

Land Cover Change Modeling Using Cellular Automata Rules Derived from Landsat Imagery

A Thesis

Presented to
the faculty of the School of Engineering and Applied Science
University of Virginia

in partial fulfillment
of the requirements for the degree

Master of Science

by

S. M. Vacik

August

2015

APPROVAL SHEET

The thesis
is submitted in partial fulfillment of the requirements
for the degree of
Master of Science

S. M. Vacik

AUTHOR

The thesis has been read and approved by the examining committee:

Gerard P. Learmonth, Ph.D.

Advisor

Barry Horowitz, Ph.D.

John Porter, Ph.D.

Accepted for the School of Engineering and Applied Science:



Craig H. Benson, Dean, School of Engineering and Applied Science

August
2015

Land Cover Change Modeling Using Cellular Automata
Rules Derived from Landsat Imagery

Thesis

Presented in Partial Fulfillment of the Requirements for the Degree Master of
Science in Systems Engineering in the Graduate School of
The University of Virginia

By

S. M. Vacik, B.S.

Graduate Program in Systems and Information Engineering

The University of Virginia

2015

Thesis Committee:

Gerard P. Learmonth, Ph.D., Advisor

John Porter, Ph.D., Committee Member

Barry Horowitz, Ph.D., Committee Chair

© 2015, S. M. Vacik

Abstract

Land cover generally describes categorical feature classes defined by their physical characteristics, such as vegetation or material type, as seen on a small parcel of surface area of the Earth. Land cover changes occur annually due to such activities as urban development, agriculture, climatic shifts, and natural disasters such as fires and hurricanes. Observing land cover change provides insight into trends due to natural and manmade annual changes that can be used to construct decision rules and to model techniques purposed with preventing or mitigating the effects of extreme weather or climatic shifts.

Annual datasets are required for such analysis and the United States Geological Survey (USGS) provides land cover datasets covering the entire United States for only the years 1992, 2001, 2006, and 2011, which cover a time frame of 19 years. One mapping technique called Variable Multiple Endmember Mixture Analysis (VMESMA) accurately maps physical characteristics of land but can be computationally intensive and slow depending on its implementation. In this thesis, Variable Spectral Unmixing (VSU) — a new and improved spectral mixture analysis technique inspired by VMESMA — is presented to produce land cover estimates from preprocessed Landsat imagery for the years 2001 through 2011. VSU results correspond to physical surface material types, such as coniferous trees and artificial substances, and are interpreted into land cover classes based on material type prior to overall classification by hierarchal rules. Agreements with the USGS National Land Cover Dataset (NLCD) of less than 40% result due to the classification rules and reflect the physical surface types that meet the first rule within the hierarchy. Future land cover mapping applications require new classification rules to improve interpretation of the VSU results and the agreement of the generated maps with the USGS

NLCD.

Land cover estimates are used to develop Cellular Automaton-based (CA) decision rules to map land cover change and to forecast such changes into future years. The CA rules are based on the analysis of Moore and von Neumann neighborhoods of a time series of VSU-generated maps. Results of the neighborhood analysis revealed potential general neighborhood structures for decision rules, which may or may not vary in time as a result to changes in the rates of change of each class. Forecast results are tested in a basic iterative fashion using the USGS NLCD 2001 map as a base case. Agreements of 66% and 77% of the von Neumann and Moore forecasts, respectively, for the year 2011 with the NLCD 2011 demonstrate the feasibility of land cover change modeling using neighborhood-based CA decision rules and a method for modeling land cover change trends based on decision rules derived from a time series of maps.

*To Sher, Jiri, Pat, & Dave -
Thank you for always believing in me.*

Acknowledgments

I would like to express my deep gratitude to my adviser Dr. Learmonth for his continued patience and encouragement and for his time and expertise in modeling. The path to get to this point traveled many circuitous parallels and ultimately resulted in this thesis, but it also culminated in a milestone in my life that taught me to have faith in myself. I hope the techniques presented in this work will benefit your research in years to come.

I wish to thank the other members of my thesis committee, Dr. Horowitz and Dr. Porter, for providing constructive insights and their expertise in remote sensing systems or image products. In particular, I also would like to express my gratitude to Dr. Porter for meeting with myself and my adviser to provide advice on the mapping procedure used in this thesis.

I would like to thank Michael Purvis for his time, support and advice for improvements to my code in Python.

Additionally, I would like to thank the other graduate students who shared the same office with me: Debo, Heimir, Mike, and Osama. Thank you for enduring my mathematical proofs and derivations and for your assistance in clarifying some points of Python coding.

Lastly, I would like to thank my family and friends. Thank you for inspiring and encouraging me to pursue a career in engineering and science. These fields continue to enrich my life.

Table of Contents

1. Introduction	10
2. Approximating the National Land Cover Dataset Using Variable Spectral Unmixing (VSU)	13
2.1 Section Overview	13
2.2 Methodology	18
2.2.1 Data Sources	18
2.2.1.1 Landsat	18
2.2.1.2 Spectral Libraries	20
2.2.1.3 National Land Cover Dataset	22
2.2.2 Image Selection	23
2.2.3 Preprocessing	24
2.2.4 Variable Spectral Unmixing (VSU)	28
2.2.4.1 Algorithm Overview	28
2.2.4.2 Extraction of Water	28
2.2.4.3 Spectral Matching	29
2.2.4.4 Spectral Mixture Analysis	31
2.2.4.5 Classification by Endmembers & Rules	31
2.2.5 Error Analysis	36
2.3 Results & Discussion	37
3. Land Cover Change Modeling (LCCM)	50
3.1 Section Overview	50
3.2 Methodology	51
3.3 Results & Discussion	55
3.3.1 Rationale of Transition Exclusions	55
3.3.2 Results of Neighborhood Analysis	56
3.3.3 Modeling Results	62
4. Conclusion & Future Work	70
4.1 Mapping Discussion	70
4.2 LCCM Conclusions	73
A. Bibliography	78

List of Tables

2.1 Landsat 5 & 7 Sensor Bandwidths	19
2.2 ESUN Solar Spectral Irradiances	26
2.3 Slope & Intercept Values	26
2.4 LS5 TM Relative Spectral Response Points	27
2.5 Parameters, Equations, & Thresholds for Spectral Matching	30
2.6 USGS NLCD Classification Classes, Codes, Criteria, & Corresponding Materials	32
2.7 Agreements with the NLCD	47
3.1 Von Neumann Neighborhoods with Greatest Frequency of Occurrence per Transition	59
3.2 Moore Neighborhoods with Greatest Frequency of Occurrence per Transition	60
3.3 Comparison of von Neumann Forecasts with NLCD	65
3.4 Comparison of Moore Forecasts with NLCD	69

List of Figures

2.1 NLCD 2011 Legend [11]	23
2.2 Code snippet showing classification rules used with VSU	35
2.3 2001 Mapping results versus NLCD 2001 by locale	41
2.4 2006 Mapping results versus NLCD 2006 by locale	42
2.5 2011 Mapping results versus NLCD 2011 by locale	43
3.1 2011 Von Neumann Forecasting Results versus NLCD 2011	64
3.2 Moore 2011 Forecasting versus NLCD 2011	67

Acronym Definitions

Acronym	Definition
C-CAP	Coastal Change Analysis Program
CA	Cellular Automata
DN(s)	Digital Number(s)
ETM+	Enhanced Thematic Mapper Plus
LCCM	Land Cover Change Modeling
LCM	Land Cover Mapping
LS	Landsat
LSI	Landsat Imagery
MESMA	Multiple Endmember Spectral Mixture Analysis
MIR	Middle Infrared
NASA	National Aeronautics & Space Administration
NIR	Near Infrared
NOAA	National Oceanic and Atmospheric Administration
OLI	Overhead Land Imager
SA	Spectral Angle
SMA	Spectral Mixture Analysis
SWIR	Short-Wave Infrared
TM	Thematic Mapper
USGS	United States Geological Survey
VMESMA	Variable Multiple Endmember Spectral Mixture Analysis
VSU	Variable Spectral Unmixing

1. Introduction

Earth continues to undergo climatic shifts in response to aggravated environmental conditions, such as increasing carbon dioxide concentrations in the atmosphere. These changes to the land already disrupt human land use, as in the case of the historic drought affecting agricultural activities in California and the unprecedented heat waves cycling around the globe, causing death in regions without access to the modern convenience of air conditioning. In the case of Louisiana, used here without loss of generality, land loss threatens to displace a large population in addition to the loss of lands used in the production of produce, petroleum products, etc. due to steady land subsidence, sea level rise, and damaging weather events. This thesis focuses on the application of land cover mapping and change modeling techniques to the Louisiana Gulf Coast, but they may be applied as is or adapted for application to the land cover of other locations and for applications other than land cover.

