'vulnerable' that have no quantitative meaning to describe a dune system, but rather a scientific approach is both possible and necessary. In Michigan, the key to achieving a more objective assessment may lie in integrating the use of remotely sensed data, such as digital slope and elevation models, with the current sand dune protection and management program. Remote sensing provides spatially dense quantitative data over regional scales that, if sufficiently accurate, would be invaluable in determining and understanding patterns and magnitudes of dune development (Sallenger et al. 2003).

The importance of accurately characterizing topography is well established in geomorphological research (Krabill et al. 2000; Wise 2000; Spinney 2001; Bowen and Waltermire 2002; Woolard and Colby 2002; Sallenger et al. 2003; White and Wang 2003; Barber and Shortridge 2004; Nagihara et al. 2004; Barber and Shortridge 2005). Two areas where topographic models have proved to be an invaluable resource include quantifying beach morphology and shoreline mapping (Krabill et al. 2000; Woolard and Colby 2002; Sallenger et al. 2003; White and Wang 2003) and drainage network and watershed delineation (Wise 2000; Spinney 2001). Gathering this data using traditional means, such as ground based surveys or photogrammetry, can be prohibitively time consuming, expensive, and sometimes inaccurate based on environmental conditions (Sallenger et al. 2003). So in response to the need for reliable and accessible topographic data, in the mid to late-nineties the USGS began developing a seamless mosaic of the best available elevation data for the conterminous United States. By the year 2000, work was complete on the National Elevation Dataset and 30-meter raster elevation files were available for download free of charge from the USGS website (Gesch et al. 2002).

Over the past few years, however, the shift has been toward models that require a more detailed and spatially explicit representation of process, such as determining the rate shoreline erosion and the volumetric change of a coastline (Woolard and Colby 2002; Sallenger et al. 2003; White and Wang 2003), and this has created the need for greater quality topographic data (French 2003; Barber and Shortridge 2004). Lidar is a technology that can produce high-resolution elevation datasets; it is also the technology that may be able to meet that need. Several collection methods as well as digital elevation products exist, and selecting one may depend on the application, purpose and/or availability to the end user. Lidar and NED are possibly the two most common sources of elevation data available yet they are considerably different from one another.

The National Elevation Dataset

NED is a seamless elevation model produced and distributed by the USGS that covers the United States at a 1 arc second (approximately 30-meter) resolution for most areas. It is a compilation of many data sources, deemed "best available", including 7.5 minute, 15-minute, 2-arc-second, and 3-arc second DEMs that, in some cases, date back to 1978. Further, the production method, horizontal datum, map projections, and elevation units of the initial elevation data may vary within the dataset (Smith and Sandwell 2003).

The USGS has used four primary methods of producing the underlying DEMs including: 1) manual profiling, 2) automatic correlation, 3) contour-to-grid interpolation, and 4) integrated contour-to-grid interpolation. Each method uses a different combination of source materials and production methods resulting in varied data quality, artifacts, and accuracy. A DEM is categorized Level 1 or Level 2 based on the way it was produced (Hodgson et al. 2003). Level 1 DEMs are typically older and created using automated correlation, while Level 2 DEMs are derived from manual profiling of contours (1:24:000 scale digital line graphs) or aerial photography and have higher accuracy specifications (Shortridge 2003).

During assembly, procedures to remove production artifacts and minimize elevation discrepancies at the transition between different sources are used to improve the quality of the dataset. Every two months NED is updated and any new source DEMs that have become available are incorporated, which typically means replacing the oldest and least accurate source data (Gesch et al. 2001; Gesch et al. 2002). NED can be downloaded free of charge or ordered on hard media for the cost of reproduction from the NED data distribution system on the USGS website.

Spatially referenced metadata included with NED allows users to calculate accuracy statistics based on source DEM characteristics such as resolution, age, level, contour interval, and production method (Gesch et al. 2002). While the USGS encourages the use of Level 2 DEMs for surface modeling, there is not a stated accuracy standard for parameters such as slope, aspect, and drainage (Hodgson et al. 2003). Further, when using an elevation dataset for a specific site or application and in regions with varying terrain or vegetation, error may differ (Bolstad and Stowe 1994). Investigations of NED include error assessments of the elevation surface and its impact on the derivative slope and aspect surfaces (Bolstad and Stowe 1994; Holmes et al. 2000; Gesch et al. 2001). More recently, studies have compared the accuracy of NED to that of lidar in areas with varying relief during leaf-on conditions (Hodgson et al. 2003) and for surface hydrologic modeling (Hodgson et al. 2003; Barber and Shortridge 2005). These studies may provide insight into the strengths and weaknesses of the different production methods of USGS DEMs, and how an older dataset with a coarser resolution compares to newer, high resolution lidar elevation data.

Bolstad and Stowe (1994) compared elevation, slope, and aspect derived from a USGS Level 1 DEM to that of a DEM developed using a proprietary stereo-correlation technique developed by a private firm. The Level 1 DEM used was created using automated stereo-correlation of a 1:40,000-scale leaf-off aerial photographs with a Gestalt Photomapper (GPM) and has a reported vertical accuracy of 7 m (target RMSE) with a maximum acceptable RMSE of 15 m. The study compared the DEM elevation values to those of 40 "ground-truth" values collected from National Geodetic Survey (NGS) control points and carrier-phase GPS surveys. Slope values were calculated as a percent using the Horn algorithm, a third-order finite difference method, and compared to slope values measured using a hand-held clinometer.

Significant findings from this study include: 1) the USGS DEM meets the reported accuracy for Level 1 DEMs, with the largest errors found in areas exhibiting the highest and lowest elevations for the study area, 2) on average, the USGS DEM under-estimated elevation, 3) a positive correlation between slope errors and slope, indicates larger errors on steeper slopes. Areas with the steepest terrain are typically forested and, using photogrammetric techniques, fewer postings are collected in forested areas. When interpolation is used to calculate elevation values it tends to smooth microtopography. Bolstad and Stowe (1994) believe that this may be the reason larger slope errors are associated with more rugged terrain . Further, when the terrain is more varied with steeper slopes, field collection methods may be hampered due to stem obstruction, understory vegetation, and other factors, which also lend to larger slope errors between sources (Bolstad and Stowe 1994).

Holmes et al. (2000) used geostatistics to evaluate the distribution of error in a USGS DEM in relation to data collected using a GPS with the intent of determining the affect of DEM error on a landslide modeling application. The USGS DEM used in their study is a Level 2 dataset produced by digitizing contours, either photogrammetrically or from existing topographic maps. In Level 2 DEMs, the maximum allowable error (RMSE) is up to one half of the contour interval, in this case 5 m. Their study area includes about 1000 m of relief, including floodplains, low rolling hills, and mountains.

The researchers (Holmes et al. 2000) did not find a statistically significant correlation between DEM error and terrain attributes, including slope, elevation and roughness. The strongest correlation value was .325 between slope and error. They report that the USGS Level 2 DEMs are generally accurate given the reported error estimates. However, they conclude that accounting for error can greatly enhance the value of an environmental model given that the error in an elevation surface can compound in derivative measures such as slope, aspect, and flow accumulation.

The Collection of Lidar Data

Lidar is an advanced remote sensing technology that is widely used to collect high-resolution terrain data (Lloyd and Atkinson 2002). Mounted to a small aircraft, the lidar instrument uses a fast-firing laser light (typically between 10,000 and 50,000 hertz) to transmit a light pulse to a target on the ground. The round-trip travel time of the pulse and the arrival of the pulse reflection are recorded by the sensor's receiver (Bowen and Waltermire 2002; Lefsky et al. 2002; Barber and Shortridge 2004). In 1999, at least 5 companies manufactured lidar systems and roughly 40 private firms offered lidar data collection services (Baltsavias 1999).

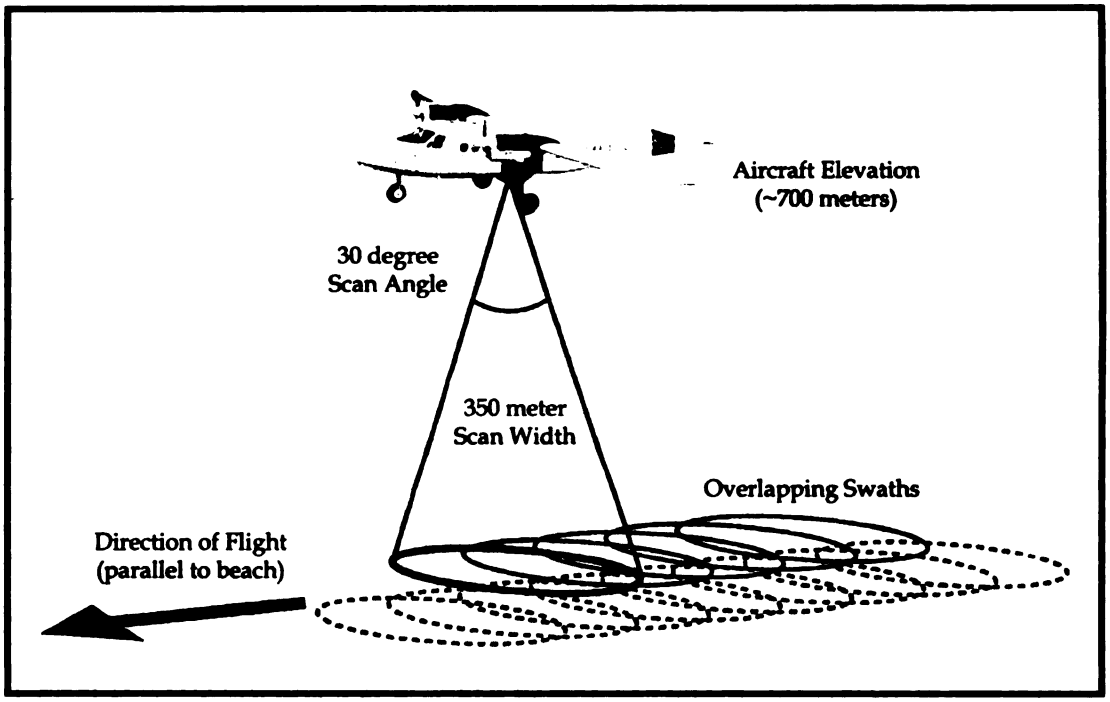


Figure 2-6 Schematic of lidar data acquisition from a twin engine aircraft (modified from Woolard and Colby 2002).

An outline of the ground surface can be captured when the vertical distance between the sensor, located in a level-flying aircraft, and the Earth's surface is repeatedly measured along a transect (Lefsky et al. 2002). The laser scans the terrain surface perpendicular to the flight path, gathering up to five measurements per square meter, although the density will vary depending upon the local topography and the scan angle (Adams and Chandler 2002; French 2003). Lidar can quickly gather spatially dense elevation data within a survey swath hundreds of meters wide (Woolard and Colby 2002; Sallenger et al. 2003; Barber and Shortridge 2004). For coastal mapping applications, the common aircraft altitude is around 700 m, with an individual swath width of roughly half of the altitude (Figure 2-6). The rotating mirror has an elliptical swath pattern with an approximately 30° scan angle (Woolard and Colby 2002; Sallenger et al. 2003).

The lidar system is typically used in combination with instruments for locating the source of the return signal in three-dimensional space (Baltsavias 1999; Lefsky et al. 2002). These instruments include a differential Global Positioning System (dGPS) receivers to obtain the position of the platform, an Inertial Navigation System (INS) to monitor the pitch, roll, and altitude of the aircraft, and angle encoders for the orientation of the scanning mirrors (Baltsavias 1999; Adams and Chandler 2002; Lefsky et al. 2002; Sallenger et al. 2003; Barber and Shortridge 2004). When combined, they may be referred to as an Airborne Laser Scanning (ALS) system. Since an ALS has the potential to determine the 3-D location of the target, it is commonly used for terrain mapping (Adams and Chandler 2002; Barber and Shortridge 2004). Several application specific laser scanning systems have been developed in recent years including the US Army Corps of Engineers' (USACE) Scanning Hydrographic

Operational Airborne Lidar Surveying (SHOALS) system for surveying bathymetry and NASA's Airborne Topographic Mapper (ATM) for global climate change applications (Sallenger et al. 2003).

Post-Processing of Lidar Data

The data returned using an ALS system is a point cloud represented by a series of xyz data points, which describes the location of the observations in three-dimensional space (Figure 2-7). The elevations, or z-values, associated with each point can represent a myriad of features including the ground, buildings, clouds, the canopy, or anything else that the laser pulse may come in contact with (Lefsky et al. 2002; Barber and Shortridge 2004). Although lidar appears to be useful in obtaining data points in leaf- on conditions, a closed canopy can result in significantly fewer ground hits and lower accuracy. Ackerman (1996) estimates an overall penetration rate of 24-29% for coniferous forests and 22-25% for deciduous forests during the growing season. While Cowen et al. (2000) report that in canopy closures of 30 – 40%, 80-90 % of lidar pulses will reach the ground, but when forest cover increases to 80-90% the number of bare ground hits significantly decreases to 10%.

For use in applications that require bare-earth results, data must undergo post-processing to remove points returned from vegetation, buildings, and any other above the ground component (Lefsky et al. 2002). This process involves a combination of highly automated processes, typically filtering algorithms, with some manual correction (Kraus and Pfeifer 1998; Bowen and Waltermire 2002; Lefsky et al. 2002; Barber and Shortridge 2004). Less is known about the affects of a dense understory on the quality of the data set. Research shows that traditional methods of removing points that hit trees and buildings may not work with vegetation below the canopy and a systematic vertical shift may be more appropriate (Pfeifer et al. 2004).

When data are collected by a contractor or engineering firm, algorithms and programs used in post-processing are generally proprietary and vary between firms (Bowen and Waltermire 2002; Lefsky et al. 2002; Barber and Shortridge 2004). Firms are often reluctant to describe their methods in detail and this can present problems for scientists using the data for research (Lefsky et al. 2002). Methods commonly used include some form of linear prediction and/or data segmentation. Linear prediction techniques use knowledge of the terrain to classify points based on their z-values. Kraus and Pfeifer (2001) developed a method they call robust interpolation or robust linear prediction. This technique combines a filtering algorithm with surface interpolation. It uses the residuals, or vertical distance between point values to the surface, to assign a weight to each z-value. The surface is then re-interpolated using the weights and if the points fall outside of a certain value, they are thrown out and the process repeated until all gross errors are eliminated. Lee and Younan (2003) further modified the work of Kraus and Pfeifer (2001) combining linear prediction with adaptive processing to derive bare earth points in vegetated areas. Data segmentation, on the other hand, separates entire features from ground points as a whole instead of point by point, by generating texture maps to identify the corners of buildings and trees (Barber and Shortridge 2004).

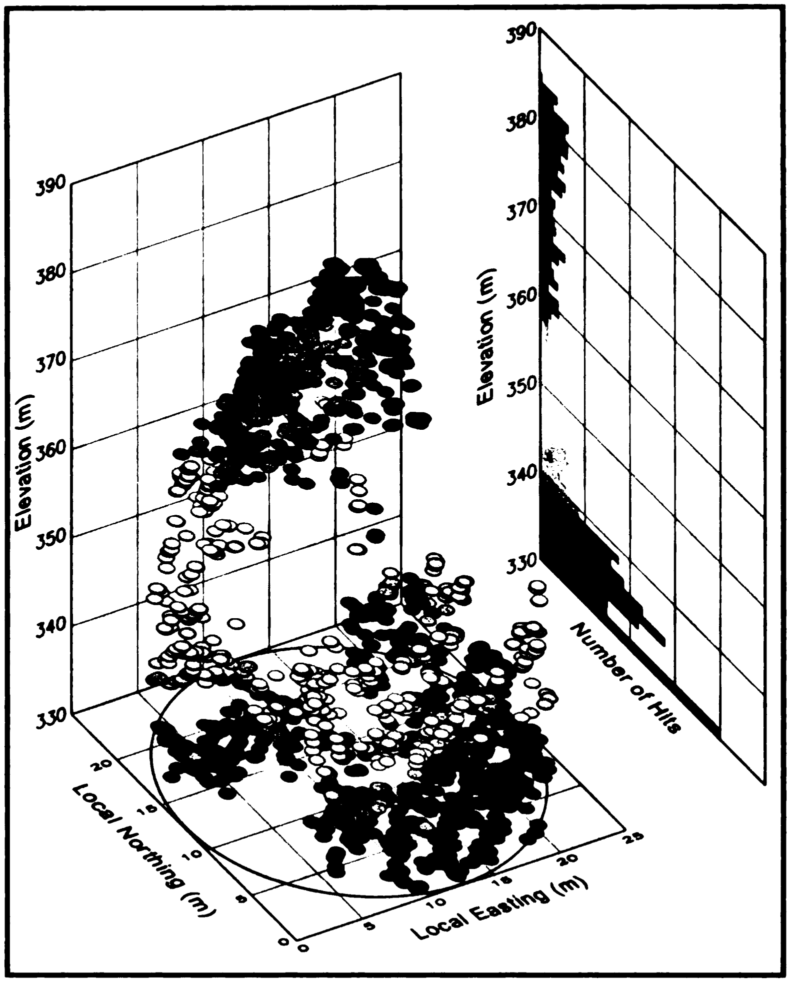


Figure 2-7 Representative illustration of the three-dimensional and vertical distributions of lidar returns within a 25-m footprint (from Lefsky et al. 2002).

A second step in post-processing, which is commonly performed by the user of the data, is interpolation to a raster grid. The data collected via lidar, being a set of irregularly spaced points, requires interpolation if an elevation surface or digital elevation model is to be derived (Lloyd and Atkinson 2002). Research, typically based on cross-validation of the data, shows that there is some discrepancy on which interpolation method of the many available is most accurate. Inverse distance weighting (IDW), ordinary kriging, and universal kriging have proven to result in a relatively high degree of accuracy in the interpolated surface (Lloyd and Atkinson 2002; Woolard and Colby 2002; Rosso et al. 2003).

