ABSTRACT

DUDLEY II, THOMAS EDWARD. The Value of Landscape Amenities: Contrasting Feature Proximity and View. (Under the direction of Roger von Haefen).

Policy makers, home owners, and land developers are all interested in how home buyers value vegetation and other landscape amenities around homes. Forward-looking environmental policy makers look to maximize public welfare in the face of increasing population concentration by strategic preservation of parks and recreational areas. Homeowners, looking to sell, try to maximize home value by modifying their homes and lots, while developers, building new subdivisions, strive to create desirable landscapes. To efficiently accomplish these goals, all three groups need to understand buyer preferences for home landscapes. This thesis uses Geographic Information Systems (GIS) simulations to quantify views from homes and then incorporates these vegetation and other landscape variables into a hedonic model to provide willingness to pay estimates to interested groups. Methods are devised that address several shortcomings of previous studies and create more robust view variables. These vegetation view variables are contrasted with the more common proximity based variables. The results suggest that improved view simulations can yield view variables that are as effective, or slightly more effective, than more common proximity measures.

The Value of Landscape Amenities: Contrasting Feature Proximity and View

by
Thomas Edward Dudley II

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Master of Arts

Economics

Raleigh, North Carolina

2012

APPROVED BY:

Dr. Roger von Haefen
Committee Chair

Dr. Helena Mitasova

Dr. Laura Taylor

DEDICATION

In memory of

Dr. C. Dewey Cooper

(1/11/1924 - 2/28/2011)

BIOGRAPHY

Thomas was born in Raleigh, North Carolina to Gary and Norma Dudley on July 8th, 1987. After graduating from W.G. Enloe High School in 2005, Thomas pursed his Bachelor's degree in Economics at North Carolina State University. As a Master's student at North Carolina State University, Thomas developed an interest in Economic research.

Thomas currently holds an internship at the North Carolina Solar Center, where he continues to conduct economic research, with a focus on renewable energy.

ACKNOWLEDGMENTS

For their willingness to sit on my committee, I thank Dr. Roger von Haefen, Dr. Helena Mitasova, and Dr. Laura Taylor.

For his valuable feedback and guidance over the course of this thesis, I personally thank Dr. Roger von Haefen.

Dr. Mitasova's expertise and enthusiasm facilitated and encouraged my adoption of advanced GIS techniques. Creating and troubleshooting the GIS approach employed by this thesis would not have been possible without her guidance.

Also, I would like to thank several professors who most notably, over my undergraduate and graduate career, helped focus and clarify my writing: Dr. David Austen, Dr. Craig Newmark and Dr. Roger von Haefen.

Most of all, I would like to thank my wife Sarah for her support; both during the writing of this thesis and before.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Introduction

An empirical GIS method is described that simulates landscape features in three dimensions for homes in a portion of Wake County, North Carolina from 1983 to 2010. To quantify landscape impacts, several environmental variables from the GIS simulation such as vegetation in view, water features in view, and views of other structures are added to a traditional property value regression model. The results of these view variables are compared to those of the more common proximity variables. Using 16,706 residential sales, the logarithm of property value is regressed on home and scenic characteristics in an ordinary least squares regression.

By classifying vegetation based on height and distance from homes, significantly different impacts are found for different types of vegetation. At a far distance from the home, tall vegetation has a positive and significant impact. That is, greater views of tall vegetation over 100 feet away from the home tend to increase property value. However, this impact is reversed and more strongly significant for shorter trees at closer distances. For example, an additional close proximity tree of typical size within 50 feet corresponds to a 0.15% decrease in property value, which for an average home in the data set, corresponds to a $418 decrease in home value.

This thesis extends previous hedonic property value view studies by using higher resolution terrain data, incorporating views of buildings based on actual building height, and using a Digital Surface Model (DSM) instead of a Digital Elevation Model (DEM) for estimating views. By comparing statistically proximity and view variables for the same

homes, view variables are shown to be marginally more effective predictors of property value.

The following literature review highlights the methodological progression of proximity and view studies which incorporate measures of landscape composition. The approach, method, and data sections elucidate how limitations of previous studies are addressed. The results section includes and interprets regression results. Lastly, the conclusion describes policy implications, study limitations, and possible extensions.

Literature Review

In order to estimate the implicit prices of product characteristics such as landscape amenities, the hedonic methods are typically used (see Palmquist 1984 for an overview). These methods decompose the value of a commodity (e.g., household residence) into part worths associated with its heterogeneous characteristics using regression-based techniques. Rosen (1974) was the first to identify the necessary conditions for interpreting these part worths as marginal willingness-to-pays for the various attributes. Variation in the observed transaction prices of similar homes, and in landscape factors around homes, allow the implicit prices or effects of scenic factors to be estimated. As a result, buyer willingness to pay for landscape features is revealed. Accurately representing landscape attributes is required for credible identification of these values in the hedonic model. Since the first efforts to include and separate the impacts of landscape features in hedonic models in the early 1980s, the complexity of included variables has increased to improve variable accuracy.

Since the mid-1980s, hedonic models have been used to estimate the value of simple landscape measures including binary variables for parks within walking distance, or property

adjacency to lakes (Milon 1984, Kohlhase 1991, McLeod 1984). In most studies these measures show significant impacts, but to separate their impacts from those of other landscape features, and to more fully understand the causality of other landscape variables, later studies include additional variables. In addition to water features and home distance to parks, another major landscape feature is vegetation cover. Vegetation variables should be included in hedonic regressions and tested for significance for several reasons: vegetation around homes is obvious to home buyers, it provides aesthetic benefits, and it reduces views of other features. An early example of the inclusion of vegetation in a hedonic model, Anderson and Cordell (1988) counts trees for a very small sample of individual residential properties. The method was time and cost intensive, but showed the positive and significant impact of on-parcel trees on property values.

To estimate the scenic value of vegetation for large data sets, computer simulations are used. The first landscape vegetation simulations used proximity methods which are still common (Netusil et al. 2010, Drake-McLaughlin and Netusil 2010, Mansfield et al. 2005, Holmes et al. 2006, Sandler et al. 2010, Geoghegan et al. 2007, etc.). Proximity measures in these studies indirectly measure landscape composition by counting vegetation within set distances from properties. Vegetation variables indicate that trees are within distance ranges, but cannot confirm if those trees have a scenic impact on views from the property. Despite this shortcoming, proximity studies typically have consistent results for vegetation value and find a near consensus that proximity to water has a positive and significant effect on property value (See Table 1).

By employing more advanced GIS methods, view studies simulate what areas of trees, open space, structures and other features are in view from study homes. These studies should more accurately quantify landscape characteristics by teasing out which features have visual impacts on home views. Unlike proximity studies, view studies typically have inconsistent findings for vegetation and the impact of water is never estimated as positive and significant as expected (See Table 2). Given the consistency of proximity-based landscape studies, technical and methodological limitations may explain the underperformance of view-based studies. Review summaries of previous proximity and view studies are provided in Tables 1 and 2.

Proximity-based Studies

Netusil et al. (2010) examine a vast data set of 30,015 housing transactions from Portland, Oregon from 1999-2001 and derive proximity-based landscape variables to quantify the impact of forest canopy on property values. The author's results suggest that 40% tree canopy coverage is ideal for maximizing property value. Their positive and significant estimate for the value of proximity to water is consistent with other proximity studies (Sandler et al. 2010, Holmes et al. 2006, Geoghegan et al. 2007, etc.). However, one possible study limitation is by including only canopy Netusil omits tree height variables which Mansfield et al. and Joly et al. have found to be significant. Also, Netusil's two-stage regression models natural features in a large one-quarter mile.

The landscape diversity of Netusil's study region makes Netusil's paper ambitious. Netusil indirectly accounts for changing topography by geographically grouping homes into six regions around the city. Her varying results for canopy appreciation among regions may

be the result of elevation variation. If hillier regions reduce the amount of canopy in view compared to flat sectors, significant attenuation bias may occur because the variable includes canopy out of view. Incorporating terrain standard deviation, total area in view, or canopy in view for each study region could model the variation in topography, and possibly explain the inconsistent canopy impacts.

Netusil's results suggest dense canopy shades out too much light and prevents growth of smaller trees or ground vegetation. Netusil's conclusion that the benefits of canopy are susceptible to diminishing marginal utility after 40% coverage is the major contribution of the paper. Measuring vegetation height as other studies do could test this hypothesis by measuring the interaction between short and tall vegetation. Netusil's land cover canopy data was derived from LiDAR (Light Detection and Ranging), which measures vegetation height in high detail. However, the preprocessed canopy layer classifies areas in binary form, as canopy or not and therefore excludes the continuous measure of vegetation height.

In addition to the height of proximate vegetation, its distance from the study home can also significantly affecting property value. Within the one-quarter mile radius Netusil uses, other studies (Joly et al. 2009, Holmes et al. 2006, Sandler et al. 2010) show marked differences in how people value vegetation and open space around their homes, and one study suggests an inner radius as small as 70 meters. Using multiple radii to model variable or diminishing returns to vegetation over distance could model buyer preference more accurately.

Improving on Netusil's previous work, Drake-McLaughlin and Netusil (2010) model the impacts of three proximity variables on property values: vegetation on the property,

vegetation in the surrounding neighborhood, and home proximity to urban amenities. Vegetation height is introduced as a categorical variable and vegetation distance is discretized into radii. In other words, in addition to one variable quantifying canopy within one-quarter mile around each home, variables represent vegetation within 200 feet, and between one-quarter mile and one-half mile. Perhaps the most significant methodological improvement of this study is the incorporation of LiDAR. This high resolution terrain and vegetation data quantifies vegetation height as a continuous instead of a binary variable, as in Netusil's previous study. Possible study short-comings are that too few categories for tree type or height are used, the policy implications of increased accessibility to amenities are unclear, and view is not considered.

Compared to Netusil's previous work, Drake-McLaughlin and Netusil utilize a slightly larger and newer property transaction data set of 30,786 single family homes. The study region, econometric techniques, and time period are all very similar. All observations are within the city of Portland, Oregon, but those within 1/2 mile of city limits are excluded. They use a high resolution, unprocessed LiDAR DSM for categorizing vegetation height. Using LiDAR vegetation data in the second regression increases R^2 values from 0.755 to 0.763. Also, on average, vegetation estimates are more significant with the more detailed terrain data. Drake-McLaughlin and Netusil are the first to use this data for a hedonic property value study incorporating vegetation proximity measures. To separate trees from bushes, vegetation is categorized as either high or low. Composition percentages are computed for several proximities, including the amount of on-property vegetation and all

vegetation within 200 feet, between 200 feet and a 1/4 mile, and between 1/4 mile and 1/2 mile.

Quadratic terms are used to model the impact of vegetation because additional vegetation is expected to increase values of property with spare vegetation, but decrease selling prices for already densely vegetated plots. This affirms the conclusion of Netusil's first paper, that in urban areas trees are desired while in more rural areas increases in tree cover crowd out lower vegetation. This is also in line with the findings of Geoghegan et al. (1997) and Sandler et al. (2010) that people value landscape diversity. Both low and high vegetation have positive impacts, but it is possible an increase in one at a cost of the other could account for higher or lower selling prices. An interaction term could be included to shed light on this relationship.

Drake-McLaughlin and Netusil's conclusions from modeling water are more robust than results from other studies. Intuitively, there are two factors to consider when modeling water: the potential for flooding and its positive aesthetic value. Netusil shows additional water on a property has a negative impact, because of an increased chance of flooding. However, if water is in the 1/4 mile or 1/2 mile buffers, it has positive value as expected. Confirming empirically that proximity to water has a positive effect on property values suggests the authors measure landscape features accurately and use an appropriate regression. Assuming view is a more accurate measure of the scenic benefits of vegetation around homes, view studies should be expected to find results as robust as those of Drake-McLaughlin and Netusil.