According to *Louisiana's Comprehensive Master Plan for a Sustainable Coast* (LCMPSC), Louisiana lost 1,880 square miles of land in the past 80 years and risks losing an additional 1,750 square miles of land by 2065 [15]. The state of Louisiana hosts a system of levees and storm drains to assist in diverting flood waters along the Mississippi and other major waterways from residential and other major infrastructure, but these flood-diversion systems also prevent naturally occurring sediment build up and fresh water from reaching wetlands suffering erosion and other damage leading to land loss [15]. Rising sea levels

further complicate this problem. According to the National Oceanic and Atmospheric Administration's (NOAA) Tide and Currents, sea level currently rises by 9.03 millimeters per year on Grand Isle, Louisiana, which equates to a rate of approximately 3 feet per 100 years [16]. If this rate continues steadily, the mean seal level could increase by 1.5 feet on Grand Isle, Louisiana, by the year 2065, placing residents and their way of life at risk.

Though these events are in the future and are inherently theoretical, current trends appear to indicate a significant risk to the livelihoods of not only the people of Louisiana. Similar problems arise around the globe. Preventive measures underway now would logically be the best course of action, but as in the case of any investment, such preventive measures require estimation of their effectiveness as proof of their ability to circumvent a risk prior to receiving any support and, more importantly, actual funding. Modeling and simulation can provide a means for demonstrating such methods.

Land cover change modeling using a cellular automata (CA) approach offers the opportunity to study Louisiana's land loss by forecasting changes in land cover area and water water area based on current trends. It will also enable testing the effectiveness of preventive techniques affecting land cover change, such as wetlands restoration programs. Prior to developing a CA-based land cover model, basic cellular automata land cover change rules must be derived, especially those related to wetlands land cover.

Such analysis requires a time series of land cover maps to determine annual land cover change in land cover cell counts. The United States Geological Survey (USGS) offers the National Land Cover Dataset (NLCD) for the years 1992, 2001, 2006, and 2011 — only four datasets for single years over a time span of 19 years. Other datasets were found, but

they have coarse spatial resolutions greater than 100 meters square. The NLCD has a spatial resolution of 30 meters and a desired high-level classification system, but it lacks data for determination of annual land cover change trends.

This thesis describes two topics: (1) creating a dataset approximating the NLCD on an annual basis for a period of at least 10 years and (2) using this dataset to determine land cover trends, particularly in the wetlands land cover, to develop decision rules for land cover change. The first topic is accomplished by using Variable Spectral Unmixing (VSU) — a modified version of the Variable Multiple Endmember Spectral Mixture Analysis (VMESMA) technique — and hierarchal classification rules, as covered in the following section. This method is used to generate a time series of maps, where the Moore and von Neumann neighborhoods of each pixel are analyzed for frequency of occurrence. The top neighborhoods per class transition are then structured into decision rules and tested in a simple iterative model. The following sections provide detail on these two topics.

2. Approximating the National Land Cover Dataset Using Variable Spectral Unmixing (VSU)

2.1 Section Overview

Spectral mixture analysis (SMA) is a sub-pixel mapping technique that estimates the fractional weights of a set of spectral endmembers whose spectral responses are believed to contribute to the response of a given pixel [13]. Its applications range from land cover mapping to mineral mapping. Different forms of SMA arise from optical properties of sensors and the atmosphere, such as linear versus nonlinear scattering, as well as due to mathematical operations proposed in previous work. This study relies on linear mixture techniques. The following section discusses research that inspired the technique proposed in this thesis.

For a given pixel i , basic linear SMA assumes a spectral response x_i is the linear weighted sum of a set of spectral responses of endmembers, which relate to a feature class, as expressed in equation (2.1):

$$\vec{x}_i = \vec{E} \cdot \vec{f} + \vec{e} \quad (2.1)$$

, where x_i is a vector of a spectral response of the i -th pixel; E is a n by m matrix with n spectral bands and m endmember responses; f is a vector of m fractional weights; and e is a vector of residual errors [20]. Equation (2.1) forms an optimization problem solvable using least-square solvers and subject to two constraints: (1) each weight f_i must be greater than

or equal to zero, and (2) all weights must sum to one [20]. Using a set of n endmembers minimizes the errors in the error vector e .

Basic SMA involves the selection and use of a constant set of a small number of endmembers for the analysis of an entire image, thereby producing sub-pixel compositional maps of each endmember. This technique is simple to implement and provides estimates of the physical composition of a pixel [19]. The number of endmembers permitted in SMA is less than or equal to the number of bands in a remote sensing image in order to limit the root mean square error of the solution to the linear system of equations per pixel. Upon selection, the set of endmembers may not be changed for an image. As a result, basic SMA fails to account for variations in cover types, permitting the exclusion of some feature classes depending on the selection of endmembers (i.e. including only endmembers of artificial and soils rather than including one or more vegetative endmembers) and the confusion between endmembers if any are spectrally similar [19]. Additionally, endmembers are selected to map a specific feature classes (such as forest, soil, or urban areas) for an entire image and do not correspond directly to physical surface materials due to the constancy of the set of endmembers. Variations in illumination (including shade) and atmospheric conditions may further increase errors.

Roberts et al. [19] introduced the technique called multiple endmember spectral mixture analysis (MESMA), which permits variations in the number and type of endmembers per pixel in an image via strict selection criteria of endmember matrices to produce different two- and three-endmember models. Their technique incorporates linear SMA and presents a rigorous method for selecting a model from several possible sets

created through combinations of endmembers within a library using three criteria related to the performance of each model. MESMA results in compositional estimates of physical endmembers rather than feature classes and allows variations in cover types, improving on the failures of basic linear SMA. MESMA is less easily implemented compared to SMA.

As noted by Garcia-Haro et al. [8], MESMA is computationally intensive due to its method of selecting the appropriate model per pixel; it allows inconsistencies in endmember selection; and it requires a library with numerous endmembers, further exacerbating computation durations. Garcia-Haro et al. proposed an alternative method called Variable Multiple Endmember Spectral Mixture Analysis (VMESMA) to address the errors inherent in MESMA and reduce its computation time. In addition to these objectives, VMESMA addresses other issues with linear SMA, including reduction of errors related to illumination and atmospheric conditions by standardizing a spectral library.

VMESMA incorporates a rigorous iterative method for selecting endmember matrices and two solutions of linear and standard unmixing from which to choose. Standardized unmixing further lowers the errors due to the influences of illumination and atmospheric conditions on the spectral response of a given pixel [8]. The selection of endmembers begins with image segmentation to target a feature class and optimize endmember matrices to identify the target. Prior to unmixing, the spectral matching algorithm further reduces the number of endmembers for a given pixel using three criteria with thresholds, including the Euclidean distance, standardized Euclidean distance, and spectral angle matching as calculated in spectral space between each endmember spectra and the response of the given pixel. Then, similar to MESMA, VMESMA reiteratively passes

through this subset of satisfactory endmembers to select those for a model which yields the least squares solution to linear SMA with the minimal root mean square error. Though this method is more accurate than linear SMA and MESMA, VMESMA requires *a priori* knowledge of a target area and is applicable to small locations to produce accurate results of a target feature class. VMESMA is less computationally intensive as compared to MESMA but still requires extensive time.

The method, called Variable Spectral Unmixing (VSU), proposed in this thesis simplifies the VMESMA approach and applies it to land cover mapping with a few notable differences. First, the algorithm segments the image into two components to extract water from land using the modified normalized difference water index (MNDWI), as discussed in detail in the following section. All pixels extracted as water are automatically assigned the corresponding land cover class code and passed over during analysis. Pixels associated with land undergo spectral matching and linear SMA prior to classification. The spectral matching algorithm uses two criteria for endmember selection, namely standardized Euclidean distance and spectral angle matching with standardized angles to attempt reduction in computational costs. Finally, it classifies each pixel by its fractional weights using a set of custom hierarchical rules.

As compared with VMESMA, the VSL method proposed here does not require *a priori* knowledge of a target location but is computationally intensive. Eliminating a few steps within VMESMA results in reduced computation duration, as implemented in this research in Python, but like VMESMA and MESMA, mapping large regions such as the entire United States, would require at least a few years on a single computer if left un-parallelized.

This method is fully autonomous with the full spectral library. Its library may be reduced to eliminate specific endmembers or endmember types that prevent detection of desired land cover classes, such as urban classes, as determined by initial tests prior to mapping an entire time series. This new method is discussed in detail in the subsequent sections. Lastly, in general, most applications of spectral mixture analysis use two or three endmembers, but this research uses six endmembers in order to increase the accuracy of the mixture analysis without comprising spectral variability.