The interpolation method that will work best is usually dependant on the data itself and terrain characteristics. If large voids exits or the data is generally sparse interpolation with kriging may be the more suitable method (Lloyd and Atkinson 2002; Barber and Shortridge 2004). However, in most cases the simpler IDW method is appropriate and may even be preferable (Rosso et al. 2003; Barber and Shortridge 2004). Although the point data may be carefully edited, once interpolated, it is important to consider that investigations into using lidar for surface modeling show that features such as elevated roadways, bridges, and railroad grades may remain leaving large sinks in the dataset (Barber and Shortridge 2004; Barber and Shortridge 2005).

The Application of Lidar Data

As previously mentioned, the primary disadvantage of ground surveys is the large time and labor expense, while the alternative photogrammetric methods are inaccurate when the ground is under a forest canopy and in areas of low relief and texture, like that of wetlands and coastal dune systems (Lefsky et al. 2002). Lidar is generally regarded as a more accurate, more efficient, and less expensive collection method for the creation of DEMs than these conventional methods (Hill et al. 2000; Bowen and Waltermire 2002; Lefsky et al. 2002; Woolard and Colby 2002; Sallenger et al. 2003; Barber and Shortridge 2004). Published accuracies for lidar data range anywhere from 5 cm to 1 m vertically, with some consensus between 15 and 20 cm, and 1 m to 2 m horizontally (Hill et al. 2000; Bowen and Waltermire 2002; Woolard and Colby 2002; French 2003; Sallenger et al. 2003; Barber and Shortridge 2004; Davenport et al. 2004). The cost of obtaining lidar data ranges from $500 to $1500 per square mile, which is substantially lower than traditional methods (Bowen and Waltermire 2002; Sallenger et al. 2003).

Given the benefits of using lidar data, its use has been and continues to be examined for many purposes. In terms of environmental applications, Lefsky et al. (2002) broadly divide lidar research into three categories: 1) remote sensing of ground topography, 2) measurement of the three dimensional structure and function of forest canopies, and 3) prediction of forest stand attributes. In addition to ecosystem management there are also several commercial uses of lidar including telecommunications, transportation, and urban development, all of which use lidar to map landscape components including buildings, infrastructure, topography, and vegetation (Hill et al. 2000). While these categories may fall short of describing the actual extent to which lidar technology has been applied, the remote sensing of ground topography alone covers numerous scientific fields and applications.

Specific topographic applications of lidar data that have been examined include hydraulic modeling and watershed delineation (Spinney 2001; Bowen and Waltermire 2002; French 2003; Barber and Shortridge 2004), analysis of the volumetric change and erosion of the coastal zone (Krabill et al. 2000; Adams and Chandler 2002; Woolard and Colby 2002; Sallenger et al. 2003; White and Wang 2003), near shore bathymetry and seafloor mapping (Irish and White 1998; Irish and Lillycrop 1999; Guenther et al. 2000), mapping the morphology and distribution of landforms (Rango et al. 2000; Nagihara et al. 2004), land cover classification (Jelaska et al. 2003; Rosso et al. 2003; Davenport et al. 2004), and habitat management (Blott and Pye 2004).

One of the most prevalent topics in the literature is coastal zone mapping. The coastal zone is the zone of interaction between land and ocean where the hydrosphere, lithosphere, and atmosphere are in constant interaction with one another (White and Wang 2003; Gabler et al. 2004). These interactions make it one of the most dynamic environments, operating on a variety of time scales ranging from a few hours to hundreds of years (Krabill et al. 2000). The U.S. coastline stretches more than 20,000 km in length, making remotely sensed data an invaluable resource for monitoring the constant shifts and changes occurring there (Krabill et al. 2000; Sallenger et al. 2003). Further driving the need for research is the heavy anthropogenic pressure placed on this margin when you consider that over half of the U.S. population resides there and the heavy economic cost of severe erosion events (Nordstrom et al. 1990; Bernd-Cohen and Gordon 1999; White and Wang 2003).

Coastal erosion is a topic well suited for the spatially dense lidar data, demonstrated by the ALACE (Airborne Lidar Assessment of Coastal Erosion) project, a joint venture of NOAA's Coastal Services Center, the USGS Center for Coastal Geology, and NASA. Using NASA's ATM, annual surveys of the Atlantic, Pacific, Great Lakes, and Gulf of Mexico shorelines were conducted between the years of 1997 and 2000 primarily to examine and measure coastal erosion (Figure 2-8) (Krabill et al. 2000; Lefsky et al. 2002; Sallenger et al. 2003). In a field experiment called Sandy Duck, researchers used an extensive set of ground measurements to verify that the data collected by the ATM was fit for coastal change applications. The researchers found that the accuracy of the data was sufficient to determine the magnitude of beach erosion following a severe storm event (Sallenger et al. 2003). A similar study was conducted on Assateague Island off the coast of Maryland, to measure the net change in beach morphology between the fall of 1995 and 1996. With the exception of areas affected by overwash, the measurements collected by the lidar system were comparable to those collected using ground survey techniques. Finding a vertical RMSE of 10 -20 cm, this study shows that lidar data can represent detailed beach morphology (Krabill et al. 2000).

While most coastal studies, like the ones described above, have focused on the validation of lidar as a mapping tool or its ability to profile erosion rates and trends, Woolard and Colby (2002) used lidar to assess the volumetric change in sand dunes between 1996 and 1997. They concluded that lidar (interpolated to a 1-2 m resolution grid using IDW) is able to accurately represent the topographic variability found on a dune covered landscape and provides sufficient information to analyze the net sediment change year over year. They suggest that their methods could be used to compare volumetric change from season to season in order to better understand the changes taking place in a dune system.

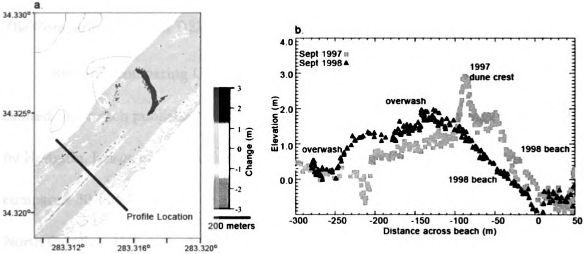


Figure 2-8 Lidar representation of pre- and post hurricane beach topography (right) for a transect on Topsail Island, North Carolina (left) (from the ALACE project in Lefsky et al. 2002).

White and Wang (2003) performed a similar study on barrier islands in North Carolina. Using lidar data from nearly one hundred study sites over three time periods (1997-1998, 1998-1999, and 1999-2000), and grid resolutions ranging from 1 m to 5 m, they were able to assess morphologic changes in the barrier islands from year to year and at different resolutions. They were also able to relate the differences in sediment volume to contrasting management techniques (developed, undeveloped, and nourished). The authors concluded that the increase in development and different management practices greatly affect the coastline's response to storm events from year to year. While both of these studies confirm the value and suitability of using lidar as a tool for evaluating and managing the coastal zone, there is still a need to study the use of lidar derivatives to further characterize topography in this setting.

The Comparison of Lidar DEMs to NED DEMs

Research comparing USGS Level 1 and Level 2 DEMs to lidar datasets is limited. Two such published studies were done by Barber and Shortridge (2005) and by Hodgson, Jensen et al. (2003). Barber and Shortridge (2005) initially intended to compare a 30-m NED dataset to a corresponding lidar dataset for two watersheds in North Carolina. However, the USGS had used lidar data in an update of the DEM for the study area, so a comparable 30-m dataset was obtained from a mosaic of USGS 7.5 minute quadrangles and used in lieu of the NED data. The second study, performed by Hodgson, Jensen et al. (2002), compares elevation data collected by lidar during leaf-on to that of both Level 1 and Level 2 DEMs in relation to slope and land cover.

Barber and Shortridge (2005) were interested in determining whether the use of lidar over traditional data collection methods improved terrain representation and hydrologic modeling. In their study three areas of investigation were considered including: 1) the dense forest cover that is often associated with riparian environments (although all survey methods may be affected by this, lidar is especially sensitive to heavily vegetated areas), 2) production artifacts that are present in any elevation dataset and are of particular challenge to hydrologic modeling; and 3) spatial scale or the difference in resolution between a lidar-based DEM and a medium-resolution datasets. They presented three "testing challenges" to resolve the differences between the two sets of elevation measures in relation to the three areas of investigation, considering whether lidar elevations were more similar to the USGS elevations in upland areas, whether lidar surfaces are hindered by manmade obstructions, and whether the differences between lidar and USGS elevations were a matter of spatial resolution as opposed to production methods.

The authors (Barber and Shortridge 2005) found that post-processing of the lidar data, in this case, was successful in removing man-made obstructions from the data set, and although there were numerous sinks, which are a reflection of non-ground hits, they were very small in area. Findings on the issue of scale versus source were less conclusive and the authors determined that this is likely dependent upon the terrain derivative in question. Slope, for example, is greatly affected by resolution and not as affected by source. The paper concludes that the while the benefits of using lidar rather than DEMs like NED are marginal, some operations may be sensitive to resolution and therefore lidar may be a better choice, particularly in areas of low relief (Barber and Shortridge 2005).

The aim of the research conducted by Hodgson et al. (2003) was to determine the overall accuracy of elevation and slope measures of four different DEMs, including one derived from lidar and two from NED (Level1 and Level 2). One further goal was to establish if land cover and terrain variability impacted the accuracy of those measures. Elevations in their study area range from 44 to 136 m, slopes range from 0° to 14°, and land cover is primarily deciduous and pine forest (60%). Sampling took place along 23 transects ranging in length from 100 to 840 m, and over 1,470 elevation points were collected using GPS and conventional surveying techniques. Slope was computed as the degree of slope between adjacent transect points, with the average distance between points being 6.88 m. Land cover data was collected at 1,195 of the survey points.

Of the four elevation surfaces studied, only the USGS Level 2 DEM fell within its reported accuracy range of 1 to 2 m. Although, the accuracy of the lidar data surpassed that of the Level 2 DEM, it more than quadrupled the RMSE that is generally advertised for the product (93 cm versus 15 to 20 cm). In forested areas the error jumped to 113 - 122 m, keeping in mind that the data was collected during the growing season. The researchers found that although elevation errors were less related to slope than they were to land cover class, error and slope were positively correlated. However, they conclude that if the environmental application in question is in an area of rugged terrain, greater elevation errors should be expected. Further, the authors recommend using either a lidar or USGS Level 2 DEM over other elevation products if an application requires accurate slope measures but, based on the topography of their study area, are unable to provide evidence that either product is suitable for such uses (Hodgson et al. 2003).

III. STUDY AREA

The study area for this investigation is a 300-meter wide strip of land, approximately 3.25 sq km in total area, spanning the shoreline of Lake Michigan in Lake Township (T6S, R20W), Berrien County, MI (Figure 3-1). Berrien County marks the southern reach of the eastern shore of Lake Michigan in Michigan and contains over 1,600 hectares of the state's 32,000 hectares designated *critical dunes*. Of those 1,600 hectares, approximately 900 are in Lake Township (MDEQ Atlas of Critical Dunes 1989). Also located within the township are Warren Dunes State Park, spanning a little more than 4 km of the shoreline, and Weko Beach recreation area.

Land use in the rest of the township is primarily residential with some industrial development, most notably the Cook nuclear power facility. Combined there are approximately 340 taxable parcels of land on record, with at least some portion of the lot contained within the study area. Although the *Sand Dunes Protection and Management Act* (Michigan State Legislature Act No. 451, Public Acts of 1994) allows the local unit of government to assume management of the dunes, at this time the MDEQ manages all of the critical dune land in Lake Township.

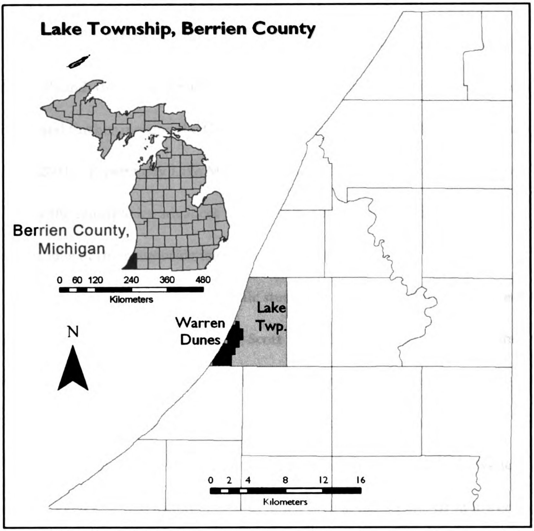


Figure 3-1 Map of the study area, Lake Township, Berrien County, MI.

History

Throughout the Pleistocene, Michigan's landscape was repeatedly overrun by glacial ice with initial ice flow into the region controlled by the local geology. In the southern, lowland region of the Great Lakes, which includes the Lake Erie and Lake Michigan basins and most of the Huron and Ontario basins, the ice was channeled over top weaker sedimentary rocks into bedrock valley systems already present in the region (Larson and Schaetzl 2001). Subsequent ice flows into the area followed a similar path and, therefore, a highly lobate front developed each time ice advanced, scouring and further incising the Great Lakes' basins (Eschman 1985; Larson and Schaetzl 2001). Present day Lake Michigan fills the large depression that was heavily eroded by the multiple ice advances and retreats, and last occupied by the Lake Michigan lobe of the Laurentide ice sheet (Eschman 1985).

Early in the history of geomorphic studies on the coastal dune complexes of Lake Michigan, researchers (Dow 1937; Scott 1942; Olson 1958d) recognized the relationship between the elevation of the lake and the adjacent landforms. Throughout the Pleistocene and into the Holocene, the level of Lake Michigan has fluctuated in response to ice advances, retreats, and isostatic changes in elevation. Some have even reported differences of as much as18 m higher and 61 m lower than the mean present day lake level of 177 m (Chrzastowski and Thompson 1992). On a smaller scale, lake level may also vary from year to year depending on temperature and precipitation and over time, patterns develop given that any year's water balance is dependant upon that of the preceding years (Olson 1958d). Both Dow (1937) and Scott (1942) hypothesized that fluctuations in lake level were the dominant control of the sand supply to the dunes on the eastern shore of Lake Michigan. Olson (1958d) further expanded this idea, and developed a cycle of dune development controlled primarily by the rising and falling of the lake level.

Although Olson's (1958d) work remains the foundational model of foredune development in the region, work by Arbogast and Loope (1999) and Van Oort et al. (2001) provided evidence in the form of radiocarbon dates from buried soils within the dunes that the evolution of the massive dunes located atop lacustrine sediments is directly linked to the historical lake level curve. Prior to this research, Dorr and Eschman (1970) and Eschman (1985) hypothesized that major dune development along the lake largely occurred during the Nipissing Phase of the Great Lakes (~ 6000 to 4000 years B.P.) when lake levels peaked at 184 m. More recent studies of Lake Michigan dune fields (Arbogast and Loope 1999; Loope and Arbogast 2000; VanOort et al. 2001; Arbogast et al. 2002), however, provide evidence that dune building began during that high stand, but that the most significant growth occurred later at episodic intervals. Van Oort et al. (2001) studied large parabolic dunes at Van Buren State Park approximately 60 km north of the study area, and their results also indicate that there were several periods of dune growth marked by intervals of stability when soils developed. A rapid decline in lake level following the Nipissing transgression led to an increase in sand supply and to a dominant period of dune building between ~ 4000 and 2500 years B.P. This time of rapid growth was followed by a long period of stability sufficient for soil development until the dunes remobilized again between 1140 and 790 cal. yr B.P.

In studies conducted elsewhere on the eastern shore it seems that while receding lake levels account for the onset of the initial growth period, later periods of dune growth (~ 3200, 2400, and 900 years B.P.) are often associated with intervals of higher lake levels (Figure 3-2; Arbogast et al. 2002). One hypothesis is that the presence of a broad beach and foredunes during lower lake levels protect the large parabolic dunes and a narrow beach and the absence of foredunes during high lake levels subject the dunes to increased wave erosion (VanOort et al. 2001). Given the similarities of dunes in Van Buren State Park and those found in the study area, this research can be used to infer a history of the large parabolic dunes found in the study area.

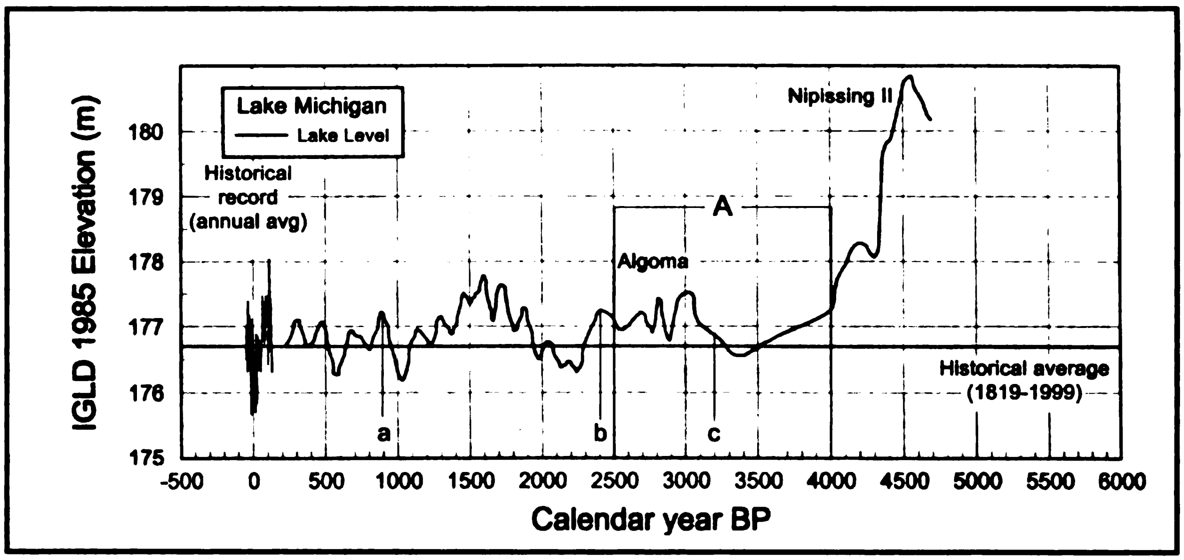


Figure 3-2 Graph showing the residual lake level curve of Lake Michigan as determined by Baedke and Thompson (2000), plotted with the historical average. A. marks the dominant time period of dune building along the eastern shore, while a, b, and c are used to note times of dune growth associated with higher lake levels (Arbogast et al. 2002; modified from Baedke and Thompson 2000).