Mansfield et al. (2007) separate the landscape benefits of trees on parcels such as increased privacy and improved aesthetics from the benefits of close contiguous forest, such as recreational opportunities and wildlife habitat. By modeling the features separately, the economic values of their corresponding amenities are independently quantified. Mansfield et al. use comparatively simple GIS proximity methods to create measures for vegetation to include in their hedonic regression. In particular, they differentiate trees on parcels from large contiguous forests by boxing off areas of large forest, classifying them as private or institutional forest and measuring their distances to homes. To quantify landscape vegetation cover on each study parcel, a low resolution GIS remote sensing measure called the Normalized Difference Vegetation Index (NDVI) is calculated for each study parcel. NDVI estimates how "green" individual parcels appear on average by measuring the reflections of near-infrared light (Prasad et al. 2005). In addition to average NDVI, the distances to private and institutional forests are included in a comprehensive economic model and a hedonic regression disentangles their impacts.

By disentangling on-parcel tree from forest impacts, Mansfield et al. find that both have significant positive effects. However, comparing Mansfield et al.'s results to those of other works is difficult because the impact of NVDI on property value is measured as a per standard deviation change in NVDI. A discussion of what would increase NVDI by one standard deviation is not provided. Also, the NVDI index cannot be translated into area of tree cover per parcel or vegetation coverages as other studies measure. An additional weakness of NVDI is that it weights conifers, such as fir, pine and hemlock trees more heavily than hardwoods because the darker green coniferous leaves reflect more near infrared

light (Torma et al. 2008). In several studies (Holmes et al. 2006, Donovan et al. 2010) conifers are shown to have smaller or even negative aesthetic value compared to hardwoods. As a result of overweighting conifers in NVDI, and conifers not being considered separately from hardwoods, Mansfield et al. may understate the positive impact of hardwoods. Higher detail remote sensing and image processing would be required to test this hypothesis. Instead of measuring NVDI, similar remote sensed data and established light reflection patterns could be used to estimate vegetation coverages and identify tree species composition per parcel.

Mansfield et al. use GIS mapping and remote sensing to estimate the economic impacts of on-parcel landscape vegetation and proximate forest. An aspect of vegetation differentiation not directly addressed is the contrast of in view and out of view forest. Mansfield et al. state, "Separating the effect of visual improvements from forest proximity can be quite difficult." The NVDI measure is appropriate for the large geographic scope of Mansfield et al.'s region, however much higher resolution data is needed to quantify the visual impacts of vegetation.

View-based Studies

Compared with counting vegetation in proximity, simulating vegetation in view should more accurately quantify landscape features around homes. Compared to proximity-based studies, view-based studies require greater programming complexity, processing time, and storage space. To work around these issues, view studies typically use lower resolution terrain surfaces, fewer land cover classifications, and smaller data sets. The inconsistent, and often weak estimates of water's impact from view studies suggest that the application of view

methods in the hedonic property literature is not fully developed (See Table 2). Two view studies are examined to show methodological improvements in the literature and to note the limitations of the most recent view-based hedonic property value study.

For a sample of 504 homes, Paterson and Boyle (2002) create view and proximity landscape variables using four land cover categories. Independent of flood probability, the negative and significant impact of hydrological features may suggest specification problems. In common with all other works that estimate landscape view, Paterson and Boyle derive views in GIS on the DEM instead of on the DSM. This leads to the same shortcomings as Joly et al. and Germino et al., most notably an overestimation of area in view. As a result of the age of the study, there are several areas for GIS methodological improvement.

Only four landcover types are used to classify areas around homes as either forest, developed, agricultural or water. As a result of vegetation in developed areas being excluded from vegetation counts, on-parcel and neighborhood landscape vegetation is ignored. For forest views, tree height, composition, and distance from each home are not considered. Given the scope of hedonic landscape studies, land classification data lacks sufficient detail to be appropriate where DEM data is available (See Figure 9).

Paterson and Boyle do not quantify variation in consumer preference within one kilometer. Variation of impacts of vegetation within one kilometer has been shown to be significant in other studies. (Joly et al. 2009, Holmes et al. 2006, Sandler et al. 2010) Paterson and Boyle's one kilometer radius is even larger than the problematic one-quarter mile radius Netusil uses.

Several of the results the authors reach raise model specification concerns. For example, the impact of hydrological features (lakes, bays and rivers) is estimated negative and significant. In most proximity studies water is found to have the expected positive impact. (See Table 1) Insignificant or counterintuitive estimates for water are common in view-based hedonic property value regressions and Paterson and Boyle's result is similar to that of Joly et al. However, Paterson and Boyle include a binary variable to model homes within FEMA (Federal Emergency Management Agency) flood zones. If the authors' assertion that the decrease in property value associated with water proximity is the result of increased flood probability, then the FEMA variable should absorb that negative impact as in Bin et al. (2008). For Paterson and Boyle, the inclusion of the FEMA term does not reverse the impact of the water term and the authors insist on their interpretation, but offer no further explanation.

The counterintuitive sign and insignificance the authors find suggests model misspecification. Another possible cause would be measurement error as the result of low resolution mapping of the landscape around homes. (See Table 2) For Paterson and Boyle, trees, buildings, and any other features besides the DEM are not capable of blocking views. The result for estimating views of water is that far more homes are credited with being within view of water than are in reality. The unmentioned and assumedly limited viewshed resolution may also conflate this problem.

Joly et al. (2009), the most advanced view-based hedonic property value study, utilizes a much more complex GIS method than that of Paterson and Boyle. Based on the sales of 4,352 homes in the urban fringe of Dijon, France, Joly et al. combine DEMs with

satellite imagery to extract the effects of nearby land uses on residential property values. Joly et al. use view simulations over a large, diverse study area, and incorporate the variables they derive into a developed economic model. However, several meaningful landscape factors are excluded and a limited resolution is used for view simulations.

The first of Joly et al.'s innovative technical considerations is the assignment of estimated average heights to different land cover types to more accurately predict landscape visibility from each property. For example, in landscape simulations all areas classified as woodland are assigned an average height so the area of woodland in view can be counted and woodlands will block views of other features. Earlier view-based studies do not consider vegetation this early in their methods (Germino et al. 2000, Paterson and Boyle 2002, Tyrvainen and Miettinen 2000). Without first simulating vegetation, most studies estimate area in view as if views are only blocked by the surface of the earth (See Figure 1). In these earlier studies, areas that are estimated to be in view are classified as vegetation, lakes or other features based on data from separate GIS layers. By including vegetation in the view simulation, Joly et al. estimates area in view much more realistically.

For Joly et al. perhaps the most surprising result is that the view of water has a negative but insignificant effect on transaction price. For most hedonic landscape studies, proximity to water is a significant and very positive variable (Leggett and Bockstael 2000, Sandler et al. 2010, Netusil 2009, etc). The authors justify this counterintuitive sign as a result of a very limited sample of properties with water visibility. Joly et al. mention the higher likelihood of flooding in areas close to water but do not strictly test this hypothesis, as Paterson and Boyle do; and therefore theirs remains a feasible explanation. Also, suggesting

they doubt their conclusions on water's impact, the authors cite several studies that show water's positive effect when estimated from larger samples. Joly et al. also acknowledge limited data availability in their study region and suggest the use of more advanced remote sensing technologies and LiDAR in future research.

In summary, proximity and view studies show progress estimating the value of vegetation and other landscape features. Proximity studies are more common and typically have stronger, more consistent conclusions. The adoption of more recent GIS view methods has improved view-based hedonic studies, while several major limitations still exist such as low terrain detail, small sample sizes, and unrealistic viewsheds. These limitations may prevent view studies from achieving results as convincing as those of proximity studies.

The following Approach section describes this paper's broad direction and how shortcomings of previous studies are addressed. Next, the Method and Data sections describe how proximity and view variables are generated in GIS software, incorporated into an economic regression, and statistically compared.

Approach

To homeowners, views of vegetation around homes yield the primary benefits of vegetation: aesthetic value and increased privacy. Vegetation in proximity is not necessarily in view and therefore may not accurately quantify the major benefits of vegetation. As a result, measures of vegetation in view should more accurately predict property values than measures of vegetation in proximity. However, this has not been confirmed by current view studies. None of the existing view studies find results that are as convincing or as well developed as those from their proximity based counterparts. This thesis includes variables

omitted from previous studies, utilizes more advanced view simulation procedures from the GIS literature, and furthers GIS vegetation visibility simulation methods to create a view-based model that marginally outperforms its proximity-based counterpart.

In order to minimize attenuation bias, ideal variables for hedonic regressions accurately quantify the intended features from which benefits are assumed to be derived. The major benefits of trees nestled around homes are improved scenic quality and increased privacy. Traditionally, to quantify these benefits, proximity methods and variables have been used to count tree cover in radii around homes. However, improved scenic quality and increased privacy are benefits derived only from trees that are in view, and not necessarily from those in proximity. Using proximity variables such as tree cover within 200 meters, is less than ideal for quantifying scenic quality because the "scene" around a home is derived from its view of, not its proximity to features. Likewise, the second major benefit of trees, increased privacy is only affected by changes in the features in view, but not necessarily proximity. Especially in diverse landscapes, trees in proximity may be a poor estimate of the primary benefits of vegetation, and therefore a view method should be used to estimate scenic quality and privacy.

To facilitate variable interpretation, all vegetation view variables in this study are derived in order to separate out other factors and to facilitate comparison between geographic features. Streets, views of other buildings, hydrological features and tree height are used to model impacts. All variable coefficients are interpreted as the percentage change in property value from one cell of that variable compared to one 3x3 feet cell of open space.

To categorize trees based on view, and to answer the questions this thesis poses, a more complex GIS method is required than that used by Mansfield et al. A tree right next to one in view may be out of view, and therefore should be excluded from that view variable. To create vegetation variables, only areas in view that correspond to vegetation are categorized based on tree height and distance from the home. For each observation, totals for each vegetation height and distance are counted. Then, in a similar manner to Mansfield et al., the variable values are included in a hedonic economic model.

Using a DEM, Joly et al. and Germino et al. overestimate area in sight and vegetation in sight because trees and buildings cannot block views. Over-counting vegetation may lead to underestimating the economic impact of that vegetation. Therefore, when using viewsheds to quantify views of vegetation from homes, view simulations (viewsheds) should be conducted on the DSM instead of the DEM. To my knowledge, this thesis is the first application of DSM viewsheds in a hedonic property value model. As an additional benefit of using DSMs, they typically are a higher resolution than DEM data. For example, the typical DEM resolution from the proximity and view studies in Tables 1 and 2 is 30 meters, while DSM LiDAR has an average resolution of approximately 3 feet. Using low resolution terrain data and the default viewshed settings in ArcGIS, total area in view is overestimated (See Figure 1). A further benefit of DSM terrain data is it can be used to map building height in the view simulation. Few of the studies reviewed included buildings in their simulations and none used the high resolution DSM to map building height. This technique maps in the simulation the slanted or irregular tops of buildings instead of assuming they are flat as

previous methods do. Tests were not conducted to examine the impact of using the DSM derived buildings compared to those from previous methods.

Compared to the articles discussed in the Literature Review, the focus of attention for this thesis is much closer to study homes. From the GIS perspective, focusing on such a small scale around homes requires views to be simulated on the DSM instead of a DEM (See Figures 1 and 2). The following Method section describes the workflow required to create view and proximity variables.