2.2 Methodology

2.2.1 Data Sources

This research utilizes imagery and spectral libraries freely offered by agencies within the U.S. government for scientific purposes. The following subsections briefly describe the sources and the data used in this research.

2.2.1.1 Landsat

The National Aeronautic and Space Administration (NASA) operates the Landsat program for the United States Geological Survey (USGS). The Landsat program consists of five retired and two currently operating satellites, each named Landsat with an identifying number in the series. Landsats 7 and 8 currently image the Earth at an altitude of 705 km with a revisit time of approximately 16 days for any given location on the surface [18]. Landsats prior to LS 7 are no longer operational, but their imagery remains available through USGS EarthExplorer (an Internet tool used to disseminate Landsat and other overhead sources of imagery for scientific purposes for free).

This research uses imagery captured by Landsat 5 Thematic Mapper (TM) and Landsat 7 Enhanced Thematic Mapper Plus (ETM+). The TM sensor preceded the ETM+ but both share the same bands that are used in this research. Both the TM and the ETM+ sensors record data for seven bands of wavelengths, as seen in Table 2.1 [18]. Each sensor records a single measurement for each band over its corresponding range of wavelengths. Each band has a square spatial resolution of 30 meters on a side [18]. NASA provides the relative spectral responses for each band of each sensor for use in scientific studies. Note,

however, that only six bands are used in this research. Band 6, which covers the thermal infrared region of the electromagnetic spectrum, is excluded.

Table 2.1 - Landsat 5 & 7 Sensor Bandwidths

Satellite	Sensor	Bands / Spectrum	Wavelengths (um)
Landsat 5	TM	1 - Blue	0.45 - 0.52
Landsat 7	ETM+	2 - Green	0.52 - 0.60
		3 - Red	0.63 - 0.69
		4 - NIR	0.76 - 0.90
		5 - SWIR1	1.55 - 1.75
		7 - SWIR2	2.08 - 2.35

The USGS processes each Landsat image for geometric and radiometric corrections prior to releasing each as a Level 1 Product, which include a single GeoTIFF image per band, a world file, a metadata file, and, if geometrically corrected, a GeoTIFF containing the ground control points. The images contained within each Level 1 Product store measurements as digital numbers, which must be converted to radiance and from radiance to reflectance prior to use in any analysis. This research utilizes the Level 1 Products for images captured by Landsats 5 and 7 over bands 1 through 5 and 7 for the years from 2001 to 2011.

2.2.1.2 Spectral Libraries

The USGS Spectroscopy Lab located in Denver, CO, offers the Digital Spectral Library (most recent version splib06a), which consists of over 1300+ spectra belonging to mineral, soil, vegetative, and artificial materials [4]. In general, each endmember spectrum includes measurements in units of percent reflectance for the range of wavelengths of 0.45 to 2.5 microns in addition to metadata consisting of the each material's classification, appearance, method of measurement, source, etc. Similarly, the NASA Jet Propulsion Laboratory (JPL) provides the Advanced Spaceborne Thermal Emission Reflection Radiometer (ASTER) Spectral Library, which include spectra taken from images captured by the ASTER sensor flying on the NASA Terra satellite in addition to spectra from other sources, including the USGS Spectroscopy Lab [2]. These spectral libraries are referred to as the USGS and JPL spectral libraries for the rest of this document.

The USGS and JPL spectral libraries overlap each other in the material types of soil, mineral, and artificial materials. The JPL spectral library, however, contains only four vegetation spectra for dry grass, healthy grass, and deciduous and coniferous trees [2]. In comparison, the USGS spectral library offers data covering a wide selection of various species of coniferous and deciduous trees and shrubs, grasses and herbaceous plants, rangeland, wetlands, cacti, and domestic plants [4]. When used together, they provide a wide selection of spectra for use in analysis. Both lack spectra for pasture, hay, and crops. The exact reasons are not known. It may be speculated, however, that these spectra were excluded based on the high spectral variability in such factors as plant height and annual illumination, varying growth and planting cycles by region, and differing soil compositions

by region. As a result, the land cover classes associated with pasture / hay and cultivated crops are not mapped. Note that pasture and hay, excluding other vegetation occurring in this class, are spectrally similar to grassland graminoid species. Thus, it may be possible to map the pasture / hay land cover class incorporating graminoid species using a classification rule.

2.2.1.3 National Land Cover Dataset

The third dataset used in this research includes the USGS National Land Cover Dataset (NLCD). The USGS preprocesses Landsat images according to a specific manual and generates land cover maps using a comprehensive approach by first spectrally clustering all pixels into target land cover classes, comparing a seasonal time series to establish classes which change rapidly throughout a single year (such as cultivated cropland), and validates it with further steps, including comparisons to ancillary datasets [21]. This process requires at least five years to completely map the entire United States, including Alaska and Hawaii, and validate results.

The NLCD includes datasets for the years 1992, 2001, 2006, and 2011, but only the last three years utilize the same mapping criteria and will be used as the control dataset for this study. Each map has a square spatial resolution of 30 meters and utilizes 16 land cover classes for the contiguous United States and Hawaii [21]. Alaska requires an additional four classes. All generated and modeled maps use the same color scheme as is captured by the NLCD 2011 legend in Figure 2.1 (below).



Figure 2.1 - NLCD 2011 Legend [11]

2.2.2 Image Selection

Land cover tends to vary spectrally due to various reasons, including variations in annual lighting conditions due to the inherent geometries of the satellite in relation to the Earth and Sun and atmospheric conditions regardless of the location of the surface target on the Earth. The best results of spectral mixture analysis occur when variations due to such parameters as temporal, luminary, and atmospheric conditions are minimized. Selecting images occurring on the same date minimizes temporal and luminary variations and ignores errors due to potential atmospheric variations. Selecting images with limited to no haze or cloud contamination may minimize errors due to atmospheric conditions.

Unfortunately, atmospheric conditions are not guaranteed to be consistent on an annual basis.

In order to minimize errors due to luminary, temporal, and atmospheric conditions, images may be selected from a range of dates under the condition that the image contains limited to no haze or cloud contamination. Glancing through the scenes available via Earth Explorer demonstrate the difficulty in selecting a scene for path-row 022-039, as Louisiana frequently lies under cloud cover. Images within the time period of late summer to late fall, or from July to November, appear to occur with cloud contamination under 20%. This research uses this time frame as the period as a criterion in image selection, preferring images with cloud contamination less than 20% occurring as close to August 1st as possible. Some images may occur as early and as late as June and November, respectively, due to availability of images with less than 20% cloud contamination.

2.2.3 Preprocessing

NLCD & Landsat Scene Clipping

The NLCD spatial reference system uses the North American Datum of 1983 (NAD83) as the geodetic model and the map projection of Albers Conical Equal Area for the Lower 48 states (AlbersL48). This geodetic model and map projection are used for all geographical images in this thesis. The World Reference System 2 (WRS2) provides shapefiles for each Landsat scene. The shapefile for path-row 022-039 was reprojected to use the same geodetic model and map projection as the NLCD using ArcGIS. Each NLCD image was clipped by the reprojected WRS2 path-row 22-39 shapefile. Prior to

preprocessing, all Landsat scenes were reprojected to the same geodetic model and map projection as the NLCD and clipped using the reprojected shapefile. This eliminates band fringing in the Level 1 Product scenes.

Landsat Scene Preprocessing

USGS offers the Level 1 Product created from raw Landsat scenes. The Level 1 Product records data in the form of digital numbers (DN). Spectral mixture analysis analyzes data in units of percent reflectance and, thus, the DN require conversion to this unit of measurement. According to the USGS' *MRLC Image Processing Manual*, the conversion process consists of two steps: conversion from DN to radiance and from radiance to percent reflectance. These conversions are performed using the following equations (2.2) and (2.3), respectively:

$$L_{\lambda} = Gain_{\lambda} \cdot DN_{\lambda} + Bias_{\lambda} \quad (2.2)$$

$$\rho_{\lambda} = \frac{\pi \cdot L_{\lambda} \cdot d^2}{ESUN_{\lambda} \cdot \sin \theta} \quad (2.3)$$

, where λ is the band number; L is the at-satellite radiance per band; $Gain$ is the band specific gain; $Bias$ is the band specific bias; ρ is the at-satellite reflectance; d is the Earth-Sun distance in astronomical units; $ESUN$ is the solar exoatmospheric spectral irradiance; θ is the sun elevation angle [21]. The resulting reflectance measurements are multiplied by 100. All bands are stacked in order of descending band number and saved to a GeoTIFF prior to analysis. The values for the solar exoatmospheric spectral irradiance are given in

Table 2.2. The value of the distance between the Earth and the Sun is referenced from a text file available from NASA.