Dune Topography

Dunes within the township are classified as barrier dunes by the MDEQ (1994), a term first defined by the MDNR (1976) and later refined by Buckler (1979) to describe an assemblage of dunes that form a physiographic boundary marked by great relief between the aesthetically pleasing shore zone and the interior. In more specific terms, Arbogast et al. (2002) noted that the dunes along the Lake Michigan shoreline can generally be divided into two universal categories 1) *foredunes* (~ 5m high), which are relatively small sublinear dunes, and 2) *large dunes* (> 20 m high), which tend to be parabolic in form and occur within many well-developed dune fields along the lake. The largest dunes in the study area are parabolic dunes over 50 m high. Other dune features that exist within the study area include but are not limited to blowouts, dune flats, and interdune lowlands (Buckler 1979). The margin of the study area is lined by small hills and low ridges of glacial till that are part of the Lake Border morainic system (Figure 3-3; Schaetzl 2006). Dunes are absent in areas where moraines approach the lake or where the shore is composed of glacial till or outwash (VanOort et al. 2001).

Where not altered by development, foredunes, parabolic dunes, blowouts, dune flats, and interdune lowlands are found throughout the study area and give dune fields the distinction of having very diversified topography. Elevation in Lake Township ranges from a low of 176 m at the shoreline to a high of 233 m on the largest dunes. The mean elevation of the study area is near 190 m, with most heights not varying more than 10 m from the mean (σ = 9.04 m). The gradient ranges from flat (0% or 0°), typically near the shore, on dune flats, and interdune lowlands, to >300% (70°) on the steepest slopes (all values were calculated using Arc Info and lidar elevation data).

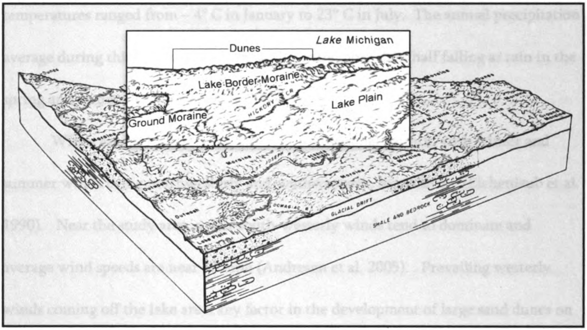


Figure 3-3 Physiography of Berrien County as depicted in the Soil Survey of Berrien County. The inset highlights the landforms found within and around the study area including the dunes, Lake Border morainic system, and glacial lake plain that underlies most of the study area (Larson 1980).

Climate

A unique characteristic of coastal locales to the east of Lake Michigan is a maritime influence on climate despite the continental location. In these areas, climate is classified as mixed marine, continental, whereas areas further inland are classified as humid continental, hot summer, which is noted for its large range in temperature (Eichenlaub et al. 1990; Gabler et al. 2004). Average annual temperatures for the area range from about -5° C in January to 21° C in July, and annual precipitation totals are near 81 cm (Eichenlaub et al. 1990). The closest coastal weather station to the study area is located approximately 20 km to the north in Benton Harbor, MI. In the 30-year period from 1951 to 1980, average daily temperatures ranged from – 4° C in January to 23° C in July. The annual precipitation average during this time was 92.7 cm with a little more than half falling as rain in the spring and summer (Andresen et al. 2005).

Winds are multidirectional and follow a seasonal pattern, with winter and summer winds being northwesterly and southwesterly, respectively (Eichenlaub et al. 1990). Near the study area, south-south-westerly winds tend to dominate and average wind speeds are near 10 mph (Andresen et al. 2005). Prevailing westerly winds coming off the lake are a key factor in the development of large sand dunes on the eastern and southern shore of the lake (Cowles 1899; Gabler et al. 2004). Climate reconstruction during the time of major dune building episodes shows both a slight decrease in precipitation and an increase in seasonality from that of the early Holocene, however, when compared to climate reconstruction elsewhere in the region, this climate change is negligible (Davis et al. 2000).

Land Use / Land Cover

According to the most recent available data (MDNR land use / land cover (LULC) classification of 1978), LULC in the study area includes areas of deciduous forest (code 41), residential (code 11), transportation, communication and utilities (code 14), and barren land (code 7) (level II categories). These classifications are based on the dominant LULC type in an area no smaller than one hectare.

An on the ground inspection of LULC classes at individual sample sites also revealed areas of dune grass and mixed forest including pines, however, there was no attempt to determine whether or not these areas met the minimum mapping unit of one hectare. Further, the limited width of the study area from the shore (300 m) combined with the selection of 46 sample points from within that area greatly limited the number of possible LULC classifications found at any of the sample locations.

In general, the distribution of land cover types moving from the shore inland is beach, barren land (including sand other than beaches), and forest (both hardwoods and pine), which is the dominant land cover on many of the largest and most well established dunes. Land use outside of the state park and recreation area includes many single family homes, which are scattered throughout, and a large nuclear power facility located in the northern half.

Soils

The predominant soil series within the study area is Oakville fine sand, which is found primarily in forested dune areas. A narrow margin of gently sloping beach stretches along the coast with projections of dune land extending back into the forested areas (Larson 1980). Areas mapped as beach remain in the swash zone and are repeatedly overrun by wave action. The primary use of the beach is recreation and that, combined with the absence of notable relief, made the areas mapped as beach an unfavorable choice for sampling.

The Oakville series corresponds with the deciduous forest land cover and developed areas (mainly residential). The Oakville series is well drained with rapid to very rapid permeability and weakly developed with A / E / B / C horizonation. In some areas, this soil is prone to instability and its best use is for woodland and wildlife habitat. In those places where it has been used for building construction, slope instability can be controlled with retaining structures (Larson 1980). Further, when compared to areas mapped dune land, in places where Oakville is found, it is generally safe to assume that the soil is older than its bare sand counterpart and has moderate to fairly stable (slopes < 18%; Arbogast 2004).

The sediment found in areas classified as dune land was derived largely from glacial drift and is composed chiefly (~90%) of quartz sand, which is extremely resistant to erosion (Santer 1993). Due to the sorting ability of wind, the dune sand is remarkably uniform in size when compared with beach sand. The sand is still actively shifting and has little to no protective vegetation cover and, therefore, soils are yet to develop (Larson 1980; Santer 1993). Vegetation where present is primarily dune grass with some trees and woody shrubs. In the areas with stabilizing vegetation, it is likely too recent for a soil profile to develop (Larson 1980).

IV. METHODS

The methods used in this study are designed to determine which digital terrain attributes, or combination of attributes, when used in a linear regression model will most closely predict true slope. Further, I will use data from two different sources, lidar and NED, to determine whether the source of the data influences the outcome of the model. My hypothesis is that lidar, when used in conjunction with calculated slope and land cover, will more accurately predict true slope than a model that uses NED and the same variables.

The first steps toward answering my research questions and testing my hypothesis are the collection of field data, primarily slope, and processing of digital data, including elevation, slope, and land cover. Once prepared, variables from each data set will be used in the design of multiple regression models. This chapter begins with an account of the methods used to collect field data, followed by a description of each digital data set and its preparation. Once this discussion is complete, I will conclude with an overview of multiple regression.

Collection of Field Data

Perhaps the most important variable in the regression equation is the set of observations (or the sample) that is used to construct the model. Generally a large sample size representative of the entire population creates an equation that is more stable or replicable across samples from that same population (Osborne 2000). While there are many different opinions on what number of observations is the appropriate amount for creating a model, several authors (Park and Dudycha 1974; Cohen and Cohen 1983; Tabachnick and Fidell 1996) agree that the size of the sample should be based on the number of independent variables used. The sample size used in this research is 46, which is a small number by some standards, but based on the literature should be sufficient to create a model using up to three predictor variables (Tabachnick and Fidell 1996; Osborne 2000).

Sites selected for field sampling needed to accurately represent the entire study area, which includes everything from expanses of level ground to near vertical slopes, but at the same time be suitable for the purpose of this study (to predict slope). Prior to entering the field 100 sample points were generated using a random point generator in Arc Map. Once in the field, a point was discarded if it was not accessible due to property rights or terrain conditions, such as dense vegetation combined with steep slopes, too close to the water line (in the swash zone), or within 5 meters of another sampling point.

Property rights were a much bigger issue than expected, specifically with respect to private roads that limited access. As a result I was unable to select points distributed over the entire study area. Given the field limitations encountered, I was ultimately able to sample 46 sites. At each of these sites, I collected three terrain attributes: slope, land cover, and elevation. Slope is the dependant variable in the multiple regression model, while elevation and land cover were collected only for assessing the accuracy of their respective digital data sets.

Figures 4-1 and 4-2 are maps showing the locations of the sample points in the northern and southern extents of the township. A majority of points in the northernmost cluster (Figure 4-1) were collected in a private residential area where I was granted access by a property owner. All points to the south of this area (Figures 4-1 and 4-2) are located either on publicly owned Weko Beach in the central part of the study area or on state land in the south, which includes Warren Dunes State Park. Large gaps represent areas were access was restricted by land ownership or terrain conditions (i.e., dense forest and/or understory combined with slopes too steep to traverse).

Vertical and horizontal measurements were collected with a Trimble Pathfinder Pro XRS GPS unit, which uses both point averaging and built-in, real-time differential correction. When point averaging of approximately 100 observations and real-time differential correction are used, the documented error of both measurements is reduced to sub-meter (Trimble Navigation Limited 2005).

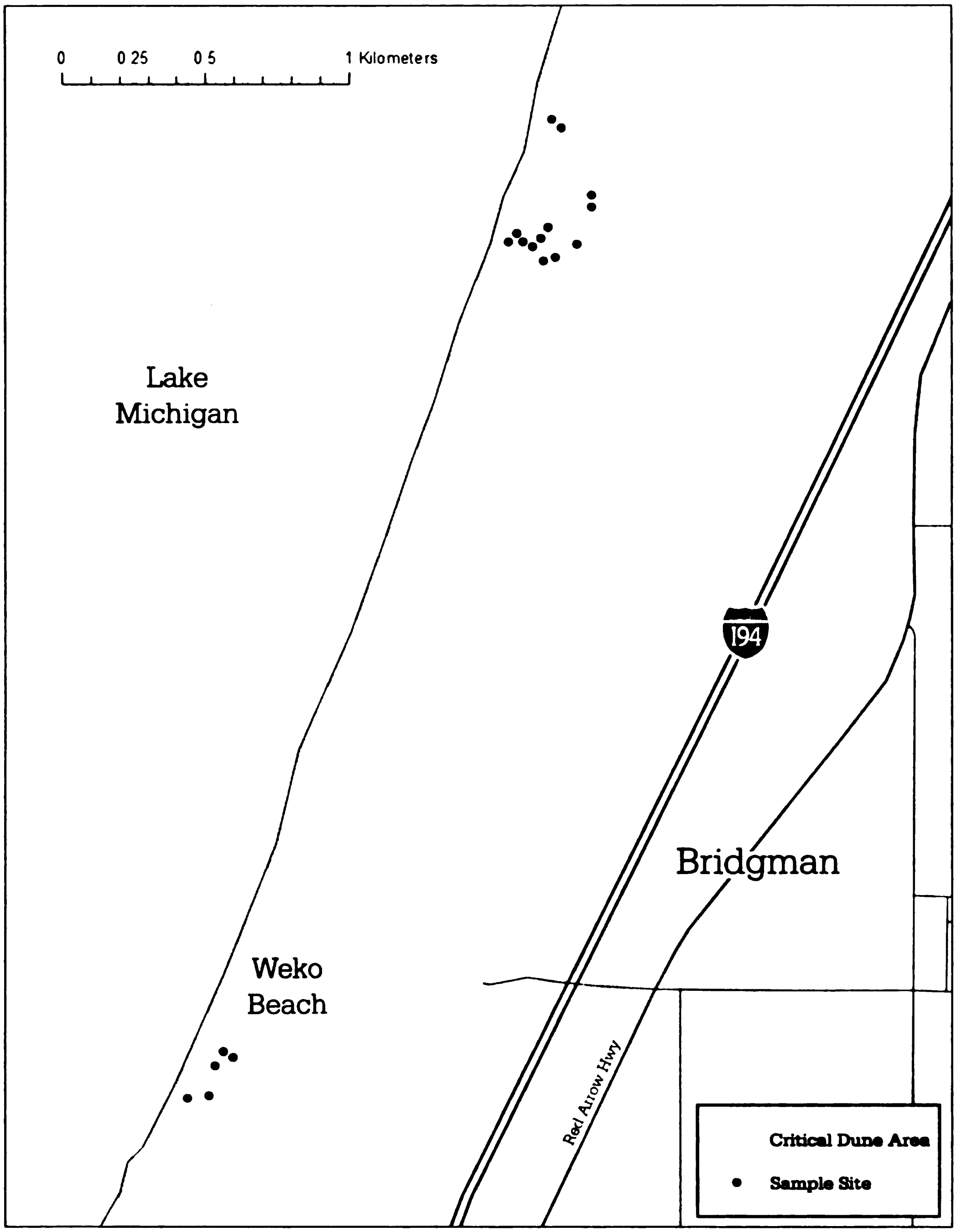


Figure 4-1 Map of the sample points collected in the northern half of Lake Township, Berrien County, MI.

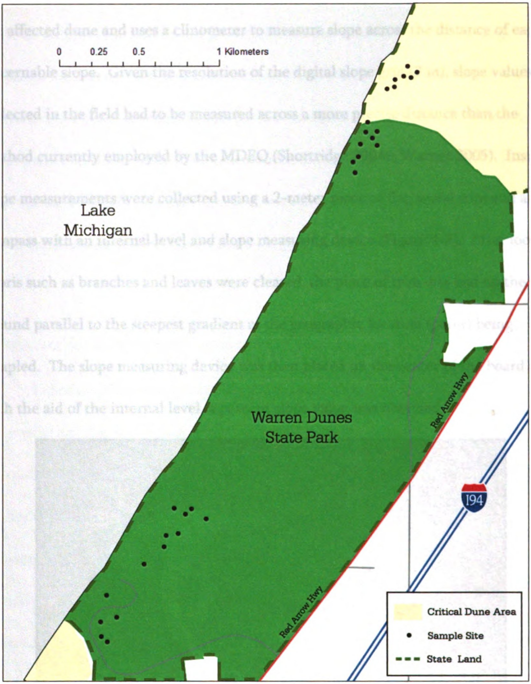


Figure 4-2 Map of the sample points collected in the southern half of Lake Township, Berrien County, MI.

When conducting a field inspection, a MDEQ agent locates visual breaks in the affected dune and uses a clinometer to measure slope across the distance of each discernable slope. Given the resolution of the digital slope grid (2 m), slope values collected in the field had to be measured across a more precise distance than the method currently employed by the MDEQ (Shortridge 2004a; Warner 2005). Instead, slope measurements were collected using a 2-meter piece of flat, wood trim and a compass with an internal level and slope measuring device (Figure 4-3). After loose debris such as branches and leaves were cleared, the piece of trim was laid on the ground parallel to the steepest gradient at the geographic location (point) being sampled. The slope measuring device was then placed on the center of the board and, with the aid of the internal level, a percent slope value was obtained.

Figure 4-3 Photograph showing the method and equipment used to collect slope measurments in the field.

The horizontal (point locations) and vertical (elevation) measures collected with the GPS were tabulated and projected from the unit into a database using Pathfinder software. As a projected database of values, the elevation measures were converted to a point coverage in Arc Toolbox and the corresponding slope and land cover values collected in the field entered into the attribute table manually. The field data was then joined to the attribute table containing the NED and lidar slope and elevation measures and the NLCD based on geographic location.

Selection of the Independent Variables

Osborne (2000) stated that, when a multiple regression equation is designed for prediction, the goal is to arrive at the model that best estimates the value of the dependant variable at unsampled locations, regardless of the relationship it has with the independent variables. Although the variables are not required to make conceptual sense, theory is quite helpful for identifying which predictor variables should be included in the equation (Osborne 2000). I chose to go with the latter approach and chose three independent variables, elevation, calculated slope, and land cover, based primarily on the potential relationship each has with the dependant variable.

The influence of land cover on remote collection methods has long been and will continue to be the subject of much research (Hodgson et al. 2003). However, more recent studies (Bolstad and Stowe 1994; Bowen and Waltermire 2002; Hodgson et al. 2003) have tended to focus more specifically on the relationship that the collection method has with elevation, land cover, and slope to gain a better understanding of how in situ terrain attributes affect the quality of a digital elevation model and its derivatives.

For example, research has found that both elevation error and modeled slope error increase when slopes on the ground are steeper (Chang and Tsai 1991; Bolstad and Stowe 1994; Hodgson et al. 2003). If the terrain surface is variable, then any horizontal errors in position created by the lidar instrument will typically result in vertical (elevation) errors (Hodgson et al. 2003). A major part of post-processing raw points collected with lidar includes identifying ground returns to weed out those from other objects. This typically involves a combination of highly automated processes, such as filtering algorithms, with some manual correction (Kraus and Pfeifer 1998; Bowen and Waltermire 2002; Lefsky et al. 2002; Barber and Shortridge 2004). In addition to affecting the instrument, sloping terrain can also make it difficult for automated weeding algorithms to reliably identify ground returns, which will result in increased error in the lidar dataset. If rugged terrain is found in conjunction with multi-story vegetation, the increase in error can be even more significant (Hodgson et al. 2003).

Research has found a positive relationship between true elevation and elevation error and true slope and slope errors in USGS level 2 DEMs as well (Bolstad and Stowe 1994). During the production process, level 2 DEMS, like the one used in this study, are checked for consistency in slope and often smoothed. This process can remove naturally occurring high frequency terrain variability (Hodgson et al. 2003).