Method

For homes in the study region, building and property characteristics from Wake County's parcel database are imported into ESRI ArcGIS 10. Based on the location of each home, GIS analysis is performed for the DSM and DEM terrain data within 1/8 of a mile of that home. Parameter values are calculated such as total area in view, area of water features in view, and most importantly vegetation. Around each home, the height of vegetation and structures are mapped for every 3x3 feet cell by subtracting the height of the DEM from the DSM. Vegetation in view is grouped based on height and distance from the home (Figures 7 and 8). The corresponding vegetation variables for proximity are also calculated and grouped. These variables are then processed by Python scripts and then imported into SAS for regression analysis. (See Data Section)

Compared to previous hedonic property value studies, this thesis uses smaller classification divisions for vegetation height and distance from the home. In previous studies the typical distance for the smallest radius in which to count vegetation is between 50 and 200 meters. Increasing the number of divisions within this distance creates more variables for

testing, and the significance of several of these variables suggest variation in vegetation impact closer to homes than in previous studies. Increased processing time and data storage space are the two drawbacks of using this number of classifications for vegetation.

Only Drake-McLaughlin and Netusil (2010) categorize vegetation based on height. The authors use two categories, short and tall. Here, vegetation is grouped into three categories, short (2-15 feet), medium (15-30 feet), and tall (50-100 feet). After initial significance tests, tall vegetation was shown to have the most variant impact over distance. To model these impacts more accurately, view and buffer variables for tall vegetation were divided using ranges, 0-10, 10-25, 25-50, 50-100, 100-200 and 200-660 feet from each home.

Deriving vegetation from the DSM requires excluding non-vegetation features. Streets, buildings and water features are extracted and not included in the vegetation counts. Cells in view that are classified as these features are counted separately and included in the regression equation.

To model changing preferences for vegetation over areas of varying population density, urban density is included in the economic model. The density term is created by a method Domanski (2010) uses which counts the number of dwellings within 1 mile of the study home.

By including variables to model local region composition, incorporating higher resolution terrain data, and using a DSM to simulate views, the method described above illustrates that it is possible for view metric to outperform the more common proximity measures.

Data

This thesis uses the geographic subset of Wake County, North Carolina shown in Figure 11. Wake County includes North Carolina's capital city, Raleigh, and several smaller municipalities. Wake County has a population of over 900,000 and is part of the Research Triangle Park, one of the largest and highest growth research and development regions in the United States. The study region includes parts of the City of Raleigh, the Town of Cary, and unincorporated areas. This extent was chosen because it corresponds to one DSM LiDAR file, which expedites the GIS analysis process. The region also includes variation in population density which is used to test for varying scenic preferences based on urban concentration.

In the study region, homes considered are limited to those sold between January 1st, 1982 and December 31, 2010. Additionally, only single family residential homes with lot sizes greater than 0.05 acres, and transaction prices greater than $30,000 are included. A total of 16,706 observations are used.

The standard or benchmark hedonic regression requires transaction and property information, such as transaction price and date, lot size, home square footage, and home geographic coordinates. (See Table 3) These and other variables from the Wake County Tax Office database are publicly available online at the Wake County Real Estate Records website. Because Wake County's database includes home coordinate information, georeferencing homes to their spatial locations is not required. To map average commute time per census tract, data from the U.S. 2000 Census is used. In GIS, census tract boundary

files are used to assign average commute times to each home based on its corresponding census tract.

The GIS analysis combines data including structure footprints, roads, hydrology, bare earth DEM LiDAR, and DSM LiDAR. Building footprints in vector polygon format, roads in vector line format, and hydrological features in vector polygon format are obtained from the Wake County GIS website. Building footprints are only available for structures in Raleigh. The population density variable counts the number of homes within one mile of the study home and is derived from counting the number of building footprints within that distance. To increase the efficiency of GIS calculations building, road, and hydrological feature layers are converted from vector to raster format (See Figures 12, 13). Computationally this approach is much faster than using the vector data and is recommended for modeling urban and mixed urban environments. (Ratti and Richens 2004)

Bare earth elevation data for Wake County is obtained from the North Carolina Flood Mapping Program website. This DEM information is LiDAR derived and therefore, highly detailed. The DEM is interpolated and converted to a raster at a resolution of 3 feet^2.

This thesis takes advantage of recently available raw DSM LiDAR data. As illustrated by Figure 10, North Carolina is an excellent location for GIS hedonic studies that require DSM LiDAR because of the state's unusually high LiDAR coverage.

Using inputs such as the DEM, DSM, water features, streets, and buildings, the GIS process creates 23 landscape variables. For each observation the full process requires 51 geoprocessing steps and ex ante and ex post data handling. Typically in ArcGIS, complex processes are completed in ModelBuilder, which gives the user an interface to visualize the

relationships between functions and their inputs and outputs. In ModelBuilder an entire multistep process can be performed with one mouse click, but only for one observation. For this thesis running the process in ModelBuilder separately for each observation would not be time efficient. Also, a goal of this thesis was to create a GIS method that other economists could adopt to create view variables for even larger datasets and for that purpose resorting to ModelBuilder would not be appropriate.

Instead, Python scripting was used to complete the GIS process, one observation at a time, for all homes within one folder. In this thesis, Python scripts are primarily used to execute ArcGIS geoprocessing functions. Additionally, because ArcGIS processes are limited in the data formats they can input and output, Python scripts are used to perform required data manipulation before and after the main GIS process.

For instance, ArcGIS requires the geographic locations of study homes to be input as individual "shapefiles." Originally, each property's information was stored as a row in Wake County's real estate database. In Python, transforming each database row to a separate shapefile required converting each row to its own text file, then to a GIS layer file, and then to a shapefile. Working with large folders significantly slows the main GIS process and therefore another script groups shapefiles into folders of 100. Typically, completing the GIS process took 8 hours per folder of 100 observations. Completing the process one folder at a time would take approximately 56 days.

Multithreading the process in Python significantly decreased this computational time. By using multithreading (or multiprocessing) all processor cores can be used to complete different tasks, as long as the separate processor cores do not simultaneously attempt to use

the same files. The computer used for this thesis had six processor cores and simultaneously ran the GIS process for five observations. (The sixth core was reserved for running the operating system.) To ensure files like the DSM and DEM data were not required by multiple processes simultaneously, five hard drives containing all files needed for the process were used, effectively dedicating a hard drive for each core. Using multiple hard drives also ensured hard drive read and write speeds were not limiting the process speed.

The output files from the GIS process are binary matrixes in text file format. Each matrix maps selected terrain characteristics around each home. Each output matrix is 440 by 440 cells; each corresponding to a 3x3 feet cell of terrain around that study home. The home itself would be in the very center of the each matrix. As an example, for one study home the view, tall vegetation, output file shows a value of "1" for each cell of tall vegetation that is in view from that home. All other matrix cells have the value "0". For each vegetation type, another Python script is used to count all cells of value "1" in distance ranges from each home. These totals are added as new variables to the original real estate characteristics database. Once the variable values are in the database SAS processes them like other property characteristics.

The data run for this thesis was prepared in separate batches and the processes were restarted manually after each batch of 500 observations was completed. For the 16,706 observations, the total processing time required for data handling, and the GIS process was not recorded, but is estimated at one month.

Results

Regression results are described for four specifications. First, is the benchmark economic model that includes the home and transaction information described in the Data section and in Table 3. Landscape variables are not included in the standard regression. The second regression adds all proximity variables reported in Table 4. The view regression replaces vegetation proximity variables with view variables reported in Table 5. The final regression tests for differences between the proximity and view variables based on vegetation height and distance from homes.

The benchmark economic specification takes the form:

$$Ln(Price_j) = \beta_0 + \sum_{i=1}^{N} \beta_i X_{ij} + \varepsilon_j$$

Where the vector X_{ij} corresponds to property and transaction variables included in Table 3. For the proximity regression, the vector of landscape proximity variables from Table 4, labeled X^P, is added.

$$Ln(Price_j) = \beta_0 + \sum_{i=1}^{N} \beta_i X_{ij} + \sum_{k=1}^{M} \beta_k X_{kj}^{P} + \varepsilon_j$$

All regressions include fixed effects for each one year period in the data set. The third regression using view variables instead of proximity variables, replaces X^P with the view variables from Table 5, labeled X^V.

$$Ln(Price_j) = \beta_0 + \sum_{i=1}^{N} \beta_i X_{ij} + \sum_{k=1}^{M} \beta_k X_{kj}^{V} + \varepsilon_j$$

Regression outputs for the three regressions are found in Table 6.

The benchmark regression yields several intuitive results. The logarithm of home heated area has a positive, large, and significant parameter value, estimating a 1% increase in home heated square footage would increase property selling price by 0.51%. Age has a negative and significant effect suggesting older homes are less desirable, ceteris paribus. Distances to major areas of interest like downtown Raleigh and the Research Triangle Park have negative and significant impacts. For this region, RTP is the major employment center, but smaller centers exist and to model buyer preference for homes close to workplaces, commute time averages per census tract are added. Average commute time also has a negative impact and is modeled with a quadratic term. Homes sold in the summer tend to sell for more and as a result, the deed date binaries for Spring, Fall, and Winter are negative and significant. The R^2 and adjusted R^2 values for the base regression are 0.6461 and 0.645.

From Table 4, eleven vegetation proximity measures are added to the base regression to create the view regression. Also, estimates for buildings, streets and hydrology are included. The R^2 and adjusted R^2 values improve slightly to .6468 and .6454. The extent of this improvement in R^2 as the result of incorporating landscape features is roughly equivalent to the improvements of Mansfield et al. (0.47 to 0.48) and Drake-McLaughlin and Netusil (0.756 to 0.763). In general the signs and levels of significance of several of the landscape variables suggest that vegetation within 50 feet of homes is undesirable, while more distant,

taller vegetation appears beneficial. Also, as the distance from the home increases, the parameter values tend to decrease. This suggests that independently of whether a certain type of vegetation has a positive or negative impact, it declines when the vegetation is farther away from the home.

As described in the Method section, building, street, hydrology, view and buffer variables are interpreted as the percentage change in home price by one additional cell of that variable. A medium size tree in the 0 to 50 feet range from a home would typically correspond to 30-40 cells. This would suggest a 0.005% to 0.067% decrease in property value, or on average, $141 to $188. Alternatively, an additional tall tree between 100 and 200 feet from the home corresponding to 40-60 cells, would tend to increase property value by between 0.0458% and 0.0687%, or between $128 and $192.

For predictive power, the view regression shows a slight improvement over the buffer regression. The R^2 and adjusted R^2 improve slightly to .6470 and .6456. The view regression finds similar, but typically more substantial estimates of the value of landscape vegetation. Using a similar example, the complete view of a small tree in the 0 to 50 feet range from the home would typically be captured as 8-15 cells of medium height vegetation. Planting a tree of this size would correspond to a 0.10% to 0.195% decrease, or for the typical price home in the data set, between a $290 and $546 decrease in property value. In contrast, the complete view of an additional tall tree between 50 and 100 feet from the home would have 15-20 cells in view, and correspond to between a 0.19% and 0.26% increase, or between a $546 and $728 increase in property value.

Correctly estimating the impact of water has been challenging in previous view studies. (Joly et al. 2009, Paterson and Boyle 2002) Using one of the most advanced view simulations, Joly et al. could not verify the expected positive and significant impact of water. The view regression here shows water's positive and significant impact. The average home with a view of water, compared to one without a view of water would sell for $2,420 more, a .86% increase. Previous studies such as Bin et al. (2008) have shown that the frequent positive correlation between the potential risks and benefits of water proximity can complicate estimating its amenity value. By studying a low flood risk region, this thesis estimates water's high amenity value.

Comparing regression results for proximity and view regressions verifies several hypotheses. A maintained assumption so far has been that proximity and view variables are not highly correlated with each other for the same height and distance vegetation. If there is high correlation, there would be little reason to justify the additional steps and computational time required to produce view variables.

For this data set, the average correlation coefficient between proximity and view variables decreases substantially over distance (See Table 7). Intuitively it seems reasonable the correlation between view and proximity variables would decrease over distance as the area of vegetation in view decreases. Low correlations such as these do not by themselves justify the use of view variables, but they do suggest view variables quantify vegetation differently from proximity variables.