Table 2.2 - ESUN Solar Spectral Irradiances

Band	1	2	3	4	5	7
LS5 TM	1957.000	1826.000	1554.000	1036.000	215.000	80.670
LS7 ETM+	1969.000	1840.000	1551.000	1044.000	225.700	82.070

Prior to these calculations, the constants are extracted and selected according to guidelines set in the USGS *MRLC Image Processing Manual* [21]. All Landsat 5 scenes occurring after May 2003 are processed with gains and biases specific to that vehicle and scenes occurring prior to May 2003 are processed with gains and biases specific to Landsat 7. This makes Landsat 5 scenes captured prior to May 2003 compatible with Landsat 7 scenes for analytical purposes [21]. To enable this processing, the Landsat 5 DN are converted to Landsat 7 DN by the following equation (2.4) prior to the conversions from DN to reflectance:

$$DN7 = DN5 \cdot m + i \quad (2.4)$$

, where $DN7$ is the Landsat 7 DN; $DN5$ is the Landsat 5 DN; and m and i are the slope and intercept values, respectively, to the measurements in Table 2.3.

Table 2.3 - Slope and Intercept Values

Band	1	2	3	4	5	7
Slope (m)	0.9398	1.7731	1.5348	1.4239	0.9828	1.3017
Intercept (i)	4.2934	4.7289	3.9796	7.032	7.0185	7.6568

Spectral Libraries Sampling

The USGS and JPL spectral libraries contain measurements of percent reflectance for wavelengths generally occurring in the bandwidths of the Landsat TM and ETM+ sensors. As previously stated, each Landsat sensor bandwidth results in one measurement. In order to simulate the spectral signature as seen by Landsat, each spectral signature of the combined spectral library is sampled at the wavelength corresponding to where the relative spectral response of each Landsat 5 TM sensor bandwidth equals or is closet to one. This provides variation per band within the library but does not fully simulate the spectral response seen with atmosphere. The wavelengths corresponding to the points in each Landsat 5 TM bandwidth at which the relative spectral response equals or approaches one are given in Table 2.4.

Table 2.4 - LS5 TM Relative Spectral Response Points

Band	1	2	3	4	5	7
Wavelength (μm)	503	594	677	800	1711	2199

2.2.4 Variable Spectral Unmixing (VSU)

2.2.4.1 Algorithm Overview

The Variable Spectral Unmixing (VSU) algorithm iterates once per pixel through the following steps prior to classifying a pixel as a land cover class:

1. Spectral Matching
2. Spectral Mixture Analysis
3. Classification

Each process involves a series of steps relying on calculations from reflectance data in an effort to reduce processing durations from requiring a few minutes per pixel to mere seconds. Additionally, the third step of classification represents an algorithm with a general purpose of interpreting the results of Spectral Matching and Spectral Mixture Analysis to identify a feature class pertaining to the given pixel. In this case, the classification algorithm identifies an appropriate land cover class, but it may be replaced with rules to map other physical surface features of interest for other applications, including but not limited to mapping minerals, burnt areas, sea ice, snow cover, or crops. The following subsections detail the underlying methodology of each step.

2.2.4.2 Extraction of Water

Large areas of water lie in some of the locations to be mapped for this analysis. In order to quicken the mapping process, areas of water are extracted prior to and passed over during mapping. The algorithm makes use of the modified normalized difference water index (MNDWI), as seen in Eq. 2.5 below, where *Green* corresponds to band 2 (*B2*)

and *MIR* corresponds to band 5 (*B5*) [22]. As recommended by Xu, the MNDWI has a threshold on the range of [0.0,1.0] in order to extract all water pixels. However, some pixels belonging to land cover classes with spectral similarities to water may also be extracted. These land features include water-infused earthen materials, such as mud, clay, and sand, and water- or oil-covered urban materials, such as some heavily trafficked roadways containing oil or other materials leaking from passing vehicles.

$$MNDWI = \frac{Green - MIR}{Green + MIR} = \frac{B2 - B5}{B2 + B5} \quad (2.5)$$

2.2.4.3 Spectral Matching

Spectral matching aims to identify and select the endmembers to include in spectral mixture analysis. This sub-algorithm begins by calculating the standard vectors of the given pixel and every endmember within the spectral library using Eq. (2.11) in Table 2.6. Eqs. (2.9) and (2.10) calculate the standard Euclidean distance and the spectral angle of each endmember's reflectance vector with respect to that of the given pixel using the standardized and unstandardized data, respectively. Spectral angles undergo standardization using Eq. (2.11) in order to simplify sorting. Results are sorted in descending order in order of preference for standard Euclidean distance values over spectral angles and then in descending order of spectral angles. This produces an array with the endmembers with minimal distance and angle values at the top.

Table 2.5 - Parameters, Equations, & Thresholds for Spectral Matching

Technique	Equation	Eq. No.	Threshold
Standard Euclidean Distance [13]	$\hat{D} = \ \hat{x} - \hat{y}\ _2$	2.9	1 unit
Spectral Angle [13]	$\theta = \cos^{-1} \left(\frac{\overline{R \cdot E}}{\ \overline{R}\ \ \overline{E}\ } \right)$	2.10	1 degree*
Standard Vector [13]	$\hat{r} = \frac{\overline{r} - \mu_r}{\sigma_r}$	2.11	N/A

* After standardization of the vector of spectral angles.

A secondary sort extracts the sorted endmembers according to those whose distances and angles meet the limits given in Table 2.6. Eqs. (2.9) and (2.10) produce values on the real line with a minimum of 0. According to de Jong, et al. [13], endmembers with distances of 0 units and angles of 0 degrees are perfect matches and endmembers with values closest to 0 are near perfect matches. Endmembers with distances and angles within the range of [0,1] are acceptable matches.

The calculations used within this sub-algorithm may produce a number of matches less than or greater than the number of bands within the image. If the number of endmembers is less than the number of bands, then all endmembers are selected. If the number of endmembers equals or surpasses the number of bands, then only the first six endmembers are selected because they possess the smallest distance and angle values that minimize the root mean square error of the least squares solution of Eq. (2.1). Prior to passing the matched endmembers to the next step, this sub-algorithm constructs a matrix from the reflectance measurements of the selected endmembers with the spectral signatures of each endmember occupying a column.

2.2.4.4 Spectral Mixture Analysis

Both MESMA and VMESMA solve Eq. (2.1) as a least-squares optimization problem with differences in the constraints. This VSU method solves Eq. (2.1) without constraints using a least squares solver provided by the linear algebra submodule in NumPy. The results are real numbers which are normalized to meet the constraints generally used in linear spectral mixture analysis and as given in Section 2.1. The least squares solver calculates the residual error (the root mean square error) associated with the solution and this error is recorded for a later calculation of the mean residual error of the modified MESMA.

2.2.4.5 Classification by Endmembers and Rules

Results of spectral mixture analysis receive preliminary sub-pixel classification based on the link between the material type to which an endmember belongs and a basic land cover class (such as coniferous or deciduous forest and cultivated crops). Endmembers within spectral libraries are classified by overarching material types, which include manmade or artificial substances; soils, rocks, and other earthen substrates; water phases; and vegetation. Vegetation is subdivided into the subcategories of coniferous and deciduous trees, shrubs and scrubs, herbaceous plants, grasses, wetlands, and crops [4][2]. Note that some categorical systems of vegetation may segment hay and pasture into its own category in recognition of its corresponding land use. For basic land cover classes belonging to a single category, endmembers of the same category may be linked to the corresponding class. For example, endmembers belonging to the material types of

manmade / artificial substances and coniferous trees lead to preliminary classification as developed and evergreen forest, respectively. These links are shown in Table 2.4.

Post sub-pixel classification, a predominant land cover class is assigned based on rules adapted from the USGS NLCD classification scheme and its criteria, as seen in Table 2.4. As calculated by SMA, the compositional weights represent the estimated percentage of a material type contributing to the land cover of a given pixel. The compositional weights belonging to the same material type are summed to calculate the cumulative compositions by material type and, hence, basic land cover types belonging to pure land cover classes. As a result of the mixed pixel problem, pixels of medium to large coarse spatial resolutions may contain multiple land cover types within their area [13]. Thus, the classification rules adapted from criteria in Table 2.6 assign a predominant class to a given pixel based on preliminary sub-pixel classification and the cumulative composition weights by material type. Note, only the classes found in the contiguous US are presented in Table 2.6. The table excludes classes found only in Alaska, as they are not encountered naturally in the study location.