Land cover may also play a role in the accuracy of both data sources for a few different reasons. First, terrain is generally more varied and rugged on steeper mountain- and hillsides, making point estimation more difficult, which leads to error in the interpolated elevation surface. This error then propagates in the derivative slope calculations. Second, larger errors are expected in forested terrain due to a lower posting density. For example, in the lidar data set, many of the returns (z-values) that would be present on barren land would be removed because they hit the canopy, and if not they would remain in the dataset. In USGS DEMs, a dense forest cover makes stereo-correlation more difficult due to the lack of clearly defined features such as roads and buildings, resulting in fewer or less accurate elevation measures (Bolstad and Stowe 1994). Land cover patterns in the study area may also play a role, as some of the oldest and largest dunes tend to be forested, while many of the younger, smaller dunes are colonized by dune grass if at all. While collecting field measurements, land cover was documented at each site and any extenuating conditions such as nearby elevated roadways that could affect the accuracy of elevation or slope measures were noted for further analysis.

It is known that DEM errors generally exhibit a distinct spatial distribution or pattern, and in most cases are not random (Fisher 1998; Liu and Jezek 1999;

Shortridge 2003; Pfeifer et al. 2004). Given that errors are systematic and influenced by patterns of land cover, elevation, and gradient, it may be possible to determine actual slope if land cover, elevation, and calculated slope are known variables.

Based on the source of the data, either lidar or NED, the elevation and slope data will be used to construct separate models. Further, from each of these sources slope was calculated in two different ways, the Horn algorithm and maximum gradient. The result being two models (lidar and NED) with four potential independent variables for each model: elevation, Horn slope, maximum gradient slope, and land cover. All of the independent variables in the model were obtained by remote sensing, reducing the time and cost of data collection and with the exception of lidar data, each dataset is available at present in a digital format to all potential users. In the near future it is highly probable that lidar elevation data will also be widely available to users. The following sections outline the steps taken to obtain and, where necessary, modify the source data of the independent variables used in the multiple regression models.

Lidar Elevation Data

The lidar point data was collected and processed in April of 2001 by Woolpert LLP, a private firm that specializes in the collection of spatial data, for distribution by the Detroit District, USACE. In its raw form, the data was three individual text files in an x,y,z format with a documented horizontal accuracy between 1 and 2 meters and a vertical accuracy of approximately 15 cm, suitable for use in applications that require bare-earth elevation measures (Lidar FGDC metadata). The data was converted from the text file tables to point shapefiles in Arc Toolbox and merged into one file in Arc Map. It was then reprojected from the Universal Transverse Mercator (UTM) projection into the Michigan GeoRef projection based on the North American Datum of 1983 (NAD83).

In contrast to the 30-meter NED data, which comes in the appropriate format (raster grid) for calculating terrain derivatives such as slope, the lidar data are a cloud of irregularly spaced points and interpolation is required to create a continuous elevation surface (Lloyd and Atkinson 2002; Barber and Shortridge 2004). Interpolating to create a grid involves the selection and use of an interpolation algorithm. Given a set of discrete points, the most appropriate spatial interpolation algorithm will be the one that best represents the known values and is able to predict the elevation values where they are unknown (Lam 1983). For this study I selected inverse distance weighting (IDW). It is a relatively simple approach that has been shown to result in minimal error in the interpolated surface (Lloyd and Atkinson 2002; Rosso et al. 2003; Barber and Shortridge 2004). The formula used to interpolate the value at an unsampled location is expressed as:

a.) $$wd = \frac{1}{d^k}$$ b.) $$Z = \frac{\sum wd_i Z_i}{\sum wd}$$

The first equation, a., is used to calculate the weight given to each known point where ***wd*** is the weighting factor applied to the known value, ***d*** is the distance from the known value to the unknown value, and ***k*** is a user-selected power factor. Equation b. uses ***wd*** to determine the value of the interpolated point, where Z_i is the known value and Z is the value of the interpolated point.

IDW is an exact interpolator, meaning that all known elevation values remain in the output surface (Lam 1983; Shortridge 2004b). The IDW equation requires that the user select two variables: *k* and *n*. The variable denoted *k* is equal to the power of the weight as a function of distance and the variable *n* is equal to the number of nearby known elevation values that will be used to compute the unknown elevation at the center of a grid cell. As the value of *k* increases, the influence of nearby points also increases and that of distant points decreases. Likewise, as the number of neighbors increases, the resulting elevation surface becomes smoother or less variable. Both *n* and *k* have a large impact on the output surface and therefore each must be chosen carefully. Cross validation was used to determine the power and number of neighbors that would minimize the RMSE in the interpolated grid (Figure 4-4; Lloyd and Atkinson 2002; Shortridge 2004b). In essence, the error introduced during the interpolation process can be ascertained using only the dataset. This method is favorable when working with a data source for which no other available dataset is known to be more accurate (Shortridge 2004a).

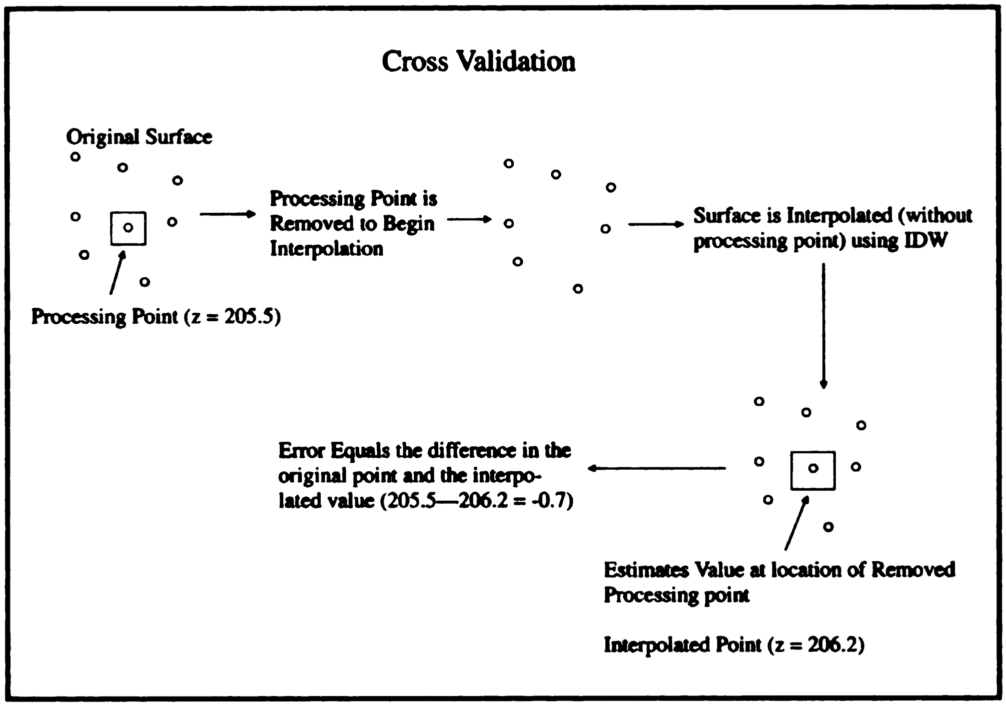


Figure 4-4 Schematic illustrating the cross validation process.

In order to use the cross validation algorithm, a representative sample of the dataset had to be selected due to limitations on computing power and time. A sample of 20,000 points was selected based upon a comparison of maximum elevation, minimum elevation, mean elevation, and standard deviation across the point sets (Table 4-1). The points selected depict a small, but representative area of the 400,000 total data points.

Next, using an R-GUI (Chambers 1991) source code written by Shortridge (2004a), I created a matrix in which the x-axis represents the power and the y-axis the number of neighbors used by the IDW algorithm. In R-GUI statistical software, I ran the code on the subset of points using a sequence of powers from .05 to 4.05, in increments of .50, and a sequence of neighbors from 1 to 10. The output was a 9 x 10 matrix of values equal to the RMSE of each combination. In this case, the resulting matrix showed that using a power of 3.05 and 9 neighbors would minimize the combined RMSE in the surface interpolated with IDW (RMSE = 39.2 cm).

Table 4-1 A Comparison of the descriptive statistics for the point subset to those of the complete dataset.

	Subset	Dataset
No. of cases	25,107	334,861
Range	175.7-232.8 m	178.3-232.8 m
Mean	190.86 m	194.57 m
Standard Dev.	9.04 m	10.43 m
Variance	81.78 m	108.79 m

In addition to the weight and number of neighbors, interpolation also requires the user to select the appropriate cell resolution of the output raster grid. As with the interpolation method, the level of error in a surface can vary greatly with different grid sizes (Smith et al. 2003). I employed a second code written by Shortridge (2004a) in R- GUI (Chambers 1991) to compute the nearest neighbor statistics of the point subset. This enabled me to determine an appropriate cell size based on the frequency interval of each data point within the representative sample. The results of the nearest neighbor operation showed that the distance between the points ranged from .326 m to 10.73 m, with a mean spacing of 1.85 m. Smith et al. (2003) suggest that the optimal resolution for minimizing error is that which is closest to the original point spacing. Therefore, based on the nearest neighbor statistics and data presented in that study as well as data storage and computation requirements, I selected a 2 x 2 m cell resolution.

The final interpolation from a point shape file to a raster grid was performed in Arc Map and the three individual grids were combined using the *merge* command in Arc Info. Figure 4-5 shows the steps taken to process the lidar data beginning at data collection and ending with the creation of the slope grids.

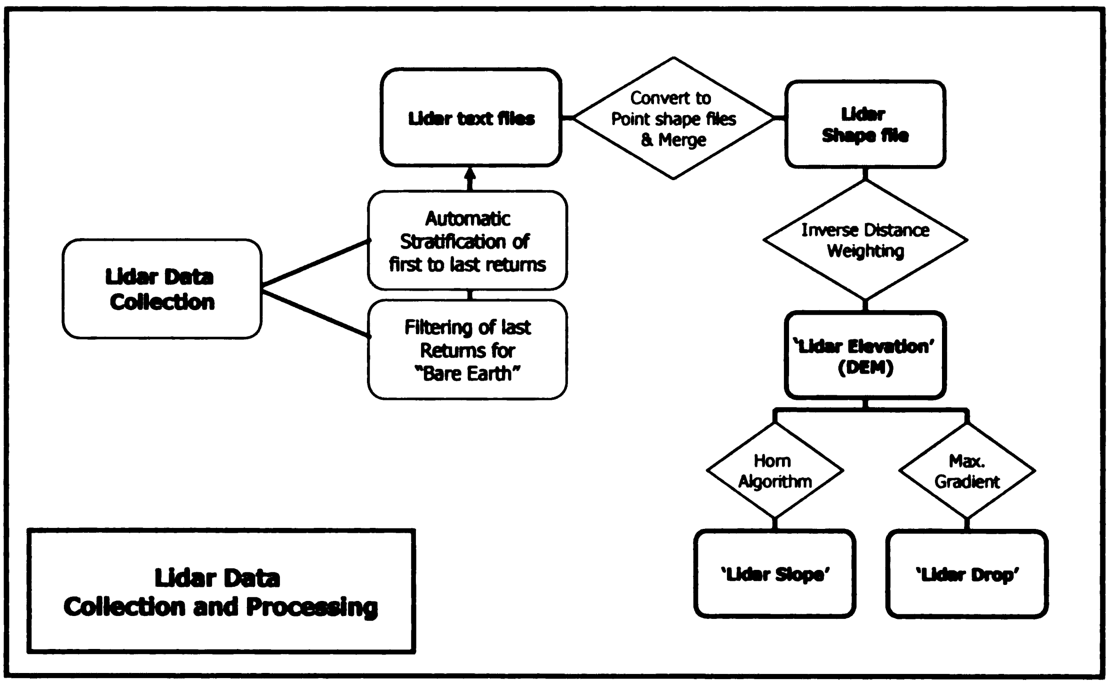


Figure 4-5 Flowchart illustrating the steps taken to create the lidar variables.

30-meter National Elevation Data

The 30-meter elevation data was collected by the USGS EROS Data Center and originally came from the USGS NED website. It is classified as a level 2 dataset, comprised of the best available elevation data (7.5 minute resolution). The data, which is available for all Michigan counties on the Michigan Center for Geographic Information (MICGI) website, underwent post-processing at MSU's RS&GIS for use in their proprietary mapping software, Michigan MapImage Viewer. Processing steps included clipping the dataset to match the extent of Michigan counties, merging the county data to create a statewide mosaic, and projecting the data in the Michigan GeoRef projection, NAD83 datum. After processing, the most recent publication date for the data is 2002.

Although the use of NED is widespread, one big disadvantage of using pre-processed DEMs is that, in general, little information is provided to the end user in terms of data collection, processing, and error distribution (Holmes et al. 2000). It is known from the metadata that the DEM used in this study was produced using manual profiling of photogrammetric stereomodels, derived from either digitized cartographic map contour overlays or scanned National Aerial Photography Program (NAPP) photographs. No estimate of accuracy is given apart from the USGS level 2 global accuracy standard of one half of the contour interval of the source DEM, which could not be found (Berrien County –NED FGDC metadata). A spot check of vertical accuracy at the 46 points sampled using a Trimble Pro XRS GPS unit revealed that, while the lidar dataset fell within its reported accuracy range, the error in the NED DEM was much higher than what is expected even in a level 1 DEM (Table 4-3). Overall, the NED elevation values at these points consistently overestimated the elevation by more than 10 m compared to both the GPS data and the lidar data. Based on the production method and the terrain, there are several possible reasons why the DEM error is larger than expected. These reasons, as well as the implications for the derivative slope values, will be discussed further in the following chapter.

The decision was made to use NED as the primary source of 30-meter elevation data over elevation data collected more recently (2001) by the Shuttle Radar Topography Mission (SRTM). For the end user, there are potential benefits of using SRTM data including its newness (although NED employs the "best available" data upon upgrades), an increase in vertical accuracy, and one common collection method (Gesch et al. 2001). However, one of the objectives of this research is investigate the relationship between land cover and two different collection methods and while SRTM technology is different than that of lidar, the reflection of the SRTM radar signal off of a tree canopy is reminiscent of the reflection of the lidar laser. Further, the systems used for lidar and SRTM data collection, ALS and Synthetic Aperture Radar (SAR), are complimentary to one another when used in conjunction (Sun et al. 2003).

Radar (SAR), are complimentary to one another when used in conjunction (Sun et al. 2003).

Table 4-2 Descriptive statistics of the elevation datasets

	Elevation	Lidar Elevation	NED Elevation
Minimum	176.87 m	177.74 m	189.38 m
Maximum	218.16 m	213.81 m	221.71 m
Mean	192.15 m	191.44 m	206.55 m
Standard Dev.	9.45 m	8.48 m	7.47 m
Variance	89.28	71.97	55.80

Table 4-3 Error statistics of the lidar and NED datasets

	Lidar Elevation (m)	NED Elevation (m)
RMSE	1.66 m	17.91 m
MAE	1.18 m	15.09 m
MAX	5.43 m	36.75 m
MIN	2.7 cm	35.4 cm

Creating the Lidar and NED Slope Grids

On a natural landscape, slope is commonly measured in the direction of the maximum perceived elevation gradient. In contrast to calculating slope in nature, however, modeling slope using a digital raster model is a bit more complicated. The direction of the steepest slope on a landscape and the straight line slope between grid cell centers most often do not match up. The fastest and simplest approach in Arc Info determines the maximum change in elevation between the processing cell and its eight surrounding neighbors, and divides this difference by the distance between the centers of the two cells (Shortridge 2003). The maximum gradient approach is not commonly used in applications that require slope measures because it is sensitive to error, creates a rough surface, and generally achieves poorer results than alternate methods.

In this study, the slope values for both elevation surfaces were calculated using both maximum gradient, the *drop* option when using the *flowdirection* command in Arc Info, and the Horn algorithm or *slope* command in Arc Info. The Horn algorithm calculates slope as the average change in elevation from north to south (rise) and east to west (rise), divided by the distance between cell centers in a 3 x 3 cell neighborhood (Figure 4-6). Although other methods of calculating slope exist, studies have shown that using a third-order finite difference approach like the Horn algorithm is best in variable terrain (Bolstad 2002). The output is a grid of slope values measured as percent slope (the unit used by the MDEQ) where values can range from 0 where elevation is constant across all cells, to values that approach infinity (or 90°). At 100% slope, the rise is equal to the run and the angle of inclination is 45°. Values measured as a percent do not have a linear relationship and a difference in values may not accurately depict the difference in on the ground slope, likewise, a difference in 1% does not equal a 1° change in gradient (Bolstad 2002).

A	B	C
D	E	F
G	H	I

$$\Delta z_x / \Delta x = \frac{(Z_a + 2 \cdot Z_d + Z_g) - (Z_c + 2 \cdot Z_f + Z_i)}{8 \cdot \text{cell resolution}}$$

$$\Delta z_y / \Delta y = \frac{(Z_a + 2 \cdot Z_b + Z_c) - (Z_g + 2 \cdot Z_h + Z_i)}{8 \cdot \text{cell resolution}}$$

$$Slope(E)\% = \sqrt{(\Delta z_x / \Delta x)^2 + (\Delta z_y / \Delta y)^2}$$

- Δz_x is the change in elevation in the x direction
- Δz_y is the change in elevation in the y direction
- Δx is the distance in the x direction (8·cell resolution)
- Δy is the distance in the y direction (8·cell resolution)

Figure 4-6 In this illustration of the Horn Algorithm, slope for the center cell E is computed from the surrounding eight cells labeled A, B, C, D, F, G, H and I.

In one final processing step, the values of the two slope grids were rounded to the nearest whole percent. This step reduces the precision of values; however, whole values are more comparable to those obtained in the field (also measured to the nearest whole percent). The descriptive statistics of each slope dataset at the 46 sample points, including the measures collected in the field and those calculated from the DEMs, are shown in Table 4-4. Table 4-5 presents the error statistics of each of the digital slope datasets compared to the actual slope measurements obtained in the field.