In addition to the results for vegetation, the regressions show interesting effects for other features. For example, *Buildings* in view are uncorrelated with urban density. We

would expect in general that a house where houses are closer together would have more of other structures in view. A possible explanation for no correlation between these factors is that in more urban areas, more trees are planted to reduce views of other structures.

In both regressions *Streets* and *Buildings* show insignificant results. This was unexpected, but is not a major concern. The primary purpose of including the *Streets* and *Buildings* variables was to exclude their counts for those of vegetation. Positive insignificant estimates for these variables may indicate that homes with larger total areas in view sell for more and these homes happen to be within sight of more streets and other buildings.

The final regression tests for differences between corresponding view and proximity variables. A new term θ_k is introduced as the difference of corresponding proximity and view variables.

$$\theta_k = \beta_k^P - \beta_k^V$$

For each vegetation type, a θ_k value significantly greater or less than zero indicates the more accurate view variables' estimate of that vegetation's impact is distinct from the proximity estimate. Significant differences between multiple view and proximity terms would suggest willingness to pay for vegetation view and proximity are distinct. When positive, the θ coefficient suggests the view variable parameter is larger than the corresponding buffer parameter and the reverse if θ is negative.

To facilitate θ_k parameter and standard error estimation in SAS, θ_k is included in the regression explicitly. This is achieved by substituting $(\theta_k + \beta_k^V)$ for β_k^P in a regression including proximity and view variables.

$$Ln(Price_j) = \beta_0 + \sum_{i=1}^{N} \beta_i X_{ij} + \sum_{k=1}^{M} \beta_k^P X_{kj}^P + \sum_{k=1}^{M} \beta_k^V X_{kj}^V + \varepsilon_j$$

$$Ln(Price_j) = \beta_0 + \sum_{i=1}^{N} \beta_i X_{ij} + \sum_{k=1}^{M} (\theta_k + \beta_k^V) X_{kj}^P + \sum_{k=1}^{M} \beta_k^V X_{kj}^V + \varepsilon_j$$

$$Ln(Price_j) = \beta_0 + \sum_{i=1}^{N} \beta_i X_{ij} + \sum_{k=1}^{M} \theta_k X_{kj}^P + \sum_{k=1}^{M} \beta_k^V X_{kj}^P + \sum_{k=1}^{M} \beta_k^V X_{kj}^V + \varepsilon_j$$

The last regression disentangles the impacts of view and proximity and takes the form:

$$Ln(Price_j) = \beta_0 + \sum_{i=1}^{N} \beta_i X_{ij} + \sum_{k=1}^{M} \theta_k X_{kj}^P + \sum_{k=1}^{M} \beta_k^V (X_{kj}^P + X_{kj}^V) + \varepsilon_j$$

Where:

$$\theta_k = \beta_k^V - \beta_k^V$$

In SAS the estimates (θ_k) for the proximity variables (X_{kj}^P) are interpreted as the difference between view and proximity variables. The parameter estimate β_k^V is calculated for the new variable, ($X_{kj}^P + X_{kj}^P$) but does not have a meaningful interpretation. However, its inclusion is necessary to maintain the stated interpretation of θ_k.

Within 50 feet from homes, compared to proximity estimates, view estimates of short (2-15 feet), and medium (15-30) height vegetation are smaller and the difference is significant at the 5% level. Close, short to medium height vegetation in view may have a more negative impact than similar vegetation at the same distance not in view because vegetation close to homes reduces total area in view significantly. Privacy provided by vegetation blocking other buildings or roads may increase property values but vegetation too

close may make the homes feel closed in. Also, too much vegetation too close would not allow space for yards. As further evidence, the view estimates for Tall vegetation between 50 and 100 feet from homes, are larger than those for proximity, and the difference is significant at the 10% level. In that range, buyers prefer trees to be in view, presumably to increase privacy while not closing in the home with vegetation. Of the vegetation categories that show significant differences between view and proximity variables, the view variables have greater magnitudes. The close variables that have negative impacts show greater negative impacts while the positive, farther tall vegetation estimates show an even more positive impact using view variables.

Conclusion

A review of GIS view methods from previous hedonic property value studies finds several common limitations. These shortcomings are addressed by using a large data set, incorporating higher resolution terrain data, including views of buildings based on actual building height, and simulating vegetation blocking views. Incorporating these view variables into an economic regression generates significant differences between several view and proximity variables. In this regression of 16,706 property transactions, landscape view variables marginally outperform proximity variables.

This result provides several policy recommendations to parties interested in buyer willingness-to-pay. When estimating the marginal economic impact of vegetation on property values, vegetation view should be considered. For local governments performing analyses to estimate the property value increases associated with creating or maintaining city parks, a relevant policy recommendation of this thesis is that view simulations should be

conducted. Instead of using simple proximity, taking into account the additional views of vegetation from homes could rank more highly park locations with greater visual impacts.

Homeowners should be aware that additional trees do not necessarily enhance their property values. Open space is preferred to additional close trees when they would block views of other features. At farther distances from homes, views of trees, especially tall trees, increase property values.

Although the GIS methods employed in this thesis represent a significant improvement over preexisting methods for representing landscape amenities, they nevertheless have several shortcomings that can be improved on in future work. For example, views are simulated for each study home at its estimated geographic coordinates in the real estate database. When these coordinates are not exactly placed on the study home, views are simulated from a side or a corner of the home. In these cases, the home itself blocks simulated views. Measurement error is introduced when vegetation is incorrectly excluded from view as a result of the inaccurate view blocking of the structure.

Instead of relying on the coordinates estimated in the real estate database, further research could address this first limitation by deriving view locations from home structure footprints. This would require several additional steps, for example, linking the footprint and property information layers with a spatial join. This would allow the view location to be set as the centroid of the structure. Temporarily deleting the study home from the simulation would be required to prevent the building from blocking the view from the new viewpoint inside it.

Another study limitation is that for each 3x3 feet cell, trees are simulated as the volume between the DSM and DEM layers. Currently, when vegetation height is derived for each cell, view blocking vegetation is simulated for the volume of that cell below that height. This simulation is not accurate for pine trees, which have nearly all of their branches and needles at the top. For example, if a line of mature pine trees separates a study home from a neighbor's home, the view simulation estimates the neighbor's home as not in view because all areas below the canopy of the pines are simulated as being view-blocking vegetation. However, in reality most of the neighboring home will still be visible, with only the trunks of the pine trees partially obscuring the view.

Adding tree species composition could test additional hypotheses about willingness to pay and could address this second study limitation. Remote sensing data derived from Wake County's high resolution orthophotos could be used to test which tree species are most desirable. Additionally, layers mapping pine locations could be used to create view simulations that assume low views are not blocked by mature pines.

To maximize the number of observations considered for the study region, real estate data from 1983-2010 were considered. Ideally, updated LiDAR derived vegetation layers would be used for all years in this time period. However, the DSM data is only available from 2001. Over this large time span, the development of previously forested areas and vegetation growth could create measurement error for vegetation terms. Further works could minimize this error by using smaller time periods and larger geographic areas to achieve sufficient data sets. Also, incorporating the time between the real estate transaction and the LiDAR measurement could be used to simulate vegetation growth over time.

REFERENCES

Anderson, L. M., and H. K. Cordell. “Influence of trees on residential property values in Athens, Georgia (U.S.A): A survey based on actual sales price.” *Landscape and Urban Planning* 15 (1988): 153-164.

Anselin, Luc. “GIS research infrastructure for spatial analysis of real estate markets.” *Journal of Housing Research* 9 (1998): 113–33.

Anselin, Luc. “Some robust approaches to testing and estimation in spatial econometrics.” *Regional Science and Urban Economics* 20 (1990): 141–163.

Bin, O., J. B. Kruse, and C. E. Landry. “Flood hazards, insurance rates, and amenities: Evidence from the coastal housing market.” *Journal of Risk and Insurance* 20 (2008): 63–82.

Bourassa, Steven C., Eva Cantoni and Martin Hoesli. “Predicting house prices with spatial dependence: A comparison of alternative methods.” *Journal of Real Estate Research* 32 (2010): 139-160.

Cheshire, Paul, and Stephen Sheppard. “On the price of land and the value of amenities.” *Economica* 62:246 (1995): 247-267.

Domanski, Adam Matthew. “Essays in nonmarket valuation with applications to environmental economics.” Diss. North Carolina State Univ., 2010.

Donovan, Geoffrey H., and David T. Butry. “Trees in the city: Valuing street trees in Portland, Oregon.” *Landscape and Urban Planning* 94.2 (2010): 77-83.

Drake-McLaughlin, Niko and Noelwah R. Netusil. “Valuing walkability and vegetation in Portland Oregon.” *Land Economics* 86.2 (2010): 281-293.

Holmes, Thomas P., Elizabeth A. Murphy, and Kathleen P. Bell. “Exotic forest insects and residential property values.” *Agricultural and Resource Economics Review* 35.1 (2006): 155–166.

Kohlhase, Janet E. “The impact of toxic waste sites on housing values.” *Journal of Urban Economics* 30 (1991): 1-26.

Geoghegan, Jacqueline, Lisa A. Wainger, and Nancy E. Bockstael. “Spatial landscape indices in a hedonic framework: An ecological economic analysis using GIS.” *Ecological Economics* 23 (1997): 251-264.

Germino, Matthew J., William A. Reiners, Benedict J. Blasko, Donald McLeod, and Chris T. Bastian. “Estimating visual properties of Rocky Mountain landscapes in GIS.” *Landscape and Urban Planning* 53 (2001): 71-83.

Irwin, Elena G., and Nancy E. Bockstael. “The problem of identifying land use spillovers: measuring the effects of open space on residential property values.” *American Journal of Agricultural Economics* 83.3 (2001): 698-704.

Joly, Daniel, Thierry Brossard, Jean Cavailhes, Mohamed Hilal, Fancois-Pierre Tourneux, Celine Tritz, and Pierre Wavresky. “A quantitative approach to the visual evaluation of landscape.” *Annals of the Association of American Geographers* 99.2: 292-308.

Lake, Iain R., Andrew A. Lovett, Ian J. Bateman, and Brett Day. “Using GIS and large-scale digital data to implement hedonic pricing studies.” *International Journal of Geographical Information Science* 14.6 (2000): 521-541.

Leggett, Christopher G., and Nancy E. Bockstael. “Evidence of the effects of water quality on residential land prices.” *Journal of Environmental Economics and Management* 39 (2000): 121-144.

Mansfield, Carol, Subhrendu K. Pattanayak, William McDow, Robert McDonald, and Patrick Halpin. “Shades of green: Measuring the value of urban forests in the housing market.” *Journal of Forest Economics* 11.3 (2005): 177-199.

McConnell, Virginia, Elizabeth Kopits, and Margaret Walls. “Farmland preservation and residential density: Can development rights markets affect land use?” *Agricultural and Resource Economics Review* (2005): 133-144.

McLeod, Paul B. “The demand for local amenity: An hedonic price analysis.” *Environment and Planning* A 16 (1984): 389-400.

Mendelsohn, Robert O., Daniel Hellerstein, Michael Huguenin, Robert E. Unsworth, and Richard Brazee. “Measuring hazardous waste damages with panel models.” *Journal of Environmental Economics and Management* 22 (1992): 259-271.

Milon, J. Walter. “Hedonic amenity valuation and functional form specification.” *Land Economics* 60.4 (1984): 378.

Netusil Noelwah R., Sudip Chattopadhyay and Kent F. Kovacs. “Estimating the demand for tree canopy: A second-stage hedonic price analysis.” *Land Economics* 86.2 (2010): 281-293.