Table 2.6 - USGS NLCD Classification Classes, Codes, Criteria, & Corresponding Materials

Class / Code	Criteria	Material Type
Open Water, 11	Areas of open water, generally with less than 25% cover of vegetation or soil.	Liquid water
Perennial Ice/Snow, 12	Areas characterized by a perennial cover of ice and/or snow, generally greater than 25% of total cover.	Solid water
Developed, Open Space, 21	Areas with a mixture of some constructed materials, but mostly vegetation in the form of lawn grasses. Impervious surfaces account for less than 20% of total cover. These areas most commonly include large-lot single-family housing units, parks, golf courses, and vegetation planted in developed settings for recreation, erosion control, or aesthetic purposes.	Manmade materials
Developed, Low	Areas with a mixture of constructed materials and	Manmade

Intensity, 22	vegetation. Impervious surfaces account for 20% to 49% of total cover. These areas most commonly include single-family housing units.	materials
Developed, Medium Intensity, 23	Areas with a mixture of constructed materials and vegetation. Impervious surfaces account 50% to 79% of the total cover. These areas most commonly include single-family housing units.	Manmade materials
Developed, High Intensity, 24	Highly developed areas where people reside or work in high numbers. Examples include apartment complexes, row houses, and commercial/industrial. Impervious surfaces account for 80% to 100% of the total cover.	Manmade materials
Barren Land (Rock/Sand/Clay), 31	Areas of bedrock, desert pavement, scarps, talus, slides, volcanic material, glacial debris, sand dunes, strip mines, gravel pits, and other accumulations of earthen material. Generally, vegetation accounts for less than 15% of total cover.	Minerals, rocks, soils, sand, clay, etc.
Deciduous Forest, 41	Areas dominated by trees generally greater than 5 meters tall, and greater than 20% of total vegetation cover. More than 75% of the tree species shed foliage simultaneously in response to seasonal change.	Deciduous trees
Evergreen Forest, 42	Areas dominated by trees generally greater than 5 meters tall, and greater than 20% of total vegetation cover. More than 75 of the tree species maintain their leaves all year. Canopy is never without green foliage.	Coniferous trees
Mixed Forest, 43	Areas dominated by trees generally greater than 5 meters tall, and greater than 20% of total vegetation cover. Neither deciduous nor evergreen species are greater than 75% of total tree cover.	Deciduous and coniferous trees
Shrub / Scrub, 52	Areas dominated by shrubs; less than 5 meters tall with shrub canopy typically greater than 20% of total vegetation. This class includes true shrubs, young trees in an early successional stage or trees stunted from environmental conditions.	Shrubs and scrubs
Grassland / Herbaceous , 71	Areas dominated by gramanoid or herbaceous vegetation, generally greater than 80% of total vegetation. These areas are not subject to intensive management such as tiling , but can be utilized for grazing.	Grasses and herbaceous plants
Pasture / Hay, 81	Areas of grasses, legumes, or grass-legume mixtures planted for livestock grazing or the production of seed or hay crops, typically on a perennial cycle. Pasture/hay vegetation accounts for greater than 20% of total vegetation.	Grasses and herbaceous plants
Cultivated, Crops, 82	Areas used for the production of annual, crops, such as corn, soybeans, vegetables, tobacco, and cotton, and also perennial woody crops such as orchards and vineyards. Crop vegetation accounts for greater than 20% of total vegetation. This class also including all land being actively tilled.	Crops

Woody Wetlands, 90	Areas where forest or shrub land vegetation accounts for greater than 20% of vegetative cover and the soil or substrate is periodically saturated with or covered with water.	Deciduous and coniferous trees
Emergent Herbaceous Wetlands, 95	Areas where perennial herbaceous vegetation accounts for greater than 80% of vegetative cover and the soil or substrate is periodically saturated with or covered with water.	Wetlands

The classification rules derived from the USGS NLCD classification criteria in Table 2.6 are shown in Figure 2.2. After the initial sub-pixel classification and the cumulative sum of compositional weights by material type, the classification rules decide on the final label for a given pixel. Each rule uses three criteria for determining the final land cover class and these include: (1) whether the target class of the rule occurs in the sub-pixel classification; (2) whether the cumulative compositional weight of the target class meets a set threshold; and (3) whether the number of material types with cumulative weights below the weight of the target class equals the number of material types minus one as an attempt to ensure the class has the maximal cumulative weight. If these three conditions are met, then the pixel is classified as the target class of the rule.

The developed and the woody wetlands classes are exempt from this rule format due to their mixed nature. Due to the varied ranges of the developed classes, the corresponding rules use only the first and second criteria for classification. The woody wetlands class is either not included due to a lack of endmembers or is validated using an ancillary dataset and a logic expression of whether the pixel in the mapped area contains one of the three forest classes. Note, in Figure 2.2, *Clazz* is a vector containing the sub-pixel classification by endmembers, *Mats* is a vector containing the cumulative sum of all weights per endmember type, and *Over* is the final classification.

```

Mats = Mixes6(W, Eye) # calculates the cumulative sum of all weights of the same type
Over = -1
#C = [Manmade, Soils/Rocks, Water, Domestic, Grasses, Herbs, Deciduous, Coniferous,
Shrubs/Scrubs, Hay, Crops, Wetlands]

if Index == 1 and Mats[2] > 0.2:
    Over = 11
else:
    if 21 in Clazz and 0.07 < Mats[0] < 0.2 and sum(x > Mats[0] for x in Mats) >= 4:
        Over = 21
    elif 21 in Clazz and 0.2 <= Mats[0] < 0.5 and sum(x > Mats[0] for x in Mats) >= 5:
        Over = 22
    elif 21 in Clazz and 0.5 <= Mats[0] < 0.8 and sum(x < Mats[0] for x in Mats) >= 5:
        Over = 23
    elif 21 in Clazz and 0.8 <= Mats[0] <= 1.0 and sum(x < Mats[0] for x in Mats) >= 5:
        Over = 24

    if 31 in Clazz and Mats[1] > 0.15 and (sum(x < Mats[1] for x in Mats) == 11 or (sum(x < Mats[1]
for x in Mats) >= 7 and sum(Mats[4:12]) < 0.2)):
        Over = 31
    elif 71 in Clazz and (Mats[4] > 0.4 or Mats[5] > 0.5 or sum(Mats[4:6]) > 0.4) and (sum(x <
Mats[4] for x in Mats) >= 8 or sum(x < Mats[5] for x in Mats) //
>= 8 or sum(x < (Mats[4] + Mats[5]) for x in Mats) >= 8):
        Over = 71
    elif 42 in Clazz and Mats[7] > 0.7 and (sum(x < Mats[7] for x in Mats) == 11 or (Mats[7] >
Mats[6] and Mats[6] < 0.5 and sum(x < Mats[7] for x in Mats) //== 10)):
        Over = 42
    elif 41 in Clazz and Mats[6] > 0.7 and (sum(x < Mats[6] for x in Mats) == 11 or (Mats[6] >
Mats[7] and Mats[7] < 0.5 and sum(x < Mats[6] for x in Mats) // == 10)):
        Over = 41
    elif 41 in Clazz and 42 in Clazz and 0.4 < Mats[6] < 0.7 and 0.4 < Mats[7] < 0.7 and sum(x <
sum(Mats[6:8]) for x in Mats) == 10:
        Over = 43
    elif 52 in Clazz and Mats[8] > 0.5 and sum(x < Mats[8] for x in Mats) == 11:
        Over = 52
    elif (71 in Clazz or 81 in Clazz) and (0.2 < Mats[4] < 0.5 or 0.2 < Mats[5] < 0.5 or 0.2 <
(Mats[4] + Mats[5]) < 0.5) and (sum(x < Mats[4] for x in Mats) == //11 or sum(x
< Mats[5] for x in Mats) == 11 or sum(x < (Mats[4] + Mats[5]) for x in Mats) == 10):
        Over = 81
    elif 82 in Clazz and Mats[10] > 0.3 and sum(x < Mats[10] for x in Mats) == 11:
        Over = 82
    elif 95 in Clazz and Mats[11] > 0.8 and sum(x < Mats[11] for x in Mats) == 11:
        Over = 95

if Over <= 0:
    Over = Clazz[0]

```

Figure 2.2 - Code snippet showing classification rules used with VSU

2.2.5 Error Analysis

As the purpose of this mapping procedure is to create a time series of land cover maps comparable to the NLCD, the years of 2001, 2006, and 2011 are mapped for error analysis and validation against the NLCD. Two methods used in the error analysis include the calculation of the percent agreement between the generated maps and the NLCD and an error matrix. The percent agreement is calculated by taking the difference of the NLCD and generated map images, counting the number of zeros signifying a perfect match, and dividing this count by the total number of pixels encompassing the mapped location. The error matrix is generated by comparing the generated maps to the corresponding NLCD and counting the land cover transitions from the mapped dataset to the NLCD. The error matrix shows where classes in the NLCD are mistaken for other classes in the generated map.