Table 4-4 Descriptive statistics of the slope datasets[1]

	Slope	Lidar Slope	NED Slope	Lidar Drop	NED Drop
Minimum	0	1%	2%	1%	17%
Maximum	79%	87%	36%	45%	67%
Mean	30.5%	29.3%	13.2%	13.9%	30.7%
Standard Dev.	21.6%	19.3%	7.8%	9.1%	13.9%
Variance	464.6	372.5	61.5	82.8	193.3

Table 4-5 Error statistics of the slope datasets

	Lidar Slope	NED Slope	Lidar Drop	NED Drop
RMSE	14.9%	27.8%	16.2%	27.2%
MAE	9.9%	21.6%	12.4%	21.8%
MAX	44%	-67%	42%	-69%
MIN	1%	0	0%	0

Land Cover Data

As previously mentioned, many studies (Lefsky et al. 2002; Rosso et al. 2003; Chasmer et al. 2004) have examined the presence of vegetation in the return signal and resultant dataset of an ALS. In some instances the goal may be to measure biomass (Lefsky et al. 2002) or map vegetation (Rosso et al. 2003), and therefore bare-earth returns are not desirable. In other applications, however, particularly those that require accurate elevation measures, such as surface hydrology, landform, or slope

[1] The second column, labeled "slope", is the dependent variable in the subsequent models, while the remaining four columns represent the independent slope variables.

mapping, bare-earth returns are necessary and feature removal is key (Bowen and Waltermire 2002; Woolard and Colby 2002; French 2003; Barber and Shortridge 2004). In either case it is clear that the presence or absence of vegetation may interfere with the collection of lidar data, influencing the accuracy of both elevation and derivative slope values, and to some degree affect USGS level 2 DEMs as well. For this reason I included land-cover, more specifically a categorical dummy variable for presence (1) or absence (0) of forest cover, as a possible variable in the multiple regression model Although herbaceous or understory vegetation can also intercept the laser signal, lower vegetation tends to have a much smaller standard deviation of height than trees and, therefore, these types of ground covers are less likely to influence slope measurements (Pfeifer et al. 2004). Likewise, it does not interfere with stereo-correlation or affect the posting density of Level 2 elevation measures.

The land cover data used in this study was adapted from the USGS National Land Cover Data (NLCD 1992) set. It uses a 21-class classification scheme and was derived from the early to mid-1990s Landsat Thematic Mapper satellite data. It is freely available from the USGS as a 30-meter raster data set for the conterminous United States (NLCD FGDC metadata). The NLCD for the state of Michigan is also available from the MICGI. It is essentially the same data except that it has been repackaged for each county as a shape file projected in the MI GeoRef projection, NAD83.

I reclassified the original land cover classes of the study area, which included Barren (31), Forested Upland (41, 42), Shrub land (51), Herbaceous Upland Natural/Semi-natural Vegetation (71), Herbaceous Planted/Cultivated (81, 82), and Wetlands (91, 92), into the forest (1)/non-forest (0) variable for the purposes of this study. Forested Upland (41, 42) and Wetlands (91 Woody Wetlands) were reclassified as forest while all other classes were considered non-forest (Figure 4-8). Table 4-6 is a land cover confusion matrix that summarizes how often the classifications of the NLCD dataset correctly matched those observed in the field. The overall accuracy of the dataset, as noted in the table, is 85%, which means that of those 46 sites sampled, 39 sites were correctly classified as forest/non-forest by the NLCD. Also noted in the table is the number of sites that have a canopy (28), compared to those without (18).

a. b.

Figure 4-8 Photographs showing two land cover types found on Lake Michigan sand dunes. Photo a. is an example of the Herbaceous Upland Natural/Semi-natural Vegetation (71) land cover class, which would be reclassified as non-forest in contrast to photo b. an example of Forested Upland (41, 42).

Table 4-6 Confusion matrix comparing ground truth land cover (observed) classifications to the NLCD (expected) classifications.

Observed Class

Expected Class	Forest	Non-forest	Row Total
Forest	25	4	29
Non-forest	3	14	17
Column Total	28	18	46
Producer Accuracy	89%	78%	
User Accuracy	86%	82%	
Overall Accuracy	85%		

Overview of Multiple Regression

Multiple regression is a statistical method designed in most cases to either predict or explain a variable of interest in order to reach one of two research goals, either: 1) to make valid projections concerning an outcome (prediction), or 2) to make an attempt to understand a phenomenon by examining a variable's correlates on a group level (explanation; Osborne 2000). The aim of this study is to predict an *actual* or *true* slope measurement based on variables that can be obtained from or calculated in a GIS. In order to do this, a multiple regression model was constructed in SYSTAT statistical software package to predict the dependant variable (Y), in this case actual slope, based on the value of at least one independent variable ($X_1,\ldots,X_n$).

The multiple regression equation takes the form of:

$$Y_i = a + \sum_{i=1}^{n} b_i \cdot X_i + \varepsilon$$

Where,

- Y_i is the value of the dependent variable;
- a is the constant or intercept;
- b_i is the standardized regression coefficient for X_i;
- X_i is the independent variable (accounting for the variance in Y); *and*
- ε is the error, resulting from the effect of unspecified independent variables and/or a totally random element.

Assumptions of Multiple Regression

In order to achieve reliable results with a regression model, Poole and O'Farrell (1971) state that six "critical assumptions" must be met. Of these six assumptions, they note, any one can be either more or less important depending on the research purpose. While it is not essential to meet the assumptions of measurement error and multicollinearity, when the research goal is prediction, the authors say that it is most important that: 1) The relationships between Y_i and each of the independent variables X_i are linear; 2) the variance of the conditional distribution of ε is constant for all such distributions (homoscedastic); and 3) the values of ε are not autocorrelated (Poole and O'Farrell 1971).

If the relationship between the dependent and independent variable(s) is non-linear, it is likely that the regression equation will not correctly estimate the relationship between variables. A graph of each independent variable plotted against the dependent variable can indicate whether or not the relationship is linear. It is also important to plot the standardized residuals to identify extreme heteroscedasticity, which can distort findings and seriously weaken the analysis (Osborne and Waters 2002).

V. RESULTS AND DISCUSSION

This chapter will begin with a discussion of the independent variables as they relate to the assumptions of regression. Next, I will present the summary statistics for each model, which will determine if my hypothesis is correct, followed by an analysis of the residuals. I will then conclude the chapter with a discussion of the results, including the error in the best model.

As first noted in the previous chapter, the variables used in a multiple regression model need to meet certain key assumptions for the results to be reliable. Again, when the research goal is prediction the most important assumptions are: 1) the relationships between Y_i and each of the independent variables X_i are linear; 2) the variance of the conditional distribution of ε is constant for all such distributions (homoscedastic); and 3) the values of ε are not autocorrelated (Poole and O'Farrell 1971). Of these three, only the first assumption can be tested prior to generating the model. To do this I plotted each of the independent variables including calculated slope (LDSLOPE, NDSLOPE, LDDROP, and NDDROP) and elevation (LDELEV and NDELEV), against the dependent variable true slope (FDSLOPE). Figures 5-1 and 5-2 show a regression line superimposed on scatterplots of LDSLOPE and FDSLOPE and NDSLOPE and FDSLOPE. The remainder of the graphs can be seen in Appendix A.

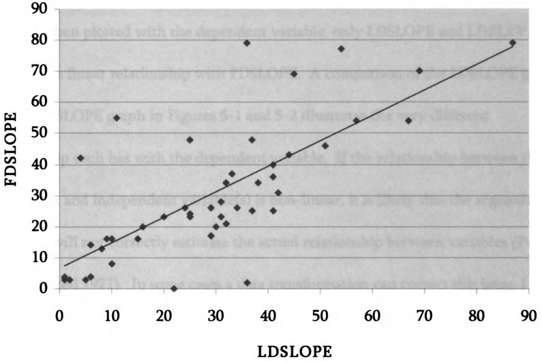


Figure 5-1 Graph showing LDSLOPE plotted against FDSLOPE. With the exception of a few outlying points, LDSLOPE exhibits a linear relationship with FDSLOPE.

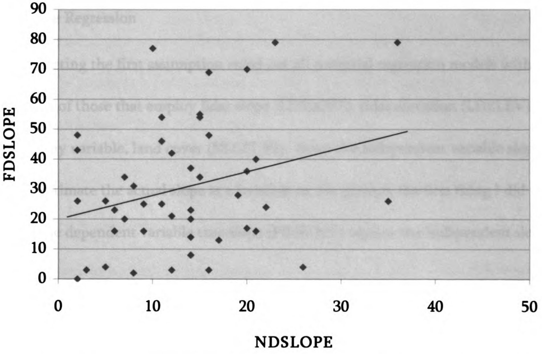


Figure 5-2 Graph of NDSLOPE plotted against FDSLOPE. The relationship depicted in this graph is not linear and therefore the NDSLOPE data does not meet the first assumption.

When plotted with the dependent variable, only LDSLOPE and LDELEV displayed a linear relationship with FDSLOPE. A comparison of the LDSLOPE graph to the NDSLOPE graph in Figures 5-1 and 5-2 illustrates the very different relationship each has with the dependent variable. If the relationship between the dependent and independent variable(s) is non-linear, it is likely that the regression equation will not correctly estimate the actual relationship between variables (Poole and O'Farrell 1971). In some cases a data transformation can correct this issue. I performed a log transformation on the FDSLOPE, NDSLOPE, and NDELEVATION data sets, yet their respective relationships remained non-linear. Similar data issues arose with the LDDROP and NDDROP datasets[2].

Univariate Regression

Testing the first assumption ruled out all potential regression models with the exception of those that employ lidar slope (LDSLOPE), lidar elevation (LDELEV), and the dummy variable, land cover (NLCD_92). Since the independent variable slope is used to estimate the actual slope at a location on the ground, the first thing I did was regress the dependent variable true slope (FDSLOPE) against the independent slope

---

[2] Both slope datasets were calculated using the maximum gradient. This method determines the maximum change in elevation between the processing cell and its eight surrounding neighbors and divides that difference by the distance between the centers of the two cells. Using a raster grid, the distance is always the same, so unless there are extreme elevation differences across cells and throughout the dataset, the result is a limited number of slope measurements.

variable (LDSLOPE). The summary statistics of the univariate analysis are shown in Table 5-1 .

Table 5-1 Summary statistics of the univariate regression model.

Variable	Constant (a)	Coefficient (B)	t-value	P (1-tail)	Adj. R^2
LDSLOPE	6.476	0.822	1.621	0.000	0.530

The R^2 value is an indicator of how well the model fits the data (e.g., R^2 close to 1.0 indicates that the model accounts for nearly all of the variance in the independent variables). The adjusted R^2 is a good statistic to summarize the fit of the regression model because it corrects for the number of coefficients and sample size, as well as a small degree of error, to more closely fit the model estimated from a sample to the entire population (Wilkinson et al. 1996). The univariate model in this study produced an adjusted R^2 of .530, or 53%.

Multivariate Regression

My initial hypothesis was that lidar elevation data used in conjunction with calculated slope and land cover would more accurately predict true slope than the same model that used NED as its source. Since the NED data did not meet the assumption of a linear relationship it may not yield a reliable model; however, I decided to proceed with the NED data for the sole purpose of comparing the two models and evaluating my hypothesis. In the second phase of my analysis, I performed a standard multivariate analysis regressing true slope against, elevation, calculated slope, and land cover for each data source. The regression equation for both the lidar and NED models can be stated as:

$$Y \text{ FDSLOPE} = a + b_1\text{SLOPE} + b_2\text{ELEV} + b_3\text{NLCD\_92}$$

The summary statistics of these models are presented in Table 5-2 and 5-3. The adjusted R^2 for the lidar model is equal to .595, meaning that slope, elevation, and land cover account for 59.5% of the variation in true slope, compared to only 14.2% for the NED model. Tolerance values for all variables are above .60, which indicates that there is little co-linearity among the variables.

In both the lidar and NED models, land cover was not a significant variable in the equation at the 95% confidence level; therefore I removed this variable and generated the models again using only slope and elevation. Again, the NED model yielded poor results, with an adjusted R^2 value of .127. Further, NDSLOPE was not found to be a significant variable. Of the three models presented, the best model for predicting true slope was one that used LDSLOPE and LDELEV, with an adjusted R^2 value of .603 (Table 5-4).

Table 5-2 Summary statistics of the multivariate regression model employing lidar source data, and slope, elevation, and land cover as independent variables.

Dep Var: FDSLOPE			N: 46			
Multiple R: 0.789			**Squared multiple R: .622**			
Adjusted squared multiple R: 0.595			**Standard error of estimate: 13.713**			
Variable	**Coefficient (B)**	**Std Error**	**Std. Coefficient**	**Tolerance**	**T**	**P (1-Tail)**
CONSTANT	-146.229	51.206	0	.	-2.856	**0.004**
LDSLOPE	0.688	0.116	0.616	0.829	5.916	**0.000**
LDELEV	0.824	0.277	0.325	0.754	2.972	**0.003**
NLCD_92	-1.967	4.567	-0.045	0.841	-0.431	**0.335**
Analysis of Variance						
Source	**Sum-of-Squares**	**Df**	**Mean-Square**	**F-ratio**	**P**	
Regression	13007.548	3	4335.849	23.058	0.000	
Residual	7897.865	42	188.044			

Table 5-3 Summary statistics of the multivariate regression model employing NED source data, and slope, elevation, and land cover as independent variables.

Dep Var: FDSLOPE			N: 46			
Multiple R: 0.446			**Squared multiple R: 0.199**			
Adjusted squared multiple R: 0.142			**Standard error of estimate: 19.969**			
Variable	**Coefficient**	**Std Error**	**Std. Coefficient**	**Tolerance**	**T**	**P (1-Tail)**
CONSTANT	-137.914	83.139	0.000	0.000	-1.659	**0.053**
NDSLOPE	0.696	0.382	0.253	0.986	1.822	**0.038**
NDELEV	0.746	0.407	0.259	0.958	1.833	**0.037**
NLCD_92	8.096	6.191	0.183	0.971	1.308	**0.099**
Analysis of Variance						
Source	**Sum-of-Squares**	**Df**	**Mean-Square**	**F-ratio**	**P**	
Regression	4157.208	3	1385.736	3.475	0.024	
Residual	16748.205	42	398.767			

Table 5-4 Summary statistics of the multivariate regression model employing lidar source data, and slope and elevation as independent variables.

Dep Var: FDSLOPE			N: 46			
Multiple R: 0.788			**Squared multiple R: 0.621**			
Adjusted squared multiple R: 0.603			**Standard error of estimate: 13.582**			
Variable	**Coefficient**	**Std Error**	**Std. Coefficient**	**Tolerance**	**T**	**P (1-Tail)**
CONSTANT	-139.989	48.645	0.000	.	-2.878	**0.006**
LDSLOPE	0.682	0.114	0.611	0.840	5.963	**0.000**
LDELEV	0.786	0.260	0.309	0.840	3.019	**0.002**
Analysis of Variance						
Source	**Sum-of-Squares**	**Df**	**Mean-Square**	**F-ratio**	**P**	
Regression	12972.676	2	6486.338	35.16	0.000	
Residual	7932.737	43	184.482			

Table 5-5 Summary statistics of the multivariate regression model employing NED source data, and slope and elevation as independent variables.

Dep Var: FDSLOPE			N: 46			
Multiple R: 0.788			**Squared multiple R: 0.621**			
Adjusted squared multiple R: 0.603			**Standard error of estimate: 13.582**			
Variable	**Coefficient**	**Std Error**	**Std. Coefficient**	**Tolerance**	**T**	**P (1-Tail)**
CONSTANT	-151.446	83.170	0.000	.	-1.821	**0.076**
LDSLOPE	.701	0.385	0.255	0.986	1.820	**0.076**
LDELEV	.836	0.405	0.290	0.986	2.067	**0.045**
Analysis of Variance						
Source	**Sum-of-Squares**	**Df**	**Mean-Square**	**F-ratio**	**P**	
Regression	3475.422	2	1737.711	4.287	0.020	
Residual	17429.991	43	405.349			

Analysis of the Residuals

The section above provided an assessment of each model based on key statistics such as the adjusted R^2. These numbers are useful for model validation, but it is also important to employ graphical methods. Graphical methods have one advantage over numerical methods in that they are able to show the relationship between the model and the data. A graph of the residuals can also uncover trends in the data that are not apparent in the statistical output (NIST/SEMATECH e-Handbook of Statistical Methods 2006). Further, the last two critical assumptions, the variance of the conditional distribution of ε is homoscedastic and the values of ε are not autocorrelated, rely solely on the residual values.

If the model's fit to the data is correct, the residuals should approximate the random errors that make the relationship between the explanatory variables and the response variable a statistical relationship. Therefore, if the residuals appear to behave randomly, it suggests that the model fits the data well. A scattergram of the regression residuals plotted against the regression estimates is shown in Figure 5-3. While there are several outlier cases, the residuals appear to be homoscedastic, which means the variance of the error values does not appear to be related to the size of the estimate. Graphs of the residuals plotted against each independent variable can also reveal patterns of heteroscedasticity. An examination of these two plots also revealed no obvious patterns (Appendix C).

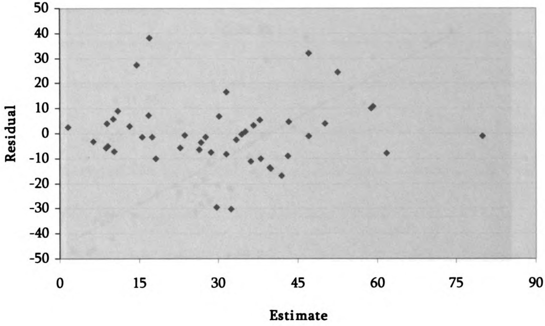


Figure 5-3 Graph showing the regression estimates plotted against the residuals.