Palmquist, Raymond B. “Estimating the demand for the characteristics of housing.” *Review of Economics and Statistics* 66 (1984): 394-404.

Paterson, Robert W. and Kevin J. Boyle. “Out of sight, Out of mind? Using GIS to incorporate visibility in hedonic property value models.” *Land Economics*. 78.3 (2002): 417-425.

Poudyal, Neelam C., Donald G. Hodges, John Fenderson, and Ward Tarkington. “Realizing the economics value of a forested landscape in a viewshed.” *Southern Journal of Applied Forestry* 34:2 (2010): 72-78.

Prasad, V. Krishna, E. Anuradha, and K. V. S. Badarinath. “Climatic controls of vegetation vigor in four contrasting forest types of India—evaluation from National Oceanic and Atmospheric Administration’s advanced very high resolution radiometer datasets.” *International Journal of Biometeorology* (2005): 6-16.

Sander, Heather, Stephen Polasky, and Robert G. Haight. “The value of urban tree cover: A hedonic property price model in Ramsey and Dakota Counties, Minnesota, USA.” *Ecological Economics* 69.8 (2010): 1646-1656.

Ratti, Carlo, and Paul Richens. “Raster analysis of urban form.” *Environment and Planning B: Planning and Design* 31.2 (2004): 297-309.

Rosen, Sherwin. “Hedonic prices and implicit markets: Product differentiation in pure competition.” *The Journal of Political Economy* 82.1 (1974): 34-55.

Tyrväinen, Lisa, and Antti Miettinen. “Property prices and urban forest amenities.” *Journal of Environmental Economics and Management* 39 (2000): 205–223.

Table 1: Proximity Studies

Author(s)	Time Period	Location	Landscape Type	Variables for View, Proximity, or Both	Proximity Vegetation Variable Form	View Vegetation Variable Form	Terrain Data Type and Resolution	Most Relevant Results	Water's Impact	Possible Limitations
Drake-McLaughlin and Netusil (2010)	2005-2007	Portland, Oregon	Mixed	Proximity	Percentages	N/A	LiDAR, 3ft^2	Tall and short vegetation have positive impacts. Advanced terrain data yields more robust results than in Netusil's previous study, which only considered tree canopy.	Positive and significant at 5% level	-View is not considered.
Geoghegan et al. (2007)	1990	Patuxent Watershed, Maryland	Mostly Rural	Proximity	Percentages	N/A	Landcover	Landscape diversity has a positive effect on property values.	Positive and significant at 1% level	-Landcover data is used. -Only data from one year is included.
Holmes et al. (2006)	1992-2002	Sparta, New Jersey	Rural	Proximity	Percentages	N/A	30 meters^2	Vegetation species composition matters, hardwood trees tend to increase property values more than pine.	Lakes and ponds are positive and significant at 5% level	-Distances to relevant points of interest and commute times are not included.
Mansfield et al. (2005)	1996-1998	RTP, NC	Mixed	Proximity	NDVI "Greenness" Index	N/A	30 meters^2	Separates the impacts of on-property vegetation from those of close forests.	Excluded	-View is not considered. -Limited resolution data is used to estimate on- parcel vegetation.
Netusil (2009)	1999-2001	Portland, Oregon	Mixed	Proximity	Tree canopy percentages	N/A	Canopy	Tree canopy on property or within 1/4 mile has a positive effect up to 40% coverage. Above 40% coverage additional canopy is detrimental	Positive and significant at 1% level	-View is not considered.
Sandler et al. (2010)	2005	Dakota and Ramsey Counties, Minnesota	Mixed	Proximity	Percentages	N/A	Tree Canopy, 30 meters^2	44% canopy within 100 meters is ideal. Total area in view has a positive and significant impact at the 1% level. Vegetation beyond 250 meters is not impactful.	Positive and significant at the 1% level	-100 meters is smallest radius for categorizing vegetation. -Only canopy is considered.

Table 2: View or View and Proximity Studies										
Author(s)	**Time Period**	**Location**	**Landscape Type**	**Variables for View, Proximity, or Both**	**Proximity Vegetation Variable Form**	**View Vegetation Variable Form**	**Terrain Data Type and Resolution**	**Most Relevant Results**	**Water's Impact**	**Possible Limitations**
Germino et al. (2000)	1997	Wyoming	Rural	View	N/A	Area in View	Landcover, 30 meters^2	Simulations can yield accurate estimates of view composition.	N/A	-Low resolution data -Vegetation and buildings cannot block views
Joly et al. (2009)	Not Specified	Dijion, France	Suburban	Both	N/A	Area in View	Landcover, 25 meters^2	View variables show positive and significant impacts on property values and they appear to be different from proximity measures.	Negative and Insignificant	-Low resolution data -Estimated vegetation height is inaccurate
Paterson and Boyle (2002)	1997-1999	Connecticut	Rural	Both	Percentages	Percentage of Total Area in View	Landcover, Unreported Resolution	Visibility of vegetation can reduce spatial autocorrelation.	Negative and Significant	-Landcover data is used -View and proximity variables are not statistically compared
Poudyal et al. (2010)	2004-2008	Outside Nashville, Tennessee	Rural	View	N/A	Area of Forest in View	30 meters^2	Total area of forests in view can have a positive and significant effect at the 1% level.	Excluded	-Low view simulation resolution -Vegetation cannot obscure views -Small sample of 546 observations
Tyrvainen and Miettinen (2000)	1984-1986	Salo, FInland	Mixed	Both	Percentages	Forest in View (Binary)	Landcover, minimum patch area 3000 m^2	Suggests view can be important. Apartments with forest views tend to sell for 4.9% more.	Excluded	-Landcover data is used -Low view simulation resolution, -Small sample of 590 observations -Only terraced apartments are included

Table 3: Property Information Descriptive Statistics

Variable Name	Mean	Standard Deviation
Sale Price (Nominal $)	279,592	229,398
Lot Size (Acres)	0.356	0.329
Heated Area (Square Feet)	2,303	1,061
Age (Days)	37,601	2,303
One or more Fireplaces (Binary)	0.79	0.41
Assessed Quality	406,599	311,046
Ranch Style Home (Non split level, conventional or townhome)	0.036	0.19
One Story Home	0.48	0.50
Two Story Home	0.32	0.47
Two Bathrooms	0.26	0.44
Split Level Style Home (Non ranch, conventional or townhome)	0.070	0.25
Spring (Binary)	0.26	0.44
Winter (Binary)	0.20	0.40
Fall (Binary)	0.24	0.42
Electric (Binary)	0.98	0.11
Gas (Binary)	0.95	0.21
Basement Percentage	10.1	20.2
Commute Time (Minutes)	2,793	1,572
Distance to the Research Triangle Park	68,312	9,068
Distance to Raleigh's Central Business District	21,940	9,566
Trend (Inflation Term)	103.2	6.31
Home Density (Number of homes within 1 Mile)	3427	1245

Table 4: Proximity Variable Descriptive Statistics

Variable Name	Description	Mean	Standard Deviation
B_close_short	Vegetation shorter than 15 feet tall and within 50 feet of the home	481	644
B_med_short	Vegetation shorter than 15 feet tall, between 50 feet and 100 feet from the home	749	339
B_far_short	Vegetation shorter than 15 feet tall, farther than 100 feet from the home	116	135
B_close_med	Vegetation between 15 and 30 feet tall and within 50 feet of the home	277	145
B_med_med	Vegetation between 15 and 30 feet tall, between 50 feet and 100 feet from the home	728	328
B_far_med	Vegetation between 15 and 30 feet tall, farther than 100 feet from the home	4,457	1,465
B_tall_0_10	Vegetation taller than 50 feet and closer than 10 feet	3.69	7.36
B_tall_10_25	Vegetation taller than 50 feet and between 10 and 25 feet	23.7	35.1
B_tall_25_50	Vegetation taller than 50 feet and between 25 and 50 feet	105	112
B_tall_50_100	Vegetation taller than 50 feet and between 50 and 100 feet	472	403
B_tall_100_200	Vegetation taller than 50 feet and between 100 and 200 feet	1,728	1,268

-Proximity parameter estimates are interpreted as the percentage change in property value by one additional 3x3 cell of that variable in proximity.

Table 5: View Variable Descriptive Statistics

Variable Name	Description	Mean	Standard Deviation
V_close_short	Vegetation in view shorter than 15 feet tall and within 50 feet of the home	53.3	41.2
V_med_short	Vegetation in view shorter than 15 feet tall, between 50 feet and 100 feet from the home	57.0	59.7
V_far_short	Vegetation in view shorter than 15 feet tall, farther than 100 feet from the home	327	612
V_close_med	Vegetation in view between 15 and 30 feet tall and within 50 feet of the home	101	107
V_med_med	Vegetation in view between 15 and 30 feet tall, between 50 feet and 100 feet from the home	89.7	83.8
V_far_med	Vegetation in view between 15 and 30 feet tall, farther than 100 feet from the home	317	583
V_tall_0_10	Vegetation taller than 50 feet, in view and closer than 10 feet	2.49	4.86
V_tall_10_25	Vegetation taller than 50 feet, in view and between 10 and 25 feet	8.77	12.2
V_tall_25_50	Vegetation taller than 50 feet, in view and between 25 and 50 feet	28.1	29.8
V_tall_50_100	Vegetation taller than 50 feet, in view and between 50 and 100 feet	72.0	63.6
Streets	Streets (Square Feet)	21.2	56.5
Buildings	Buildings (Square Feet)	359	676
Hydro	Hydrology (Square Feet)	9.73	187

-All view parameter estimates are interpreted as the percentage change in property value by one additional 3x3 cell of that variable in view.

Table 6: Regression Output

	Regression Name		
Variable	**Base**	**Proximity**	**View**
Lot Size	0.025**	0.020**	0.022**
	(0.0106)	(0.0107)	(0.0107)
Heated Area	0.516***	0.512***	0.509***
	(0.0128)	(0.0129)	(0.0129)
Age	-0.0060***	-0.0062***	-0.0062***
	(2.59e-4)	(2.68e-4)	(2.62e-4)
One or more Fireplaces	0.070***	0.069***	0.068***
	(0.0077)	(0.0077)	(0.0077)
Assessed Quality	6.59e-7***	6.57e-7***	6.58e-7***
	(1.65e-8)	(1.66e-8)	(1.66e-8)
Ranch Style Home	0.021	0.019	0.020
	(0.016)	(0.016)	(0.016)
One Story Home****	-0.036***	-0.036***	-0.036***
	(0.0096)	(0.0096)	(0.0096)
Two Story Home****	-0.049***	-0.049***	-0.049***
	(0.0085)	(0.0085)	(0.0085)
Two Bathrooms	-0.025***	-0.026***	-0.026***
	(0.0076)	(0.0076)	(0.0076)
Split Level Style Home	-0.077***	-0.079***	-0.079***
	(0.014)	(0.014)	(0.014)
Spring	-0.030***	-0.030***	-0.030***
	(0.0077)	(0.0077)	(0.0077)
Winter	-0.064***	-0.064***	-0.0640***
	(0.0083)	(0.0083)	(0.0083)
Fall	-0.045***	-0.044***	-0.044***
	(0.0079)	(0.0079)	(0.0079)
Electric	0.174	0.171	0.172
	(0.043)	(0.043)	(0.043)
Gas	-0.023***	-0.021***	-0.024***
	(0.017)	(0.017)	(0.017)
Basement_Percentage	4.62e-4***	4.27e-4***	4.36e-4***
	(1.77 e-4)	(1.77e-4)	(1.77 e-4)

Table 6 (Continued)