Direct comparison between the generated maps and the NLCD is recognized to be a rough measure of agreement due to differences in mapping methods. The USGS produces the NLCD using a rigorous procedure of preliminary unsupervised spectral clustering of multi-temporal Landsat scenes from the same target year, followed by extensive validation processes using human expertise in pattern recognition and ancillary datasets whereas the VSU method is a single-pass, autonomous mapping method using hierarchal classification rules derived from the NLCD classification criteria. Due to this, the results of this method should be further validated using photo-interpretative techniques.

2.3 Results & Discussion

Figures 2.3, 2.4, and 2.5 show the mapping results for the same years corresponding to the NLCD for the image spatial coordinates of the upper left hand corner of each image: [2100,1900], [3620,1060], and [5450,4650], respectively. Refer to the NLCD legend in Figure 2.1 in Section 2.2.1.3 to see the class-color coordination. Each location mapped for the years 2001, 2006, and 2011 show changes in the following land cover changes respectively: forest, shrub land, grassland, pasture/hay, and woody wetlands; developed, pasture / hay, cultivated crops, and woody wetlands; and emergent herbaceous wetlands and water. The following presentation of results and discussion focuses primarily on graphical comparisons of the VSU-generated maps with the USGS NLCD.

When compared visually, there are areas of similarly classified and misclassified pixels between the VSU-generated maps and the NLCD. For example, in the upper left-hand corner of the first locale (A) in Figures 2.3, 2.4, and 2.5, the NLCD features large areas of land belonging to the grassland, shrub land, and coniferous forest classes in 2001, which gradually change to predominantly coniferous forest in 2011. In comparison, the same area of the VSU-generated maps of the first locale (B) feature a mixture of grassland, shrub land, and coniferous forest in 2001 that transmutes to predominantly coniferous forest in 2011. While the same trend of land cover change from a mixture of grassland, shrub land, and coniferous forest to paramount coniferous forest is seen in both datasets, differences exist between the two time series for the single locale. The 2001 VSU-image in Figure 2.3 (B) of the upper left-hand corner features a mixture of shrub land and deciduous forest amid grassland and coniferous forest where the NLCD contains shrub land surrounded by

grassland and coniferous forest. For the upper left-hand corner in Figure 2.4 (B), the region in 2006 shows mixtures of shrub land, deciduous forest, and coniferous forest whereas the NLCD 2006 in Figure 2.4 (A) shows a similar mixture to its previous state in 2001. In Figure 2.5, the same region in the upper left-hand corner of the VSU-generated image (B) and the NLCD 2011 (A) changed to predominantly coniferous forest, but the NLCD 2011 and VSU-generated 2011 image feature shrub land and grassland respectively in a few surrounding areas, respectively. Similar examples may be found upon examination of the images in Figures 2.3, 2.4, and 2.5, and they will be discussed briefly in the rest of this section in addition to other observed errors.

These overarching classifications for vegetation types of deciduous and coniferous trees, shrub land, and grassland occur due to spectral similarities between vegetation types and due to the hierarchal structure and order of the classification rules. Shrubs and scrubs cover a wide range of plants consisting of woody stems and of either coniferous or deciduous foliage characterized by needles and broad, flattened leaves, respectively. As stated in Table 2.6, shrub and scrub land plants grow to maximal heights under five meters and these classes may include young trees under this height limit. As a result, shrub land possesses similar spectral signatures to deciduous and coniferous forest, leading to misinterpretation in classification as a result of these vegetation types occurring in the unmixing results. Similarly, grassland includes graminoid species which may reach heights of low to moderately high scrubs and shrubs and these two vegetation types may occur naturally in grassland, causing spectral similarity between grassland and shrub land classes. The results of SMA reflect estimates of physical materials identified as endmembers

detected during spectral matching [13], and VSU permits multiple endmembers per material type, allowing cumulative weights of material types to vary per pixel in an attempt to improve classification results. Thus, misinterpretations may occur due to the cumulative weight of multiple endmembers of the same material type occurring in the VSU results and a lack of proper rules or another system to resolve these incidents of spectral similarities into an appropriate land cover class.

Additionally, the classification results in Figures 2.3, 2.4, and 2.5 correspond to the material types whose cumulative weights met the conditions in the rules shown in Figure 2.2 but they also reflect the order and hierarchal structure of those rules. Grassland and coniferous forest occur early in the hierarchal structure and experience high proliferation in all three images of Figures 2.3, 2.4, and 2.5 (B) as shown by the large swaths of land in the colors corresponding to these classes. For example, in the case of the upper left-hand corner of the first locale (B) in Figures 2.3, 2.4, and 2.5, coniferous forest arises due to having a high cumulative weight or a cumulative weight greater than the weight of the deciduous forest material type and this rule occurs prior to the rules of deciduous and mixed forest and shrub land but after grassland. Similar structures of the other rules allow those classes to occur in the final results, though a class lower in the hierarchy may be a better interpretation, but their order in the hierarchy contributes to their proliferation in the final map. Thus, the mapped land cover classes represent a base material type with a large cumulative weight or the dominant physical characteristic of the land, namely pure land cover classes, rather than their intended or correct classes due to the hierarchal classification rules.

These trends are seen in the other locales, but limited detection of urban classes and wetlands also appear. In the third locale (NLCD and VSU-generated images (E) and (F), respectively, in Figures 2.3, 2.4, and 2.5), grassland and shrub land were classified in the VSU-generated images as opposed to being classified as emergent herbaceous wetlands. Woody wetlands include large quantities of coniferous and deciduous trees, shrubs, and scrubs whereas emergent herbaceous wetlands include graminoid and herbaceous species, which may include herbaceous shrubs. Thus, the grassland and shrub land classified in the third locale are the same vegetation types as those encountered in emergent herbaceous wetlands. They are misclassified due to the VSU results including these endmembers and meeting the conditions required for these classes in the hierarchical rules prior to the rule for emergent herbaceous wetlands. Another possibility is the low spectral similarity between the few wetlands endmembers and the pixel reflectance measurements, which may be countered by acquiring more spectra for wetlands. Further problems with this locale are discussed later in this section. In the second locale (NLCD and VSU-generated images (C) and (D), respectively, in Figures 2.3, 2.4, and 2.5), vegetation types of grassland, shrub land, and forest approximate the general vegetation encountered in urban areas, including trees and grasses. This signifies the dominance of the vegetation types over artificial and manmade materials associated with the developed classes as a result of greater spectral similarity between the vegetative endmembers with the pixel reflectance measurements and the limited classification of urban areas due to the break in the hierarchical rules between the urban classes and the following rules.