The regression output can indicate if there are any outliers including large residuals, high leverage points, and influential points, all of which can distort the regression model. The studentized residuals, leverage, and Cook's distance are three statistics that can identify if any of these problems are present in the data. In general, 95% of the studentized residuals should fall between ± 2, and 99% between ±3 (Wilkinson et al. 1996). There were five points that fell outside of this range. An examination of the Cook's D statistic showed one case that was influential at the 95% level (compared to an *F*-distribution), while there were no cases that had unusually high leverage (Figure 5-4).

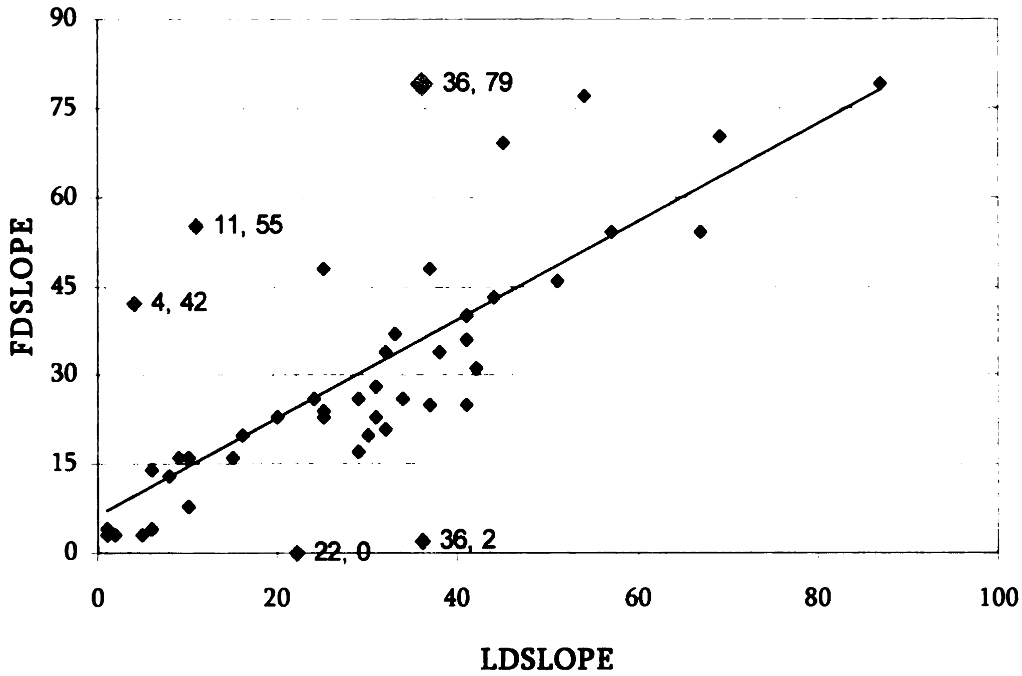


Figure 5-4 Graph showing the independent variable LDSLOPE plotted against the dependent variable FDSLOPE with outlier cases highlighted. The red points note cases that have a large studentized residual, with the larger red point indicating the case that also has a large influence.

Discussion

The first conclusion that can be drawn from the results of this research is that the NED level 2 DEM for the study area is not fit for use in applications that require accurate slope measurements. While the sampling method used to collect the field slope measurements was more comparable to the resolution of the lidar data, it is still reasonable to expect that the relationship between the two datasets, which represent the same variable, to be linear with a fairly strong positive relationship. A graph of FDSLOPE and NDSLOPE showed that this was not the case[3].

[3] The Pearson correlation coefficient further confirmed the lack of a positive linear relationship between the two variables.

There is a large difference in the resolution of the lidar data (2 m) versus the NED data (30 m) and in a study comparing lidar data to NED, Barber and Shortridge (2005) found that slope in particular is more affected by resolution and less affected by source. To test this theory on my data, I resampled the NED dataset to a finer resolution (5 m), calculated the slope variables, and plotted the resampled NED slope variable against true slope to see if there were any obvious improvements. An examination of the resampled NED data and the new graph (Appendix B) showed no improvement whatsoever, nor did the models that employed this data. This leads me to conclude that the resolution of this dataset is inherently tied to its source and, therefore, the source of the data, in this case, is as important as the resolution.

Although some researchers (Holmes et al. 2000; Gesch et al. 2002) have found USGS Level 2 DEMs to be generally accurate, the elevation measures used in this research on average were in error by more than 10 m. Compounding this problem, the error that exists in the elevation surface typically propagates in derivative measures such as slope (Holmes et al. 2000). While I cannot determine why the NED data was of such poor quality or conclude anything about the overall accuracy of the entire DEM, terrain attributes within the study area (i.e. slope and land cover), and the resolution (posting density) of the source data may account for the lack of correlation between FDSLOPE and NDSLOPE.

Research has found that USGS Level 2 production methods can be hampered when the ground is under a forest canopy. A dense forest makes stereo-correlation more difficult due to the lack of clearly defined features such as roads and buildings and when using photogrammetric techniques, fewer postings are collected in forested areas (Bolstad and Stowe 1994; Lefsky et al. 2002). A positive correlation has been shown between true elevation and elevation errors and true slope and slope errors[4]. Areas with the steepest terrain are often times forested, which further compounds this issue (Bolstad and Stowe 1994).

One final problem is post-processing. If interpolation has been used to calculate elevation values, it tends to smooth microtopography and this could be the reason why such large slope errors are associated with more variable terrain. Most of the sample points were collected on the sides of dunes and point estimation in these areas is very difficult. Again, this leads to error in the interpolated elevation surface, which then propagates itself in the derivative slope (Bolstad and Stowe 1994).

The results of the regression analysis showed that the best model for predicting true slope employs slope (calculated using the Horn algorithm[5]) and elevation. The land cover dummy variable was found to be insignificant and did not improve the model. I found this surprising because larger errors are expected in forested terrain due to a lower posting density. Many of the returns that would be present on barren land would be removed from forested areas because they hit the canopy and if not, they would remain in the dataset (Bolstad and Stowe 1994).

---

[4] The Pearson correlation coefficient for FDSLOPE and slope error indicated a very strong positive correlation between the two variables. I found a strong negative correlation between elevation error and true elevation, indicating that as elevation increased, error decreased.

[5] For a review of the Horn Algorithm, see chapter IV.

In this particular data set, it is probable that post-processing was successful in removing man-made obstructions from the data set, similar to findings made by Barber and Shortridge (2005). When you consider the collection method, a lower posting density for lidar is still much greater than that of NED. So while similar issues with terrain attributes and error persist in lidar data, the collection method (source), posting density (resolution), and post-processing of the raw data points could explain why land cover was not a significant variable as well as the difference between lidar and NED.

The second point to address is why LDELEV was a significant variable and improved the equation. Research has found that both elevation error and modeled slope error increase when slopes on the ground are steeper, suggesting that there is a connection between true slope and elevation (Chang and Tsai 1991; Bolstad and Stowe 1994; Hodgson et al. 2003). Further establishing that relationship, both FDSLOPE and absolute slope error (equal to |FDSLOPE – LDSLOPE|) were strongly correlated with LDELEV at the 99% confidence level. This means that steeper slopes are associated with higher elevations in the study area and the difference between FDSLOPE and LDSLOPE is at least in part explained by LDELEV, as the regression equation would suggest. Although I found that land cover did not have an impact on the model, based on past research it still may be reasonable to expect that certain land cover attributes are associated with increase in error in the digital data sets. Forest is the dominant land cover on many of the highest and most well established dunes, which could explain the positive linear relationship between elevation and error in the slope dataset.

One final point to address is the problems that could arise by using elevation and one of its derivatives, slope, in the same model. Slope, as discussed in the methods section, is calculated as the average change in elevation from north to south and east to west, across a nine cell neighborhood. Although the tolerance statistics of the model that used LDSLOPE and LDELEV indicated that there was not a statistical relationship between the two variables, they certainly are interdependent. When two variables are collinear, or providing the same information to the regression equation, the standard errors may be inflated and, therefore, the estimated coefficients may not be representative of the population coefficients (Wilkinson et al. 1996).

Conceptually speaking, slope, as it is calculated, is dependent on the difference in elevation values, not the actual elevation values. For example, the gradient from 10 m to 12 m over a distance of 1 m and the gradient from 100 m to 102 m over a distance of 1 m are the same; however, the elevation at each of those points is very different. Aside from the model's tolerance values, there is no way to know if the relationship between elevation and slope has affected the model apart from conducting further research from the same population.

Discussion of Model Error

Like graphical methods, analysis of the error terms including describing, classifying, and mapping the residuals, can uncover trends in the data. The error in remotely sensed data sets and models derived from them is as important, if not more important, than the data itself, which is made evident by previous research (Chang and Tsai 1991; Bolstad and Stowe 1994; Fisher 1998; Liu and Jezek 1999; Holmes et al. 2000; Lefsky et al. 2002; Woolard and Colby 2002; Hodgson et al. 2003; Rosso et al. 2003; Sallenger et al. 2003; White and Wang 2003; Nagihara et al. 2004). Key statistics of the model's error values are found in Table 5-6 shown below. While some error values were very high, the RMSE for the model was 13% with over half of the residuals below 7%. As noted in the previous chapter, values measured as a percent do not have a linear relationship and a difference in values may not accurately depict the difference in on the ground slope, likewise, a difference in 1% does not equal a 1° change in gradient (Bolstad 2002). Consequently it is difficult to quantify the error terms[6].

Table 5-6 Model error statistics

Minimum	Maximum	Mean	MAE	RMSE	Standard Dev.
0.007	38.009	0.00	9.38	13.132	9.29

[6] I will address this point again in the "Research Problems" section of Chapter 6.

Three of the five largest studentized residuals are associated with slopes greater than 40% (Figure 5-5). This finding is similar to previous research which found that modeled slope error increases when slopes on the ground are steeper (Chang and Tsai 1991; Bolstad and Stowe 1994; Hodgson et al. 2003). The two remaining outliers are found where on the ground slope is less than 5%. Collectively these findings suggest that the model is more accurate when true slopes are moderate and less accurate when slopes are either gentle or steep. A graph of the average root mean square error broken down by slope class illustrates this point (Figure 5-6).

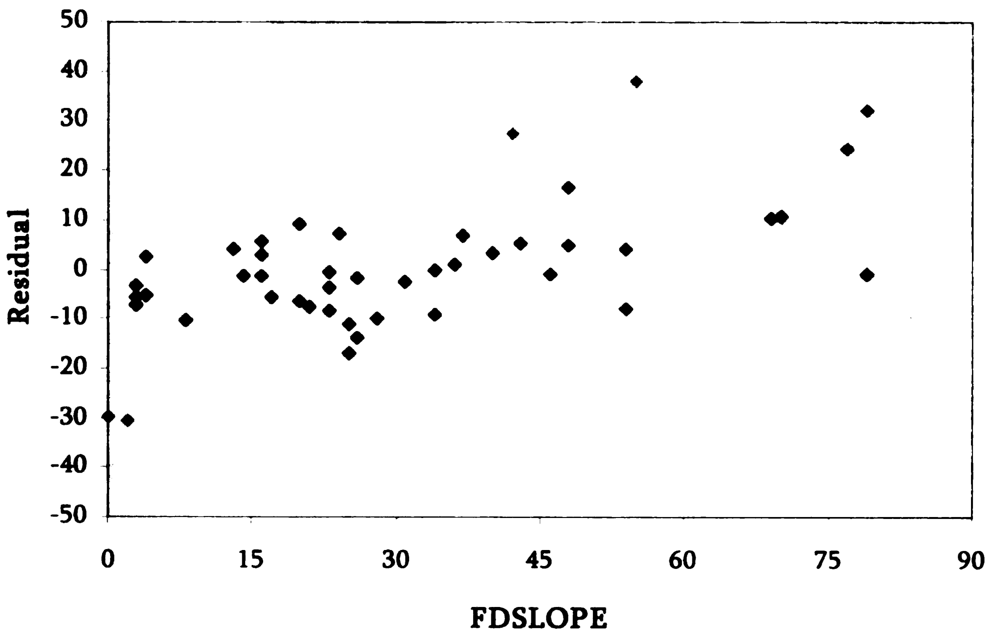


Figure 5-5 Graph showing FDSLOPE plotted against the residuals; points with the largest studentized residuals are shown in red.

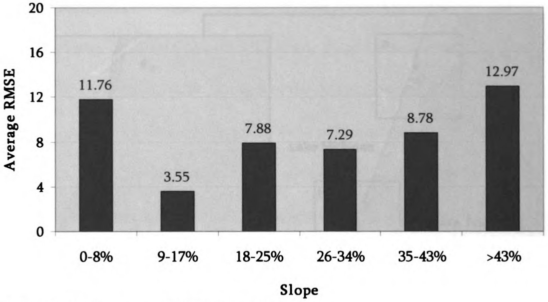


Figure 5-6 Graph of the average RMSE by slope class.

A map of the residuals shown in Figure 5-7 revealed that several larger residuals are clustered in the northern portion of the study area (inset a.). Here, more than half of the points have error values that are more than one standard deviation above the mean. However, an investigation revealed that there is no spatial autocorrelation between the residuals here or elsewhere in the study area.

The two northernmost points were collected off of a utility access road beneath a forest canopy. One of these points had a residual value of 27.44%, with the model significantly underestimating the slope of this point. The cluster of points to the immediate south was collected in a residential area and within highly variable, forested terrain. Given the number of houses present there, this area is perhaps the most representative of sites inspected by the MDEQ. The mean absolute error (MAE) in this area was 10.88%, which is slightly higher than the overall average of 9.38%.

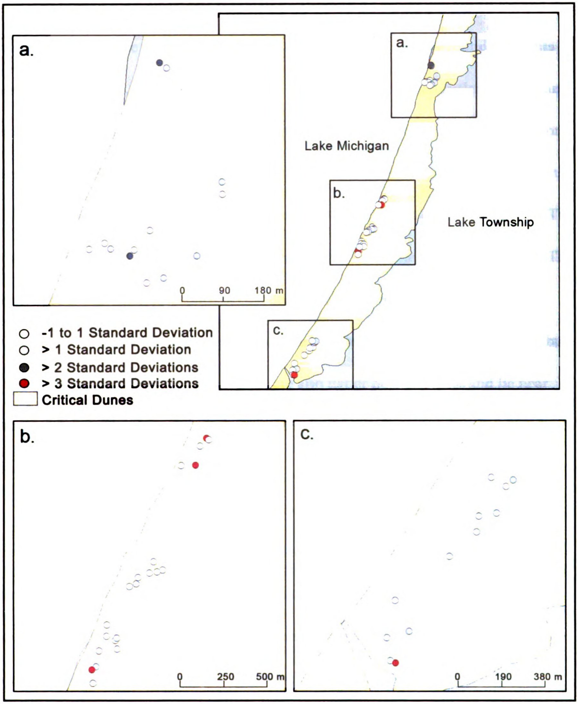


Figure 5-7 Map showing the spatial distribution of the residuals. Points were classified by the number of standard deviations the residual at that point fell above or below the mean. The map was broken into three geographic extents (a. north, b. central, and c. south) based on the proximity of the points to one another.

Three of the four largest residuals (> 3 Standard Deviations) are all located in the same geographic extent (Figure 5-7, inset b.). Based on the surrounding points and my field notes, this is likely the result of site specific characteristics rather than spatial autocorrelation. Two of the three points are located within 200 meters of one another near Weko Beach. Of those, one was collected on a large dune bordering the lake. The front of the dune is primarily dune grass, while the slipface is forested. The dune is also topped by an elevated boardwalk, which runs the length of its crest. The other point was taken just south of that dune in a deflation hollow, surrounded on three sides by forest. Both points are classified by the NLCD as having a canopy, but neither is located under a canopy. Both are also flat or near flat areas and lie near steeper slopes. The three points closest to them all have slopes greater than 23%, yet have small residuals.

The southernmost extent had the lowest overall error and only one large residual value out of the twelve points collected there (Figure 5-7, inset c.). It is also important to note that of those twelve points, nine were classified as bare sand and dune grass. The one point that exceeded the mean error by more than three standard deviations (32.02%) was collected on the side of an exceptionally steep dune (FDSLOPE = 79%), beneath a forest canopy. In this case the lidar dataset and the model both underestimated the slope. That being said, the estimate of a nearby point with nearly identical slope and terrain characteristics was in error by less that 1%.

This indicates that the problem with the lidar dataset may be inconsistency as opposed to terrain characteristics.

Summary

The results of my study revealed several findings that are both interesting and relevant. The lack of a linear relationship between NDSLOPE and FDSLOPE indicated that slope derived from the NED DEM for the study area does not accurately depict true slope. To a degree this could be a reflection of the method used in this study to collect slope measurements. However, the production method of USGS level 2 DEMs, resolution (posting density), post-processing, and terrain characteristics within the study area are most likely the cause of this discrepancy.

The univariate model, which employed lidar slope, and the multivariate model, which employed lidar slope and lidar elevation, produced adjusted R^2 values of .530 and .603 respectively. While the elevation variable improved the model, the land cover dummy variable was not found to be a significant part of the regression equation and therefore discarded. The strong positive correlation between elevation and slope error is likely the reason that elevation improved the model. However, given the collection method and past research (Bolstad and Stowe 1994; Bowen and Waltermire 2002; Hodgson et al. 2003), it is unclear why land cover did not account for a greater part of the variance. I believe this is due to a superior resolution, even when taking into account fewer postings, and successful post-processing methods.

Based on the studentized residuals, six points were classified as outliers. While some of the terrain characteristics at these locations were described, the reason for these large errors cannot be determined. It is plausible that there may be greater inconsistencies in the lidar data based on certain terrain characteristics. It is also interesting to note that the largest errors occurred on slopes less than 5% or greater than 40% and on average, error was larger for these two slope classes than for any other. Slope error is expected to increase when slopes are steeper and terrain is more variable, but finding two of the largest errors at places on the landscape with little or no slope is unusual.

VI. CONCLUSION

Collectively, the State of Michigan contains what may be the largest complex of freshwater dunes in the world (VanOort et al. 2001; Arbogast et al. 2002). The conditions that existed at the time these sand dunes formed were unique and led to the development of some of the most magnificent parabolic dunes. Since the early 1900's, dune sand has been prized for its use as foundry cores in the automobile industry. From that time on, the sand dunes found along the eastern shore of Lake Michigan have been threatened by sand mining and, in some cases, completely destroyed by intensive mining practices (Buckler 1979; Lake Michigan Federation 1999).