Commute	1.26e-6	1.52e-6	7.36e-7
	(8.36e-6)	(8.41e-6)	(8.41e-6)
Commute^2	-3.57e-9***	-3.54e-9***	-3.65e-9***
	(1.055e-9)	(1.06e-9)	(1.06e-9)
DistanceRTP	-1.49e-5***	-1.47e-5***	-1.48e-5***
	(7.80e-7)	(7.84e-7)	(7.84e-7)
DistanceRaleigh	-1.13e-5***	-1.16e-5***	-1.17e-5***
	(6.47e-7)	(6.47e-7)	(6.51e-7)
Trend	0.0061***	0.0063***	0.0063***
	(0.0026)	(0.0027)	(0.0026)
Home Density	4.965e-5***	4.71e-5***	4.81e-5***
	(3.49 e-6)	(3.88 e-6)	(3.83e-6)
B_close_short		-1.68e-5***	
		(6.18e-6)	
B_med_short		-2.21e-5*	
		(1.34e-5)	
B_far_short		7.89e-7***	
		(3.84e-7)	
B_close_med		-1.33e-5	
		(2.91e-5)	
B_med_med		-9.92e-6	
		(1.59e-5)	
B_far_med		5.147e-7	
		(3.865e-7)	
B_tall_0_10		0.00032	
		(0.0008)	
B_tall_10_25		-0.00012	
		(0.00023)	
B_tall_25_50		1.33e-5	
		(6.41e-5)	
B_tall_50_100		-1.38e-5	
		(1.8e-5)	
B_tall_100_200		1.1e-5***	
		(4.07e-6)	
Streets		2.47e-6	1.93e-6
		(7.28e-5)	(7.35 e-5)

Table 6 (Continued)

Buildings		5.39e-6	6.53e-7
		(1.39e-5)	(1.38e-5)
Hydro		3.0e-5**	2.95e-5**
		(1.53e-5)	(1.53 e-5)
V_close_short			-0.000304***
			(8.97e-5)
V_med_short			-1.9e-5
			(6.30e-5)
V_far_short			-2.71 e-5**
			(1.06e-5)
V_close_med			-0.0001327**
			(6.30e-5)
V_med_med			1.69e-5
			(4.21e-5)
V_far_med			1.31e-5
			(1.05e-5)
V_tall_0_10			-0.000670
			(0.0008)
V_tall_10_25			-1.1e-6
			(0.000332)
V_tall_25_50			0.000231*
			(0.0001)
V_tall_50_100			0.000134**
			(5.59e-5)
V_tall_100_200			5.15 e-5**
			(2.58 e-5)
Adjusted R^2	0.6450	0.6454	0.6456

****-Excluded are 2.5, 3 and 3.5 story homes.
-Parameter Values and their (non-robust standard errors) are reported.
-View and proximity parameter estimates are interpreted as the percentage change in property value by one additional 3x3 cell of that variable either in view or proximity.
-All regressions include year fixed effects.
-*, ** and *** signify significance at the 10%, 5% and 1% levels.

Table 7: Proximity and View Comparison Regression Output

Variable	Difference Regression
Lot Size	0.019**
	(0.011)
Heated Area	0.51***
	(0.013)
Age	-0.0062***
	(2.65e-4)
One or more Fireplaces	0.068***
	(0.0077)
Assessed Quality	6.59e-7***
	(1.66e-8)
Ranch Style Home	0.012
	(0.016)
One Story Home	-0.036***
	(0.0096)
Two Story Home	-0.049***
	(0.0085)
Two Bathrooms	-0.026***
	(0.0076)
Split Level Style Home	-0.079***
	(0.014)
Spring	-0.030***
	(0.0077)
Winter	-0.065***
	(0.0083)
Fall	-0.044***
	(0.0079)
Electric	0.18
	(0.043)
Gas	-0.027***
	(0.017)
Basement_Percentage	4.38e-4***
	(1.76e-4)
Commute	-1.77e-6
	(8.41e-6)
$Commute^2$	-3.48e-9***
	(1.06e-9)

Table 7 (Continued)

DistanceRTP	-1.47e-5***
	(7.83e-7)
DistanceRaleigh	-1.20e-5***
	(6.57e-7)
Trend	0.0063***
	(0.0026)
Home Density	5.09e-5***
	(3.85e-6)
Theta_close_short	1.98e-5
	(4.97e-5)
Theta _med_short	-3.65e-5*
	(2.04e-5)
Theta _far_short	6.92e-7
	(7.67e-7)
Theta _close_med	-1.64e-5
	(4.41e-5)
Theta _med_med	4.83e-6
	(1.97e-5)
Theta _far_med	-1.31e-6**
	(5.69e-7)
Theta _tall_0_10	7.29e-4
	(0.0042)
Theta _tall_10_25	4.12e-4
	(7.44e-4)
Theta _tall_25_50	7.09e-5
	(2.61e-4)
Theta _tall_50_100	1.47e-4
	(9.18e-5)
Theta _tall_100_200	-7.84e-6
	(3.09e-5)
Streets	-1.09e-4
	(9.31e-5)
Buildings	7.93e-6
	(1.41e-5)
Hydro	3.09e-5**
	(1.53e-5)
Adjusted R^2	0.6459

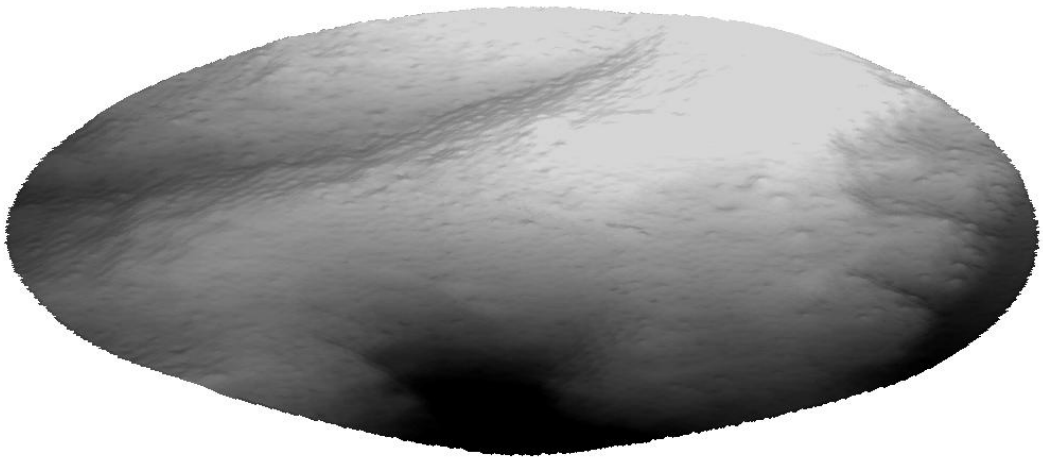
Table 7 (Continued)

-Parameter Values and their (SEs) are reported.
-View and proximity parameter estimates are interpreted as the percentage change in property value by one additional 3x3 cell of that variable either in view or proximity.
-All regressions include year fixed effects.
-*, ** and *** signify significance at the 10%, 5% and 1% levels.

Table 8: Proximity and View Correlations			
	Distance Range		
	0-50 Feet	50-100 Feet	100-660 Feet
Average Correlations Between View and Proximity Variables	0.188	0.011	0.00363

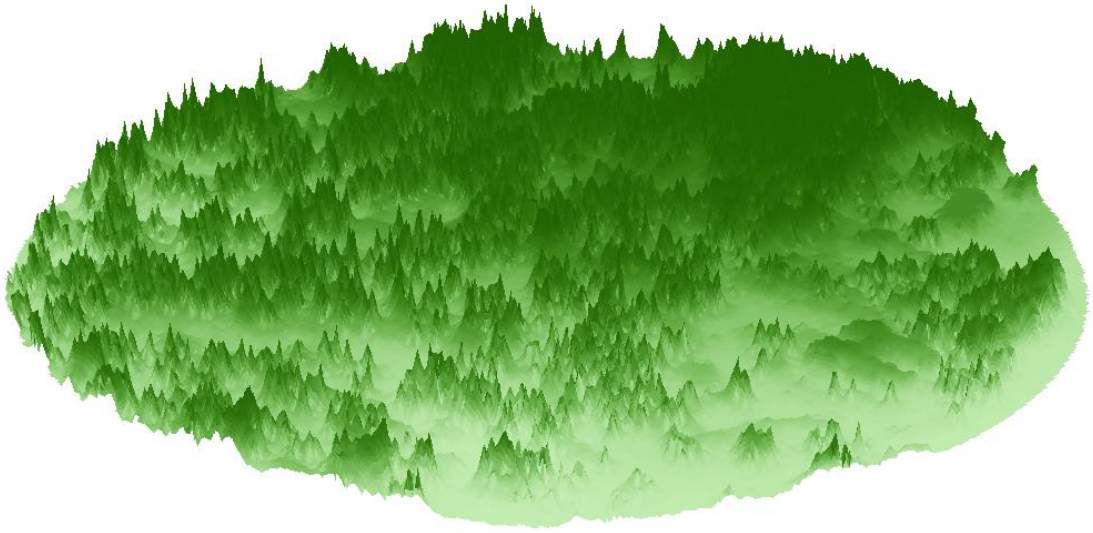
APPENDIX

Figure 1: DEM



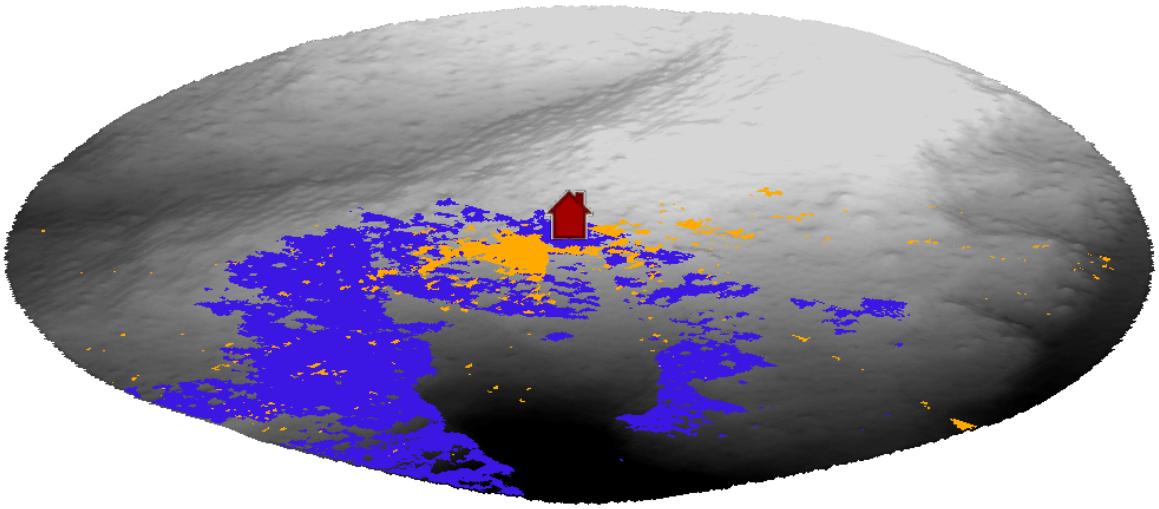
Above is the Digital Elevation Model (DEM) for the terrain within 1/8 of a mile around one study home. The elevation changes about 100 feet. The lowest elevation is displayed as black while the lightest grey is the highest. The DEM maps, as closely as possible, the bare earth surface. Commonly views are simulated on DEMs to estimate areas in view, which can work well when only the bare earth surface can block views.