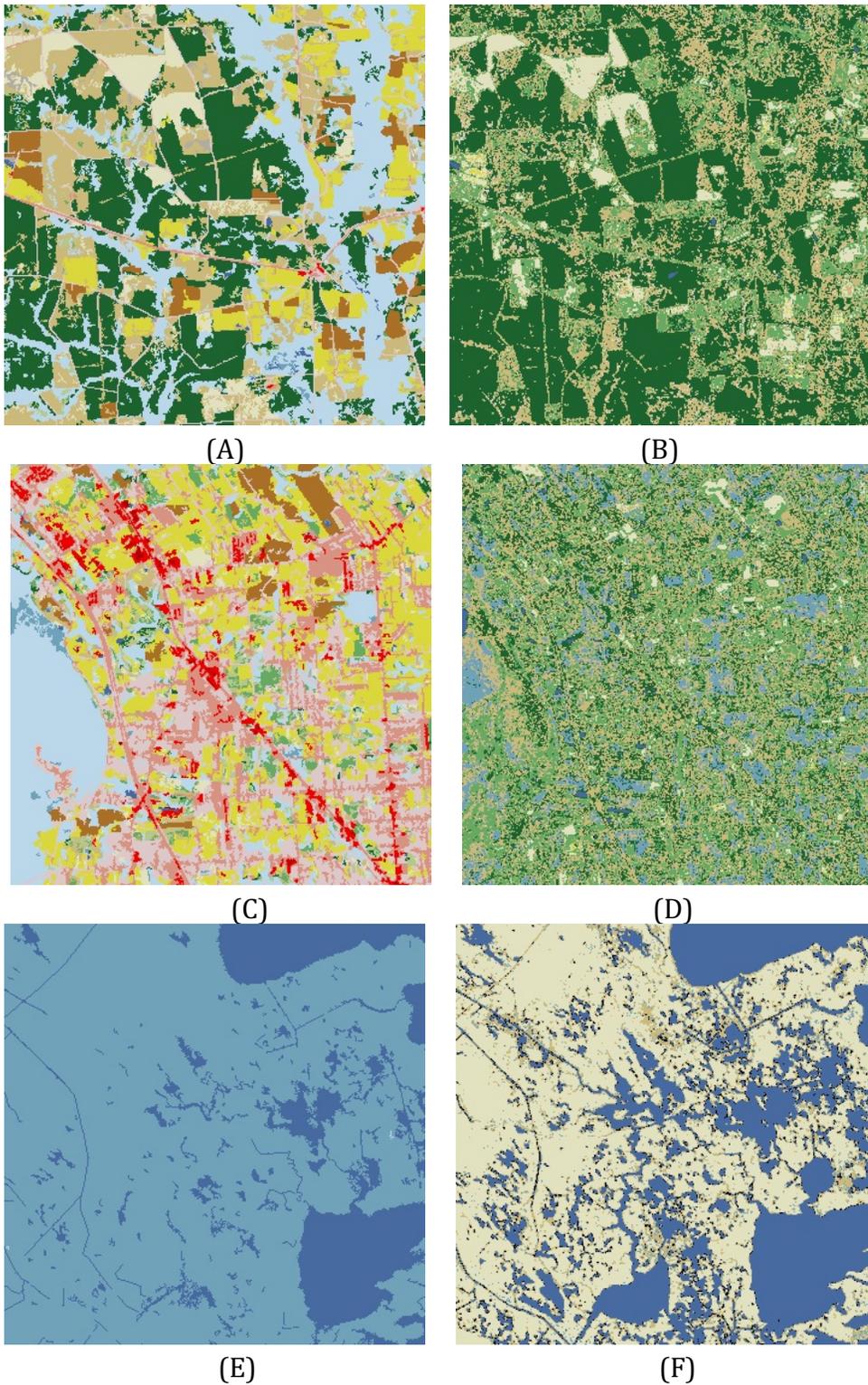


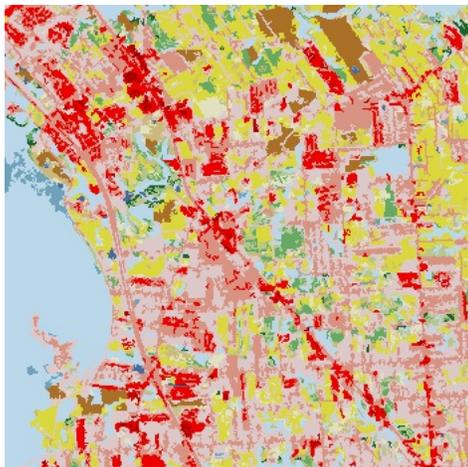
Figure 2.3 - 2001 Mapping Results versus the NLCD 2001 by Locale: (A) NLCD 2001 [2100,1900]; (B) Mapped 2001 [2100,1900]; (C) NLCD 2001 [3620,1060]; (D) Mapped 2001 [3620,1060]; (E) NLCD 2001 [5450,4650]; (F) Mapped 2001 [5450,4650]. Black pixels are unclassified.



(A)



(B)



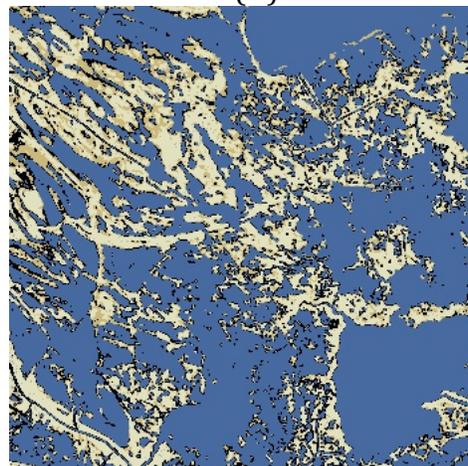
(C)



(D)



(E)



(F)

Figure 2.4 - 2006 Mapping results versus NLCD 2006 by locale: (A) NLCD 2006 [2100, 1900]; (C) NLCD 2006 [3620,1060]; (E) NLCD 2006 [5450,4650]; (B) Mapped 2006 [2100,1900]; (D) Mapped 2006 [3620,1060]; (F) Mapped 2006 [5450,4650]. Black pixels are unclassified.

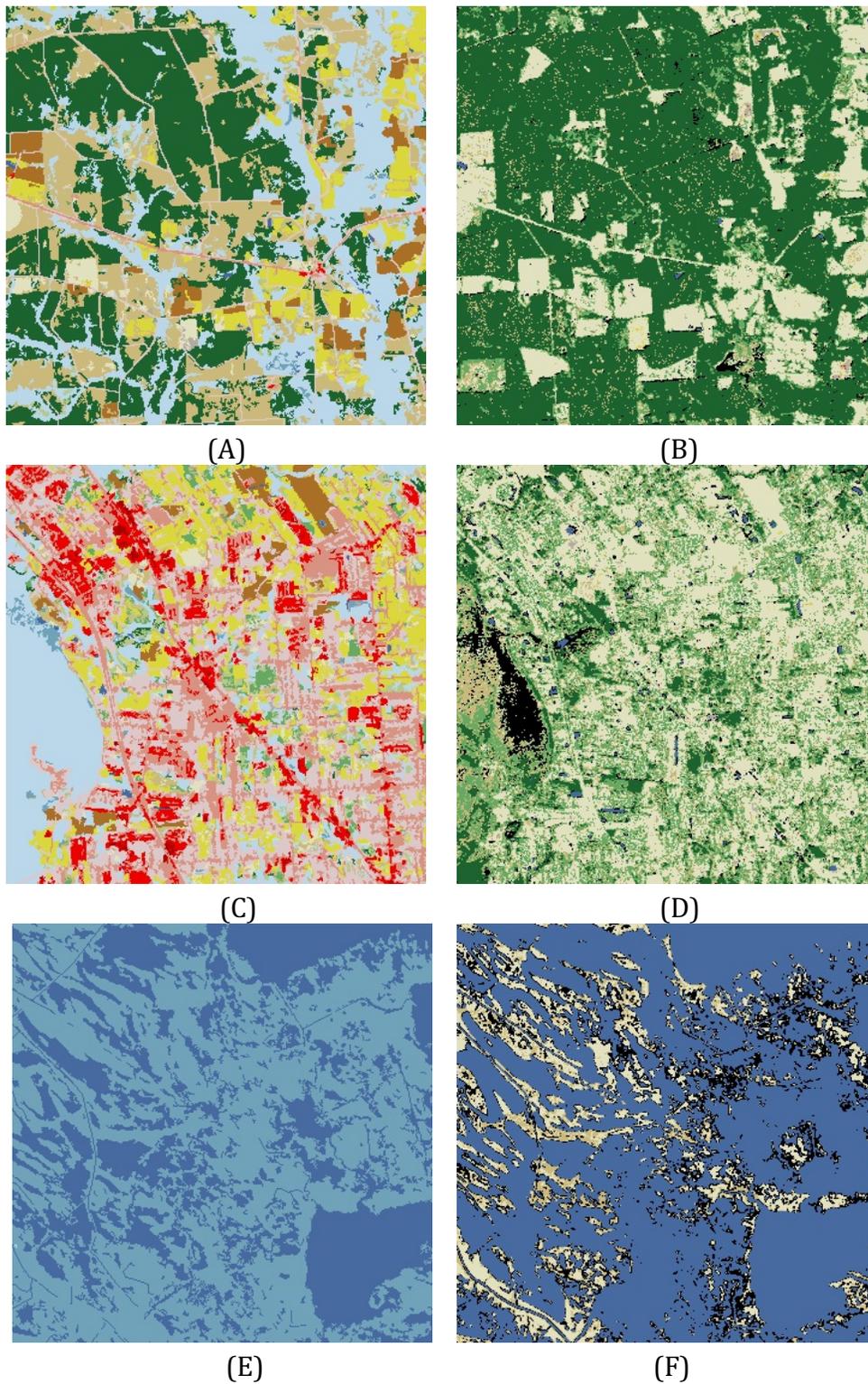


Figure 2.5 - 2011 Mapping Results versus NLCD 2011 by locale: (A) NLCD 2011 [2100,1900]; (C) NLCD 2011 [3620,1060]; (E) NLCD 2011 [5450, 4650]; (B) Mapped 2011 [2100,1900]; (D) Mapped 2011 [3620,1060]; (F) Mapped 2011 [5450,4650]. Black pixels are unclassified.

These results occur primarily due to the hierarchal classification rules seen in Section 2.2.4.5, which attempted to resolve sub-pixel classifications resulting from the mixed pixel problem into a high-level classification to be placed in the resulting maps in Figures 2.3, 2.4, and 2.5. When examining the sub-pixel maps in descending compositional weight, developed land cover codes (21, 22, 23, and 24) often occurred in the lower layers of the sub-pixel map, such as a matrix depth of 4 and 5 that have very low compositional weights, and generally belonged to the open space developed class (21). As a result, the agreements in Table 2.7 between the generated and NLCD maps per each year at locale [3620,1060] feature very low agreements because the developed (21 - 24), pasture / hay (81), and cultivated cropland (82) dominate this locale in the NLCD. It is possible to partially fix this problem by changing the classification system to use different criteria, which will be further discussed below and in Section 4.