Initially it was the public's concern over mining that led to the passing of the *Sand Dunes Protection and Management Act* (Act No. 222, Public Acts of 1976), but as human and developmental pressures increased, the act was amended to include the *critical dune* designation (Lake Michigan Federation 1999; Michigan State Legislature Acts No. 146 and 147, Public Acts of 1989). These amendments also established standards and a permitting program for development in areas designated *critical dunes* (Michigan State Legislature Acts No. 146 and 147, Public Acts of 1989). At present, the MDEQ is charged with the management and administration of critical dune policy and employs eleven agents to implement the permitting program in nearly forty counties (Warner, personal comm., 2002). Today, the sand dunes that those agents work to protect are highly valued for the scenic, recreational, ecological, and economic opportunities that they provide (Buckler 1979; Lichter 1995; Arbogast and Loope 1999; VanOort et al. 2001; Arbogast et al. 2002).

Intrigued by the management issues surrounding Lake Michigan's sand dunes and inspired by their evolution, I designed a research project that could potentially be beneficial to agencies like the MDEQ as well as other environmental applications. In Michigan, the key to achieving a more objective assessment of the coastal zone may lie in integrating the use of digital and remotely sensed data with the current sand dune protection and management program. While the results of this study were focused on one regression model, the research I conducted fell directly at the juncture of several different disciplines including geomorphology, terrain modeling, and environmental policy and management.

I began by selecting a study area that was both suitable for what I intended to study and for which two different sources of digital data were available, including low resolution NED and high resolution lidar data. I found both of these attributes in Lake Township, which is located along the southern extent of the lakeshore in Berrien County, Michigan and contains 2,200 acres of the state's critical dunes.

Next, I outlined several research questions that I intended to explore in the hope of better understanding how digital data could be used to aid management decisions that typically require data collection in the field. The first question I considered was: What terrain variables, or combination of variables, when used in a regression model, most closely predict true slope? Of the three models I presented in the Results and Discussion chapter, the best model for predicting true slope was one that used lidar as the source data and calculated slope and elevation as the independent variables. This model had an adjusted R^2 value of .603.

The next question I planned to address was: What is the error in the predicted slope values? While some error values were very high, the RMSE for the model was about 13% with over half of the residuals below 7%. The MAE was equal to 9.38%, with a minimum error value of .007 and a maximum error value of 38.009. The RMSE and MAE as well as a majority of the individual errors seem to be reasonable, but there were several residuals in excess of 25 %. This indicates that the model is not consistent at predicting slope.

Lastly, I hoped to determine to what degree the source, or collection method, and/or resolution of the elevation data influenced the model. I originally planned to use both lidar and NED as the source of the elevation and slope variables and compare the models generated from each. After plotting the slope values calculated using NED against the field slope values, I discovered that the relationship was not linear and therefore violated one of the critical assumptions of multiple regression. Despite this violation, I generated two models using the NED data strictly for comparison. However, in the end, the lack of a linear relationship as well as the results of those models led me to the conclude that the NED level 2 DEM for the study area is not fit for use in applications that require accurate slope measurements. Further, since the resolution of the data is inherently tied to its source and collection method, I also concluded that the source of the data is as important as the resolution.

Research Problems

Over the course of conducting and documenting my research I encountered several problems that I could not overcome. I discovered the first major obstacle to my research the day I entered the field to collect sample points. Prior to entering the field, I generated 100 random sampling points across the entire study area. Although I knew that I would need to eliminate some of those points due to their proximity to one another or site-specific characteristics, I had no idea just how many of them would be inaccessible. Primarily due to private ownership, there were large portions of the study area that I could not even get close to. Moreover, many of these areas were residential and therefore perhaps the most representative of sites visited by the MDEQ.

The second problem is one of measurement. I chose to measure slope as a percent, both in the field and digitally, because I felt that I needed to use the same scale used by the MDEQ when measuring slope in the field. Unfortunately calculating it as a percent eventually led to issues in the analysis. Measured as a percent, slope has an upper limit of infinity and the difference between an 8% and 9% slope and a

70 and 71% slope may not be the same. On the other hand, when measured in degrees, the difference between an 8° and 9° slope and a 70° and 71° slope is an equal difference. Once I recognized the problem I could have converted all of my data from percent slope to degree slope, but at that point I chose not to. For future research I would recommend using degree slope from start to finish.

The next research problem I encountered was bureaucratic. No matter what the outcome of my research was, there is no way to judge whether or not it is good enough to be used by the MDEQ, or even what good enough is. Regardless, given my results, I would not recommend that agents replace measures obtained in the field with those from digital data sources. However, with that said, it is possible that there is an alternate use of lidar for predicting whether or not slope measures are above or below a certain threshold, which may be more useful to an agency like the MDEQ.

The last issue I will discuss is not so much of a problem as it is a reality. While some dunes exist within large, stabilized dune fields, dunes are transient features that can migrate up to 40 meters in a year and a dune will continue to migrate as long as winds are strong enough to carry sand up the windward slope to the crest (Nordstrom et al. 1990; Gabler et al. 2004). In addition to the wind, when working near the shoreline, water is also a factor in destabilizing the coast. What this means is that, on a variety of time scales, what is true of the coastal zone today may not be true tomorrow, and slopes that are present today may not be present in the future.

Directions for Future Research

The goal of this study was to determine which terrain attributes, or combination of attributes, when used in a regression model would most closely predict true slope. Predictability, in this context, also refers to the replication of the regression results (such as the R^2) for other samples from the same population. In an ideal situation, a second sample would exhibit nearly identical results. More often, however, the test sample's results are very different (usually worse, not better) than those for the model's sample, because the regression model is data dependent (Osborne 2000). This may also suggest that the sample used to build the model is not representative of the population in question. Neter et al. (1996) state that, "By far the preferred method to validate a regression model is through the collection of new data." While gathering additional data to test the models' predictive ability was desirable, given the considerable time and effort required to collect more data, it was not possible. Future research would benefit from collecting a much larger dataset that could be used to test this model or generate a new model.

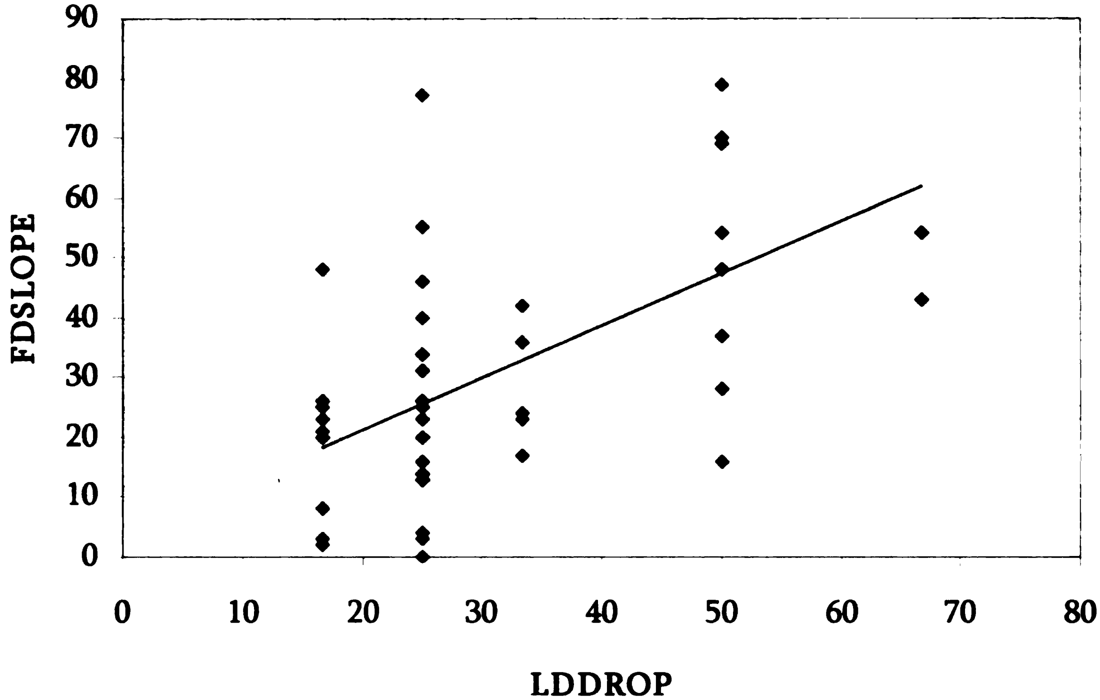
Although lidar appears to be useful in obtaining data points in leaf-on conditions, a closed canopy can result in significantly fewer ground hits and lower accuracy. Ackerman (1996) estimates an overall penetration rate of 24-29% for coniferous forests and 22-25% for deciduous forests during the growing season. While Cowen et al. (2000) report that in canopy closures of 30 – 40%, 80-90% of lidar pulses will reach the ground, but when forest cover increases to 80-90% the number of bare ground hits significantly decreases to 10%. The results of this study did not show conclusive evidence that a canopy or lack there of had any influence on the quality of the lidar data; however, I did not measure the forest cover. Future studies may benefit from classifying the forest by the degree of canopy closure.

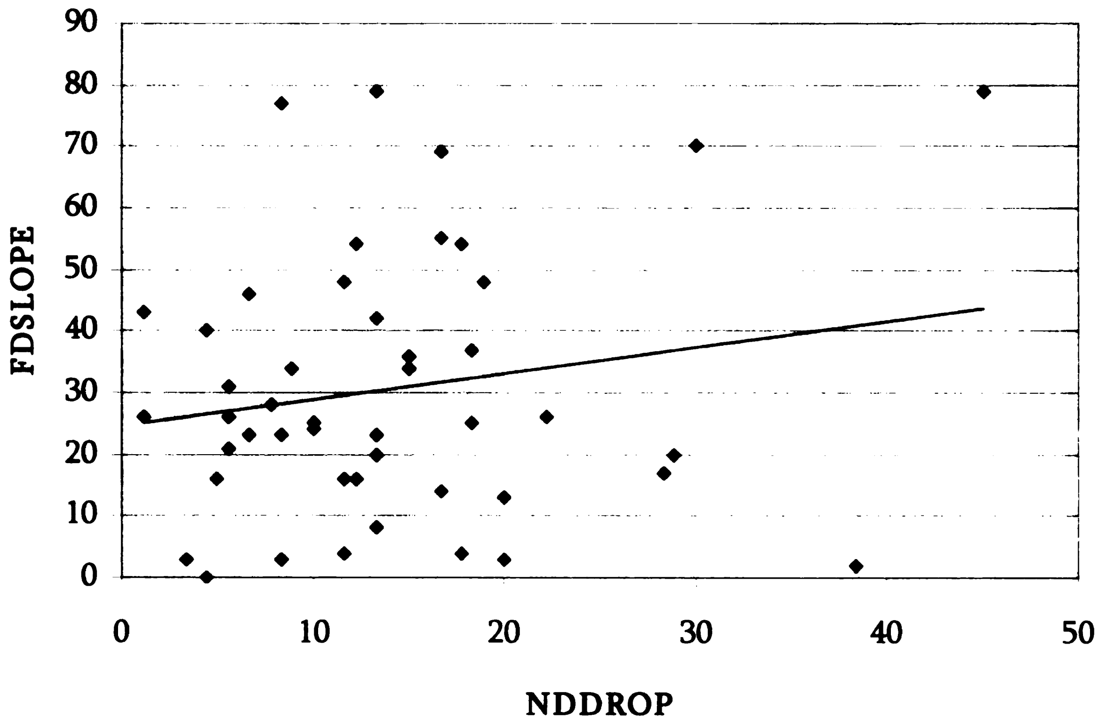
While I ruled out NED as a potential data source for the model, every two months NED is updated and any new source DEMs that have become available are incorporated (Gesch et al. 2001; Gesch et al. 2002). In a study conducted by Barber and Shortridge (2005), the researchers sought to compare lidar data to its NED counterpart. Once they had access to the NED data they discovered that the source of the NED DEM was lidar. There may come a time when all NED DEMs use higher resolution source data, thereby eliminating the need to compare lidar to NED. While this may not be a direction for future research, it is certainly something that future research should take into consideration.

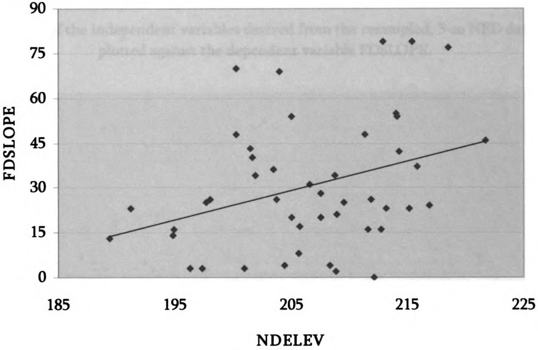
The U.S. coastline stretches more than 20,000 km in length, making remotely sensed data an invaluable resource for monitoring the constant shifts and changes occurring there (Krabill et al. 2000; Sallenger et al. 2003). Further driving the need for research is the heavy anthropogenic pressure placed on this margin, considering that over half of the U.S. population resides near the coast (Nordstrom et al. 1990; Bernd-Cohen and Gordon 1999; White and Wang 2003). Lidar has proven potential as a valuable a tool for evaluating and managing the coastal zone, yet there is still a need to study using lidar derivatives to further characterize topography in this setting.

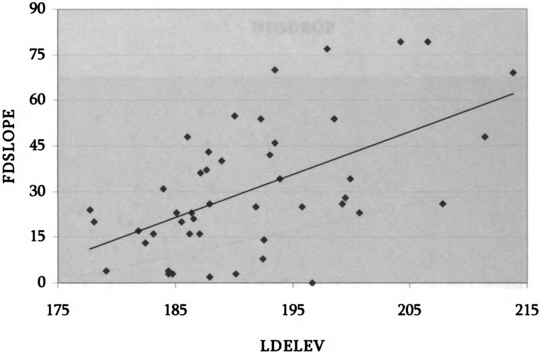
APPENDIX A

Graphs of the independent variables plotted against the dependent variable FDSLOPE.



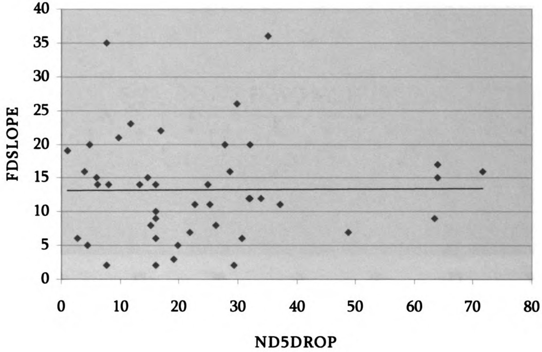


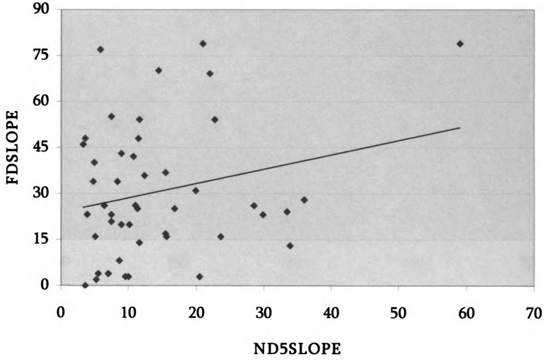

90
75
60
45
30
15
0
FDSLOPE
185
195
205
215
225
NDELEV


90
75
60
45
30
15
0
FDSLOPE
175
185
195
205
215
LDELEV

APPENDIX B

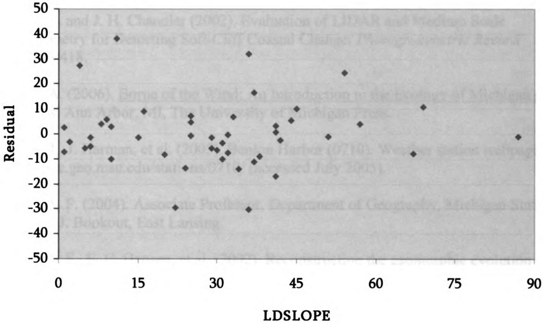
Graphs of the independent variables derived from the resampled, 5-m NED data plotted against the dependent variable FDSLOPE.

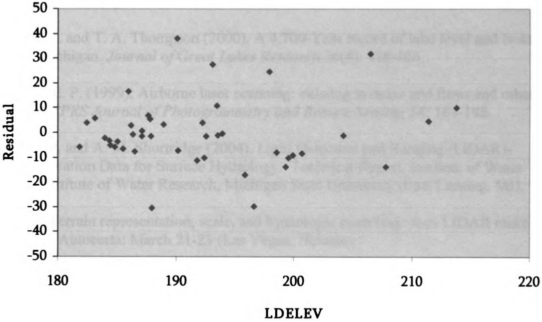




APPENDIX C

Graphs of the residuals plotted against each independent variable.





BIBLIOGRAPHY

Ackerman, F. (1996). Airborne laser scanning for elevation models. *Geomatics Info Magazine* 10(10): 24-25.

Adams, J. C. and J. H. Chandler (2002). Evaluation of LIDAR and Medium Scale Photogrammetry for Detecting Soft-Cliff Coastal Change. *Photogrammetric Record* 17(99): 405-418.

Albert, D. A. (2006). Borne of the Wind: An Introduction to the Ecology of Michigan's Sand Dunes. Ann Arbor, MI, The University of Michigan Press.

Andresen, J., J. Harman, et al. (2005). Benton Harbor (0710). Weather station webpage, http://climate.geo.msu.edu/stations/0710/ (accessed July 2005).

Arbogast, A. F. (2004). Associate Professor, Department of Geography, Michigan State University. J. Bookout, East Lansing.

Arbogast, A. F., E. C. Hansen, et al. (2002). Reconstructing the geomorphic evolution of large coastal dunes along te southeastern shore of Lake Michigan. *Geomorphology* 46: 241 - 255.