Figure 2: DSM



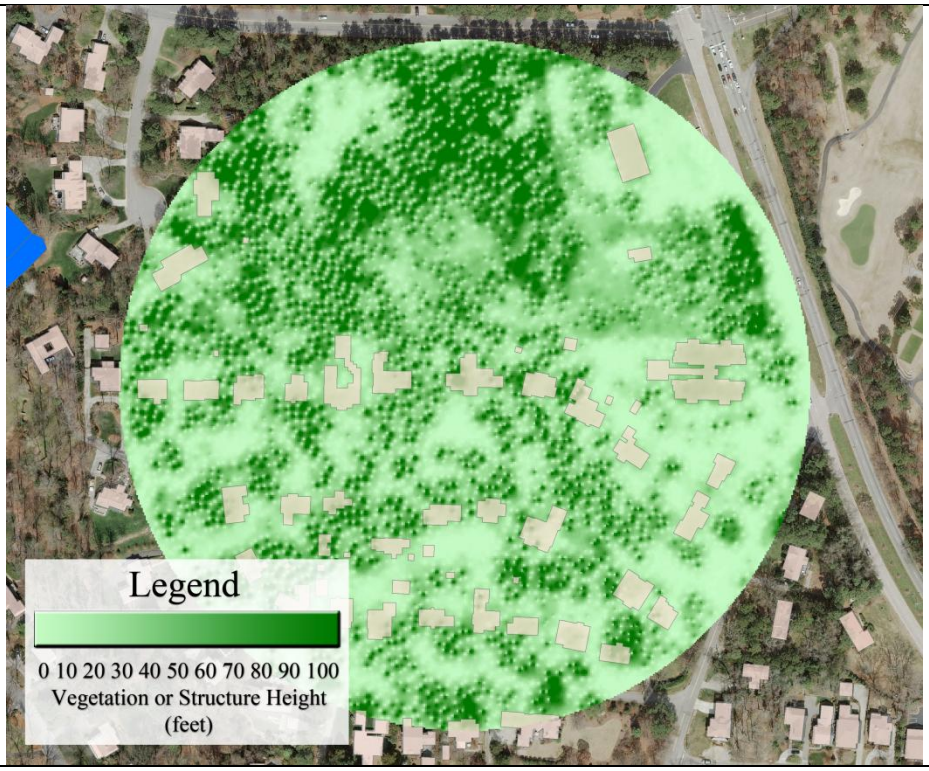
Above is the Digital Surface Model (DSM) for the same region, from the same perspective. DSMs map the bare earth surface as DEMs, but they also include vegetation and structures. Using these more complex models of the features around homes, more realistic estimates of area in view can be found. Most of the features visible above are trees.

Figure 3: DEM and DSM Viewsheds



Viewsheds are the estimated total areas in view from GIS simulations. Above are the two types of simulated viewsheds, both simulated from the home. The purple viewshed is from using the Digital Elevation Model (DEM), and the orange overlaid is derived from the Digital Surface Model (DSM). As a result of the DEM view only being blocked by the surface of the earth, the estimated viewshed is much larger than for the DSM. The DSM viewshed shows most area around the home is in view and a few patches of far tall trees as well. The DEM viewshed estimates that a large area is visible to the southeast. View of this area is blocked by trees and therefore is modeled more correctly by the DSM viewshed. Overall, DEM viewsheds overestimate area in view.

Figure 4: Height Distance Raster



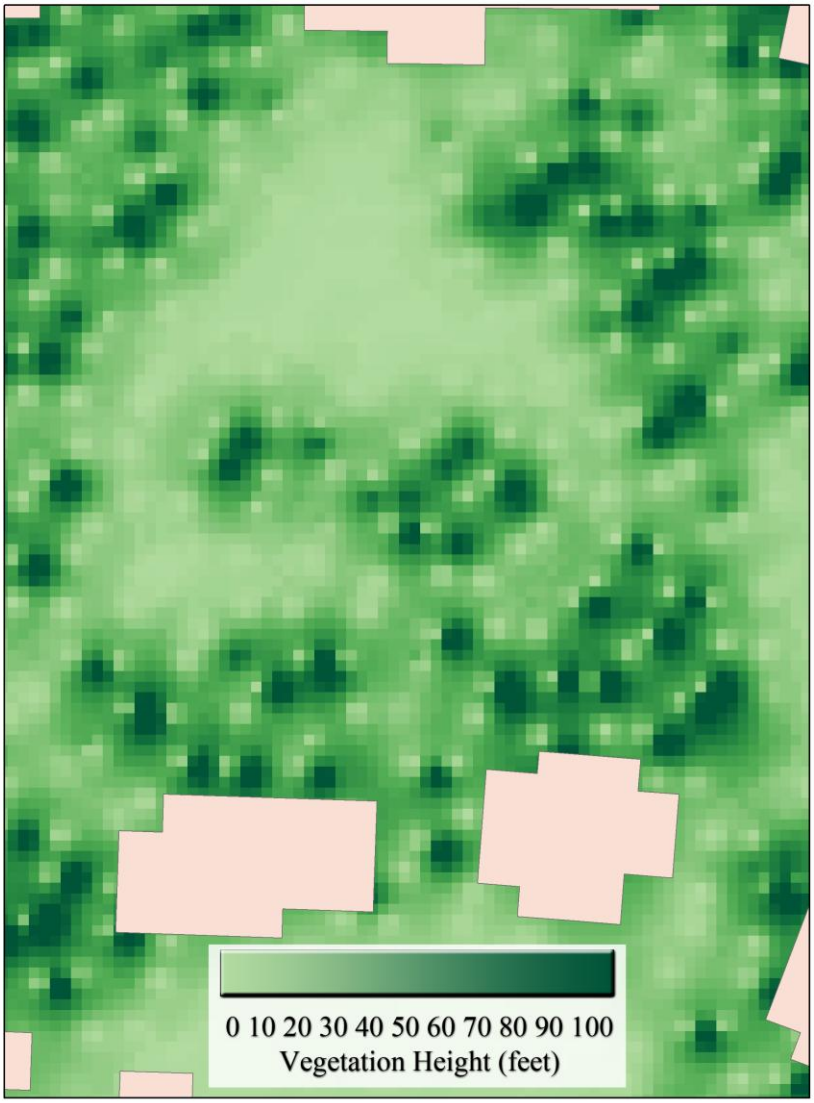
In the image to the left the height of trees is derived for every cell from the height of the DSM above the DEM. The values range to 0 (light green) to 100 feet (dark green).

Figure 5: Overlaid DSM Viewshed

Displaying the viewshed derived from the DSM over an orthographic photo helps illustrate what the view simulation captures. The view to the left is the viewshed from the home at the top of the image. Orange transparent areas are estimated by the simulation as being in view. The lawn south of the home is in view and most of the view farther south is blocked by trees along the road. To the east of the home most of the view is blocked by the close tree line.

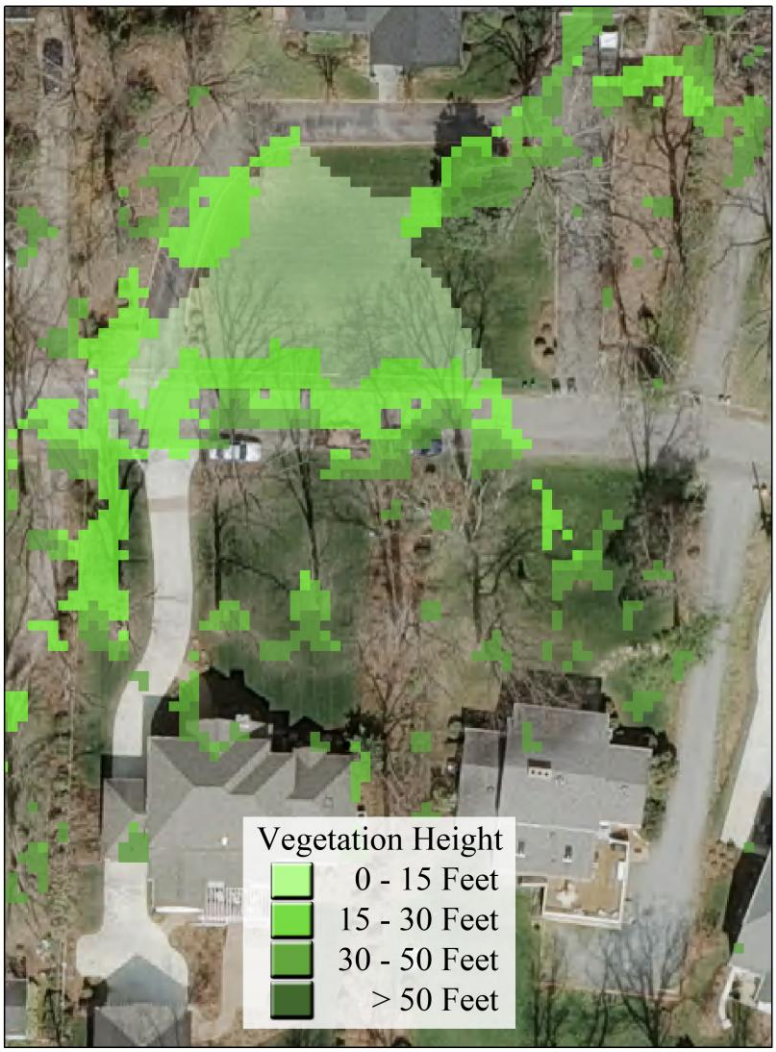
Also displayed here are the streets and buildings layers. Streets are shown as partially transparent grey and buildings as beige polygons. These layers are used to estimate view composition of non-vegetation features.

Figure 6: Vegetation Height



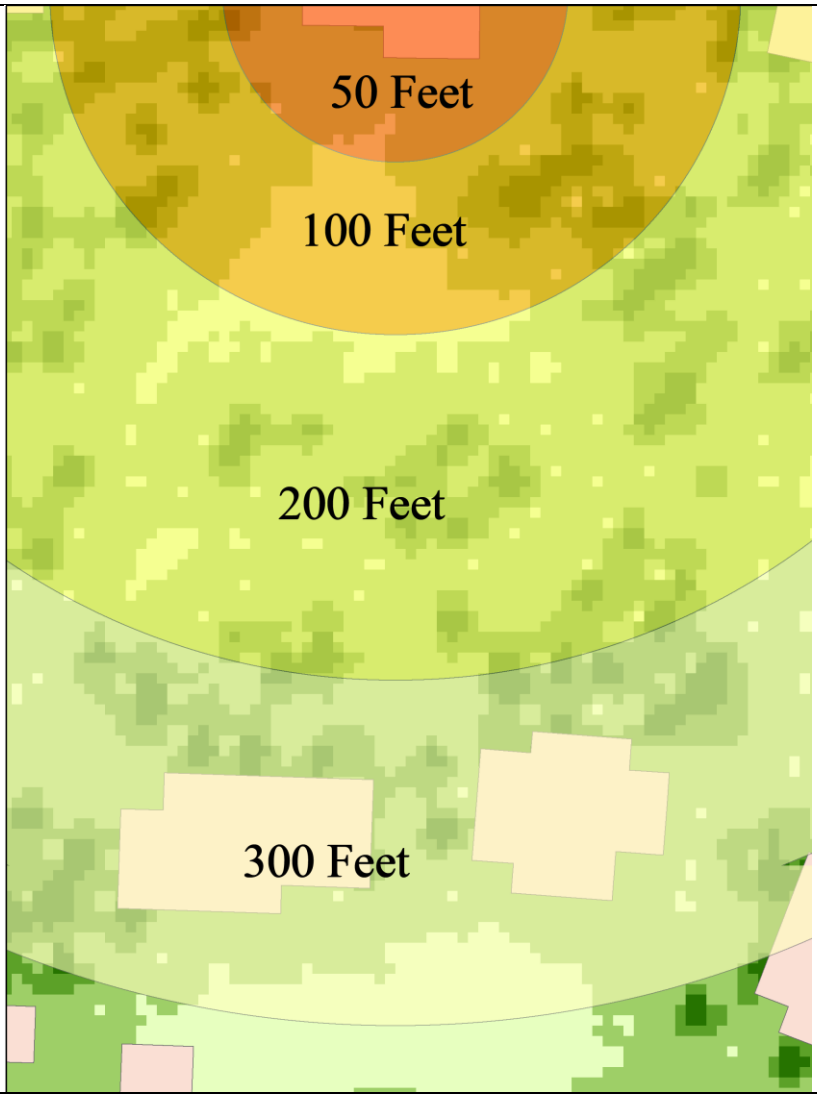
To the left is the height difference raster for the area shown in Figure 4 above, also building footprints are displayed. For each 3x3 cell the elevation of the DSM above the DEM is shown. Where the difference is larger the color is darker green, indicating taller vegetation.