Arbogast, A. F. and W. L. Loope (1999). Maximum-Limiting Ages of Lake Michigan Coastal Dunes: Their Correlation With Holocene Lake Level History. *Journal of Great Lakes Research* 25(2): 372-382.

Baedke, S. J. and T. A. Thompson (2000). A 4,700-Year record of lake level and isotasy for Lake Michigan. *Journal of Great Lakes Research* 26(4): 416-426.

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

Barber, C. P. and A. M. Shortridge (2004). Light Detection and Ranging (LiDAR)-Derived Elevation Data for Surface Hydrology. Technical Report, Institute of Water Research Institute of Water Research, Michigan State University (East Lansing, MI).

--- (2005). Terrain representation, scale, and hydrologic modeling: does LiDAR make a difference? Autocarta: March 21-23 (Las Vegas, Nevada).

Bernd-Cohen, T. and M. Gordon (1999). State Coastal Program Effectiveness in Protecting Natural Beaches, Dunes, Bluffs, and Rocky Shores. *Coastal Management* 27: 187-217.

Blott, S. and K. Pye (2004). Application of Lidar Digital Terrain Modelling to Predict Intertidal Habitat Development at a Managed Retreat Site: Abbotts Hall, Essex, UK. *Earth Surface Processes and Landforms* 29: 893-905.

Bolstad, P. (2002). GIS Fundamentals: A First Text on Geographic Information Systems. White Bear Lake, MN, Eider Press.

Bolstad, P. and T. Stowe (1994). An Evaluation of DEM Accuracy: Elevation, Slope, and Aspect. *Photogrammetric Engineering & Remote Sensing* 60(11): 1327-1332.

Bowen, Z. H. and R. G. Waltermire (2002). Evaluation of light detection and ranging (LIDAR) for measuring river corridor topography. *Journal of the American Water Resources Association* 38(1): 33-41.

Brown, A. C. and A. McLachlan (2002). Sandy shore ecosystems and the threats facing them: some predictions for the year 2025. *Environmental Conservation* 29(1): 62-77.

Buckler, W. R. (1979). *Dune Type Inventory and Barrier Dune Classification Study of Michigan's Lake Michigan Shore* Michigan Department of Environmental Quality, Geological Survey Divison: Investigation 23 (Lansing, MI).

Chambers, J. (1991). *R-GUI*. GNU Operating System, Boston, MA, Free Software Foundation, Inc.

Chang, K.-T. and B.-W. Tsai (1991). The effect of DEM resolution on slope and aspect mapping. *Cartography & Geographic Information Systems* 18(1): 69-77.

Chasmer, L., C. Hopkinson, et al. (2004). *Assessing the Three-Dimensional Frequency Distribution of Airborne and Ground-Based Lidar Data for Red Pine and Mixed Deciduous Forest Plots*. International Archives of Photogrammetry, Remote Sensing, and Spatial Information Sciences, Freiburg, Germany.

Chrzastowski, M. J. and T. A. Thompson (1992). Late Wisconsinan and Holocene Coastal Evolution of the Southern Shore of Lake Michigan. Quaternary Coasts of the United States: Marine and Lacustrine Systems. C. H. Fletcher, III and J. F. Wehmiller. Tulsa, OK, Society for Sedimentary Geology. Special Publication No. 48.

Cohen, P. and J. Cohen (1983). Applied multiple regression/correlation analysis for the behavioral sciences. Hillsdale, NJ, Lawrence Erlbaum Associates, Inc.

Cowen, D. J., J. R. Jensen, et al. (2000). A GIS-assisted rail construction econometric model that incorporates LIDAR data. *Photogrammetric Engineering and Remote Sensing* 66(11): 1323-1326.

Cowles, H. C. (1899). The Ecological Relations of the Vegetation on the Sand Dunes of Lake Michigan. Part I -Geographical Relations of the Dune Floras. *Botanical Gazette* 27(2): 95-117.

Davenport, I. J., N. Holden, et al. (2004). Characterizing Errors in Airborne Laser Altimetry Data to Extract Soil Roughness. *IEEE Transactions on Geoscience and Remote Sensing* 42(10): 2130-2141.

Davis, M., M. G. Winkler, et al. (2000). Holocene climate in the western Great Lakes National Parks and Lakeshores: Implications for future climate change. *Conservation Biology* 14(4): 968-983.

Dorr, J. A., Jr. and D. F. Eschman (1970). Geology of Michigan. Ann Arbor, Michigan, University of Michigan Press.

Dow, K. W. (1937). The origin of perched dunes on the Manistee moraine, Michigan. *Michigan Academy of Science, Arts, and Letters* 23: 427-440.

Eichenlaub, V. L., J. R. Harman, et al. (1990). The Climatic Atlas of Michigan. Notre Dame, IN, University of Notre Dame Press.

Eschman, D. F. (1985). Summary of the Quaternary History of Michigan, Ohio and Indiana. *Journal of Geologic Education* 33(3): 161-167.

Fisher, P. (1998). Improved Modeling of Elevation Error with Geostatistics. *GeoInformatica* 2(3): 215-233.

French, J. R. (2003). Airborne LIDAR in Support of Geomorphological and Hydraulic Modelling. *Earth Surface Processes and Landforms* 28: 321-335.

Gabler, R. E., J. F. Petersen, et al. (2004). Essentials of Physical Geography. Belmont, CA, Brooks/Cole- Thomson Learning.

Gares, P. A. (1990). "Eolian processes and dune changes at developed and undeveloped sites, Island Beach, New Jersey". Coastal Dunes: Forms and Processes. K. F. Nordstrom, N. P. Psuty and B. Carter. West Sussex, England, John Wiley & Sons Ltd.: 361-380.

Gesch, D., M. Oimoen, et al. (2002). The National Elevation Dataset. *Photogrammetric Engineering & Remote Sensing* 68(1): 5-11.

Gesch, D., J. Williams, et al. (2001). *A comparison of US Geological Survey seamless elevation models with Shuttle Radar Topograpy Mission data*. Geoscience and Remote Sensing Symposium. IGARSS '01. IEEE 2001 International, Sydney, NSW, Australia.

Guenther, G. C., M. W. Brooks, et al. (2000). New capabilities of the 'SHOALS' airborne lidar bathymeter. *Remote Sensing of Environment* 73(2): 247-255.

Hershman, M. J., J. W. Good, et al. (1999). The Effectiveness of Coastal Zone Management in the United States. *Coastal Management* 27: 113-138.

Hill, J. M., L. A. Graham, et al. (2000). Wide-Area Topographic Mapping and Applications Using Airborne LIght Detection and Ranging (LIDAR) Technology. *Photogrammetric Engineering & Remote Sensing* 66(8): 908 - 914.

Hodgson, M. E., S. R. Schill, et al. (2003). An evaluation of LIDAR- and IFSAR-derived digital elevation models in leaf-on conditions with USGS Level 1 and Level 2 DEMs. *Remote Sensing of Environment* 84(2): 295-308.

Holmes, K. W., O. A. Chadwick, et al. (2000). Error in a USGS 30-meter digital elevation model and its impact on terrain modeling. *Journal of Hydrology* 233: 154-173.

Holt, P. (2002). *The Sand Dunes Program* World Wide Web address: http://www.michigan.gov/deq/0,1607,7-135-3311_4114_4236-9832--,00.html (accessed October 2005).

Irish, J. L. and W. J. Lillycrop (1999). Scanning laser mapping of the coastal zone: the SHOALS system. *ISPRS Journal of Photogrammetry and Remote Sensing* 54: 123-129.

Irish, J. L. and T. E. White (1998). Coastal engineering applications of high-resolution lidar bathymetry. *Coastal Engineering* 35(1-2): 47-71.

Jelaska, S. D., O. Antonic, et al. (2003). Estimating plant species occurence in MTB/64 quadrants as a function of DEM-based variables--a case study for Medvednica Nature Park, Croatia. *Ecological Modeling* 170: 333-343.

Krabill, W. B., C. W. Wright, et al. (2000). Airborne Laser Mapping of Assateague National Seashore Beach. *Photogrammetric Engineering & Remote Sensing* 66(1): 65-71.

Kraus, K. and N. Pfeifer (1998). Determination of terrain models in wooded areas with airborne laser scanning. *ISPRS Journal of Photogrammetry and Remote Sensing* 53: 93-203.

Lake Michigan Federation. (1999). *Vanishing Lake Michigan Sand Dunes: Threats from Mining* Lake Michigan Federation(Chicago).

Lam, N. S.-N. (1983). Spatial Interpolation Methods: A Review. *The American Cartographer* 10(2): 129-149.

Larson, G. and R. Schaetzl (2001). Origin and Evolution of the Great Lakes. *Journal of Great Lakes Research* 27(4): 518-546.

Larson, J. D. (1980). Soil Survey of Berrien County, Michigan. U.S. Department of Agriculture, Soil Conservation Service, Washington, D.C.

Lefsky, M. A., W. B. Cohen, et al. (2002). Lidar Remote Sensing for Ecosystem Studies. *Bioscience* 52(1): 19-30.

Lichter, J. (1995). Lake Michigan Beach-Ridge and Dune Development, Lake Level, and Variability in Regional Water Balance. *Quaternary Research* 44: 181-189.

Liu, H. and K. C. Jezek (1999). Investigating DEM Error Patterns by Directional Variograms and Fourier Analysis. *Geographical Analysis* 31(3): 249-266.

Lloyd, C. D. and P. M. Atkinson (2002). Deriving DSMs from LiDAR data with kriging. *International Journal of Remote Sensing* 23(12): 2519-2524.

Loope, W. L. and A. F. Arbogast (2000). Dominance of an ~150-Year Cycle of Sand-Supply Change in Late Holocene Dune-Building along the Eastern Shore of Lake Michigan. *Quaternary Research* 54: 414-422.

Michigan Department of Natural Resources. (1978). *Land Use / Land Cover Classification.*

Michigan State Legislature. (1976). Sand Dune Protection and Management Act: Act No. 222, Public Acts of 1976.

Michigan State Legislature. (1989). Sand Dune Protection and Management Act: Acts No. 146 and 147, Public Acts of 1989.

Michigan State Legislature. (1994). Natural Resources Environmental Protection Act: Part 353, Act No. 451 Public Acts of 1994, Lansing, MI.

Nagihara, S., K. R. Mulligan, et al. (2004). Use of a Three-Dimensional Laser Scanner to Digitally Capture the Topography of Sand Dunes in High Spatial Resolution. *Earth Surface Processes and Landforms* 29: 391-398.

Neter, J., M. H. Kutner, et al. (1996). Applied Linear Regression Models. Chicago, Irwin.

NIST/SEMATECH e-Handbook of Statistical Methods (2006). World Wide Web address: http://www.itl.nist.gov/div898/handbook/ (accessed September 2006). Section 4.4.4.1.

Nordstrom, K. F. (1994). Beaches and dunes of human-altered coasts. *Progress in Physical Geography* 18(4): 497-516.

Nordstrom, K. F., N. L. Jackson, et al. (2002). Municipal initiatives for managing dunes in coastal residential areas: a case study of Avalon, New Jersey, USA. *Geomorphology* 47: 137-152.

Nordstrom, K. F., R. Lampe, et al. (2000). Reestablishing Naturally Functioning Dunes on Developed Coasts. *Environmental Management* 25(1): 37-51.

Nordstrom, K. F., N. P. Psuty, et al., Eds. (1990). *Coastal Dunes: Forms and Processes.* West Sussex, England, John Wiley & Sons Ltd.

Olson, J. S. (1958a). Lake Michigan Dune Development: 1. Rates of Succession and soil changes on southern Lake Michigan sand dunes. *Botanical Gazette* 119(3): 125-168.

--- (1958c). Lake Michigan Dune Development 2. Plants as agents and tools in geomorphology. *The Journal of Geology* 66(4): 343-351.

--- (1958d). Lake Michigan Dune Development. 3. Lake-level, beach, and dune oscillations. *journal of Geology* 66(5): 473-483.

Orme, A. R. (1990). "The instability of Holocence coastal dunes: the case of the Morro dunes, California". Coastal Dunes. K. F. Nordstrom, N. P. Psuty and R. W. G. Carter. West Sussex, England, John Wiley & Sons Ltd.: 315-336.

Osborne, J. W. (2000). Prediction in multiple regression. *Practical Assessment, Research, & Evaluation* from http://PAREonline.net/getvn.aasp?v=7&n=2 (accessed October 2005).

Osborne, J. W. and E. Waters (2002). Four assumptions of multiple regression that researchers should always test. *Practical Assessment, Research, & Evaluation* from http://PAREonline.net/getvn.asp?v=8&n=2 (accessed October 2005).

Park, C. and A. Dudycha (1974). A cross-validation approach to sample size determination. *Journal of the American Statistical Association* 69: 214-218.

Pfeifer, N., B. Gorte, et al. (2004). *Influences of vegetation on laser altimetry - Analysis and correction approaches.* International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Freiburg, Germany, Institute for Forest Growth, Institute for Remote Sensing and Landscape Information Systems, Albert Ludwigs University.

Poole, M. A. and P. N. O'Farrell (1971). The assumptions of the linear regression model. *Transactions of the Institute of British Geographers*(52): 145-158.

Rango, A., M. Chopping, et al. (2000). Morphological Characteristics of Shrub Coppice Dunes in Desert Grasslands of Southern New Mexico derived from Scanning LIDAR. *Remote Sensing of Environment* 74: 26-44.

Robertson-Rintoul, M. J. (1990). "A quantitative analysis of the near-surface wind flow pattern over coastal parabolic dunes". Coastal Dunes. K. F. Nordstrom, N. P. Psuty and R. W. G. Carter. West Sussex, England, John Wiley & Sons Ltd.: 57-78.

Rosso, P., S. Ustin, et al. (2003). *Use of LIDAR to produce high-resolution marsh vegetation and terrain maps*. ISPRS Workshop -Three Dimensional Mapping from INSAR and LIDAR, Portland, OR.

Sallenger, A. H., W. B. Krabill, et al. (2003). Evaluation of Airborne Topographic Lidar for Quantifying Beach Changes. *Journal of Coastal Research* 19(1): 125-133.

Santer, R. A. (1993). Geography of Michigan and the Great Lakes Basin. Dubuque, IA, Kendall/Hunt Publishing.

Schaetzl, R. (2006). GEO 333 Geography of Michigan course webpage, http://www.geo.msu.edu/geo333/retreat.html (accessed February 2006). Michigan State University, East Lansing, MI.

Scott, I. D. (1942). The dunes of Lake Michigan and correlated problems. *Michigan Academician* 44: 53-61.

Sherman, D. J. and S. Hotta (1990). "Aeolian sediment transport: theory and measurement". Coastal Dunes. K. F. Nordstrom, N. P. Psuty and R. W. G. Carter. West Sussex, England, John Wiley & Sons Ltd.: 17-37.

Shortridge, A. M. (2003). GEO 428 Lecture Notes, Fall 2003. Department of Geography, Michigan State University, East Lansing, MI.

--- (2004a). Assistant Professor, The Department of Geography, Michigan State University. Personal Communication: J. Bookout, East Lansing.

--- (2004b). GEO 425 lecture notes, Spring 2004. Department of Geography, Michigan State University, East Lansing, MI.

Smith, A. W. (1988). The Behaviour of Natural Sand Dunes. *Shore and Beach* 56(4): 24-29.

Smith, B. and D. Sandwell (2003). Accuracy and resolution of shuttle radar topography mission data. *Geophysical Research Letters* 30(9): 1467-1470.

Smith, S. L., D. A. Holland, et al. (2003). *The Effect of Changing Grid Size in the Creation of Laser Scanner Digital Surface Models*. International Conference on Geocomputation, Southampton, England.

Spinney, J. (2001). Report prepared for Applied Geomatics Research Group. *Environmental Application of LIDAR Data: Exploring High-Resolution Watershed Delineation.* Centre of Geographic Sciences.

Sun, G., K. J. Ranson, et al. (2003). Validation of surface height from shuttle radar topography mission using shuttle laser altimeter. *Remote Sensing of Environment* 88: 401-411.

Tabachnick, B. G. and L. S. Fidell (1996). Using Multivariate Statistics. New York, Harper Collins.

Thompson, T. A. (1992). Beach-ridge development and lake-level variation in southern Lake Michigan. *Sedimentary Geology* 80: 305-318.

Thompson, T. A. and S. J. Baedke (1995). Beach-ridge development in Lake Michigan: shoreline behavior in response to quasi-periodic lake-level events. *Marine Geology* 129: 163-174.

--- (1997). Strand-plain evidence for late Holocene lake-level variations in Lake Michigan. *Geological Society of America Bulletin* 109(6): 666-682.

Trimble Navigation Limited (2005). World Wide Web address: http://www.trimble.com (accessed February 2005), Copyright 2005.

United States Federal Government. (1972). Coastal Zone Management Act (CZMA) of 1972: 1451-1464.

VanOort, M., A. Arbogast, et al. (2001). Geomorphological History of Massive Parabolic Dunes, Van Buren State Park, Van Buren County, Michigan. *Michigan Academician* XXXIII: 175-188.

Warner, M. (2005). Staff Member, Land and Water Management Divison, Michigan Department of Environmental Quality. Personal Communication: J. Bookout, E. Lansing.

White, S. A. and Y. Wang (2003). Utilizing DEMs derived from LIDAR date to analyze morphologic change in the North Carolina coastline. *Remote Sensing of Environment* 85: 39-47.

Wilkinson, L., G. Blank, et al. (1996). Desktop Data Analysis with Systat. Upper Saddle River, NJ, Prentice-Hall, Inc.

Wise, S. (2000). Assessing the quality for hydrological applications of digital elevation models derived from contours. *Hydrological Processes* 14: 1909-1929.

Woolard, J. W. and J. D. Colby (2002). Spatial characterization, resolution, and volumetric change of coastal dunes using airborne LIDAR: Cape Hatteras, North Carolina. *Geomorphology* 48: 269-287.