Figure 7: Vegetation Classification by Height



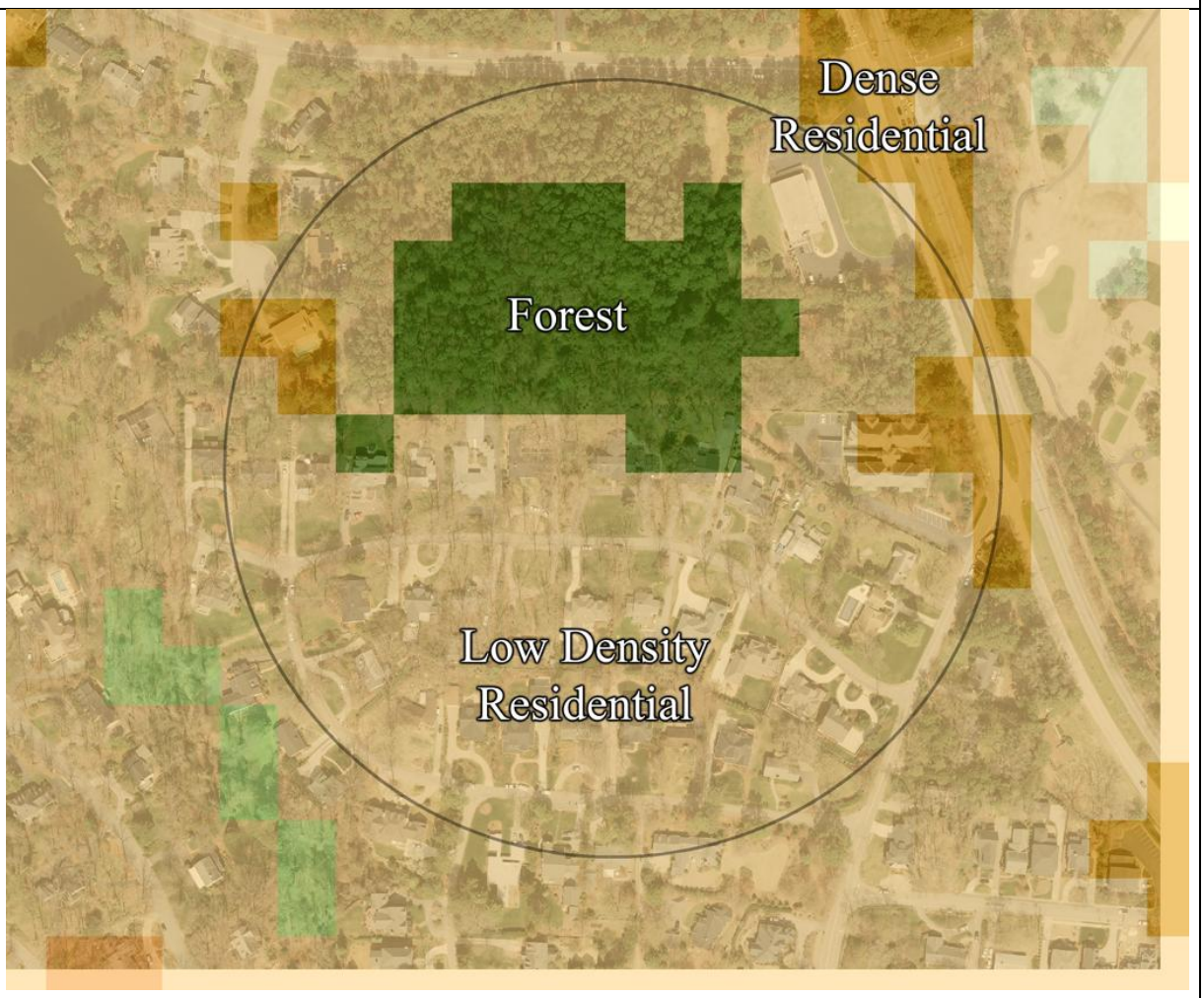
Using the height of the DSM layer above the DEM for each 3x3 cell, each cell in view is categorized based on the height of the vegetation in it. In the image to the left, the total area in view from Figure 5 is categorized based on vegetation height. Light green cells indicate vegetation lower than 2 feet. The darkest green cells signify trees taller than 30 feet.

Figure 8: Vegetation Classification Distances



To the left are the radii used for classifying vegetation into distance ranges around each home. Centered on the study home at the top of the image are four buffers corresponding to 50, 100, 200 and 300 feet. Vegetation is considered up to 660 feet, or 1/8 of a mile from each home.

Figure 9: Land Classification or "Landcover" Data



Above is an example of landcover data for the area surrounding one home in the data set.
The resolution for landcover data used by previous studies, and above, is 30 meters^2. Green blocks indicate forest while orange signifies residental areas. Previous view studies use similar landclass data to estimate areas in view at very low resolutions. To minimize measurement error, this thesis uses LiDAR terrain data at a 3ft^2 resolution. In addition to estimating areas in view, LiDAR data is used to classify vegetation based on height instead of estimating total forest area in view.The radius represents the local extent considered around the home, which is 1/8 mile.

Figure 10: National LiDAR Coverage (USGS)

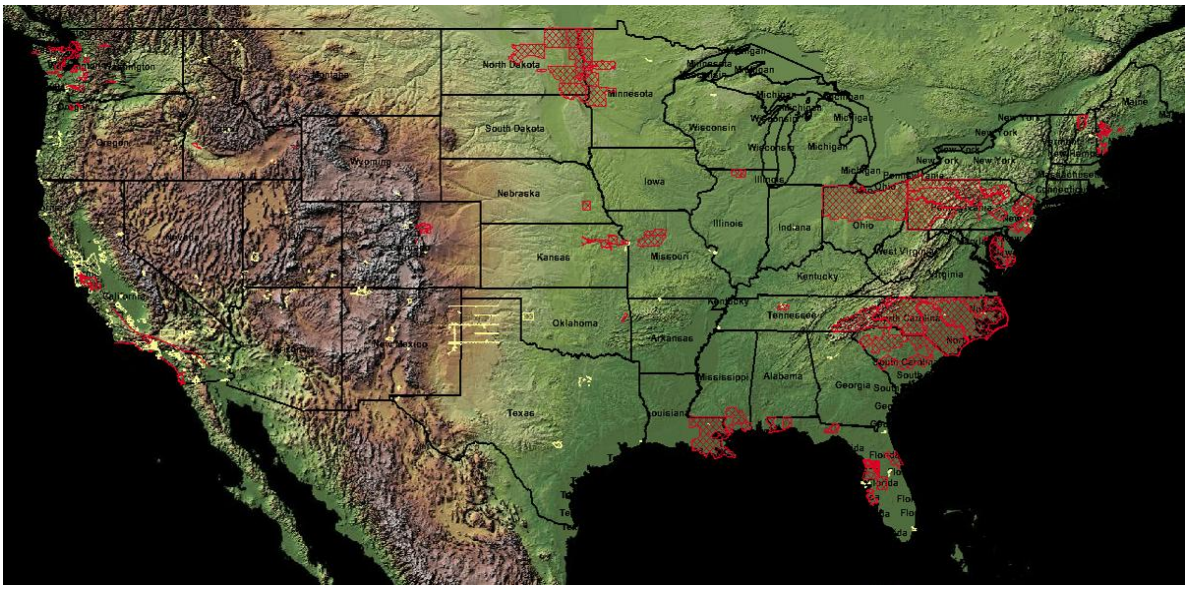
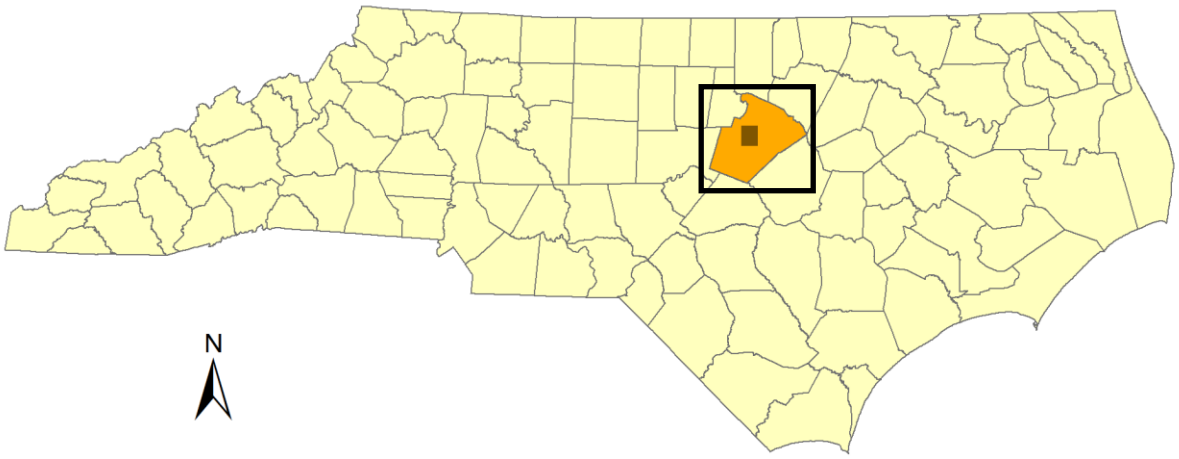
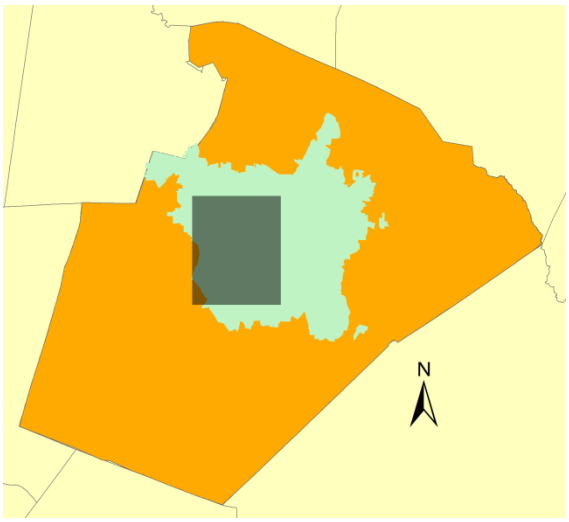


Figure 11: Study Region Extent



The grey rectangle in both images is the study region. The light green shape is the outline of the City of Raleigh.



Vector to Raster Conversion	
Figure 12: Vector Image 	Shown to the left is the orthographic photo of a neighborhood overlaid building, street and water layers from the Wake County GIS website. These files are in vector format.
Figure 13: Raster Image 	To expedite simulations the original vector files are converted to raster. This process creates cells at a resolution of 3 feet. By comparing the before and after images, the conversion process does not significantly change the data.

Workstation Specifications

Operating System:
Windows 7 (64-bit)

Software:
ESRI ArcGIS 10
Eclipse 3.7.1
Python 2.6
PyDev 2.2.4

Processor:
AMD Phenom X6 1090T 3.2 GHz

Motherboard:
MSI MS-7623

RAM:
8 Gb Crucial Dual Channel DDR3 1333

Hard Drives:
1x 60 Gb Crucial SATA III SSD
1x 40 Gb Mushkin SATA II SSD
3x 500 Gb Western Digital 7200 rpm