The mapping results signify a low spectral similarity between the spectral urban targets in the original Landsat images and the endmembers belonging to urban materials for at least three reasons. First, this occurs due to low representation of various artificial materials and their spectra at different points in their life cycles in addition to the lack of endmembers for pasture / hay and crops at various stages in their annual growth cycles in the spectral library. This problem is easily fixed by incorporating such endmembers into the library. Second, this suggests the endmember sampling method may be partially to blame, which contributes to the high proliferation of vegetated classes in the generated maps. Louisiana features high abundances of the vegetation types within the library and Landsat scenes of the area contain high spectral similarity to vegetation classes, increasing

the difficulty of mapping urban areas within Louisiana. Lastly, temporal disparities in illumination contribute to the limited detection of urban areas and the selected Landsat scenes were acquired on different days within the seasons of late summer to late fall. Using images from the same day per year does not guarantee detection of urban areas due to annual variations in atmospheric conditions, including large annual variations in cloud contamination.

These last two problems may be solved using two methods either together or separately. Though, these methods should be explored and tested prior to application with VSU. Regarding the spectral library, the endmember library could be simulated and sampled in order to account for radiometric and temporally-caused illumination differences between the endmembers and each Landsat scene. For example, endmember responses may be simulated to correspond to a response seen from space. One such method is to convolve each endmember with the relative spectral response of the Landsat sensor and then convolve the resulting signal with a signal for atmospheric interference. This method, however, is experimental, requires sampling (such as at the wavelength where the relative spectral response of the sensor equals or approaches one), and would require knowledge of atmospheric conditions for the day and location of the region imaged by Landsat. The time series of Landsat scenes could be radiometrically normalized to account for differences in daily and annual illumination and for disparities in annual atmospheric conditions, but the radiometrically normalized scenes would likely still contain at least marginal spectral distinction from the library. However, detection and classification of urban areas would likely improve if these two methods or a similar

strategy were employed to minimize the spectral variance between the time series of Landsat scenes and the spectral library. Another means of improving urban detection and classification may include using multiple images from different seasons. Winter images lacking snow show increased urban areas due to lack of leafy coverage. Urban areas classified in winter imagery could be used to correct misclassified urban areas in maps generated from scenes acquired in other seasons of the same year.

In the last locale, each year shows more water coverage than occurs in the corresponding NLCD images. This can occur for at least two reasons. These areas lie on the coast and consist of wetlands, which may be covered with water periodically throughout a single year due to either or both drainage and oceanic tides. As a result, the mapping algorithm may have picked up on standing water and classified it as open water accordingly. Additionally, the use of the MNDWI with the threshold of 0.0 may have caused more pixels to be extracted automatically as water. As a result of these errors, the agreements between each mapped location and the corresponding NLCD are low, as seen in Table 2.7. In the first locale, the agreement corresponds predominantly to the areas of vegetation belonging to the types of forest, shrubs, and grasses and the agreement in the last locale occurs primarily due to water coverage. However, misclassification due to the hierarchal classification rules also persist in the low agreements.

Table 2.7 - Agreements with the NLCD

Locale Image Coordinates	[2100,1900]	[3620,1060]	[5450,4650]
NLCD 2001	33.76%	2.03%	20.69%
NLCD 2006	27.72%	1.73%	18.53%
NLCD 2011	32.35%	3.10%	39.64%

According to the NLCD classification criteria in Table 2.7, all classes represent mixtures of physical surface characteristics with varying percentages. The exact limits for every coverage type are not explicitly stated except for the coverage predominant per class. For example, the predominant type in deciduous forest are deciduous trees and the criteria does not specify the limits for other coverage types, such as bare land, artificial materials, coniferous trees, gramanoids like grass, and other materials, occurring within the same land area. This constitutes the mixed pixel problem. In the case of the third locale, vegetation occurring in emergent herbaceous wetlands (95 — dark blue as seen in the maps) includes both herbaceous and gramanoid species, which are shared with the grassland and shrub land. A few of the herbaceous species in the spectral library are herbaceous scrubs and contribute spectrally both to grassland and shrub land. As a result of the material type and the hierarchal classification system, the wetlands were misclassified as grassland in third locale. Similar effects are seen in the maps produced for the first and second locales.

Thus, altering the hierarchal classification rules may improve classification results. The easiest fix may be to raise the classification rule for wetlands to a higher level in the rule structure, but this may lead to confusion with other classes below it in the hierarchy. The hierarchal classification rules could be replaced with rules based on explicit mixtures

of material types occurring within each NLCD class. These mixtures may be generated through analysis of the VSU output of the endmember ids and material types with their corresponding fractional weights to establish limits on the cumulative weights by material type occurring for each class in the NLCD. Classification rules could be structured based on these limits to further guide and improve the accuracy of classification into NLCD classes. If VSU is used with other applications requiring classification, a similar approach may be used with the target classification scheme. Additionally, implementing such mixture rules as a decision tree would assist in minimizing error due to a single hierarchy structure. Though, there are other means of classification, including maximal likelihood, fuzzy, or score-based systems, which could be adapted to the NLCD criteria and used in place of a hierarchal system.

Overall, the results of VSU are very promising. Despite agreements less than 40%, the VSU method produced land cover maps with classes corresponding to physical material types with cumulative weights meeting the conditions within and the order of the hierarchal classification system as opposed to classes representing land cover classes with mixed physical material types, but these results may be improved to result in mixed classes corresponding to physical coverage. In comparison to the NLCD method, this method is less involved. The NLCD methodology requires images from multiple seasons, days, and years; ancillary datasets; at least a moderately sized group of personnel; and multiple years to create one accurate land cover dataset for a single year. This method using VSU requires a single person, one image per year mapped, and one workstation. It is fast, autonomous and adaptive based on the requirements per application. If improved to the point where its

accuracy is at least 75% for mapping land cover, the method using VSU would rival the NLCD methodology.

3. Land Cover Change Modeling

3.1 Section Overview

Land cover change modeling attempts to predict the future state of land cover in response to various stimuli, such as environmental factors, previous states, etc. This second contribution of this thesis is concerned with alterations in land cover due to differences or changes in surrounding land cover and aims to determine, if any, the trends in land cover change due to the neighborhood of a pixel in the previous state after a transition of the target pixel to a future state. In this research, the Moore and von Neumann cell neighborhoods of a time series of land cover maps are analyzed to determine land cover change trends due to neighboring land cover and derived rules derived are used to model changes in classes using the NLCD 2001 as a starting point. In analyzing a time series of maps and modeling land cover change using derived decision rules, this part of the thesis aims to demonstrate how modeling using neighborhoods and previous states as criteria for change may produce good forecasts of future land cover with an accuracy of at least 50%.

Note, this portion of the thesis relies on data created using the mapping algorithm developed in Section 2, but this section specifically deals with the development of a method for deriving decision rules from a time series and their implementation. Any mentions of the mapping algorithm are to indicate potential errors in deriving land cover change decision rules and are intended to discuss ways to improve the detection of changes in land cover due to neighborhoods rather than to directly improving the mapping algorithm, which is discussed in depth in Sections 2.3 and 4.1. The following sections address the

methodology used in land cover change modeling and analysis prior to delving into modeling results.

3.2 Methodology

This study will examine land cover change trends as a function of changes in neighboring land cover. In the proposed model, a two-dimensional array represents a large land surface composed of pixels with a square spatial resolution of 30 meters and annual changes to land cover result from changes in its spatial neighborhood, such as the Moore or von Neumann neighborhoods of a given pixel. The spatial orientation of the land cover classes of the neighborhoods is assumed not to contribute to the change in a target pixel.

As the NLCD contains datasets only for four years over 19 years, this study begins by generating a land cover dataset according to the method outlined in Section 2 of this thesis. The VSU algorithm generated maps for three locations for the years 2001 through 2011 for a total of 11 years. The locations were chosen after examining the land cover change between the NLCD 2001 and 2006, the 2006 and 2011, and the 2001 and 2011 datasets after calculating the difference between each dataset.

The maps generated using VSU tend to be strong in pure land cover classes (such as forest, grassland, and shrub land) and weak in the other land cover classes. As discussed in Section 2.3, urban and agricultural areas suffer low propagation in the VSU-generated images due to urban areas occurring with low compositional weights that do not survive post-classification and a lack of pasture/hay and crop endmembers. Wetlands occur due to mixtures of vegetation types and water drainage systems that were not accounted for in the

classification algorithm used with VSU. Thus, prior to neighborhood analysis, the VSU-generated maps need to be corrected in order to derive decision rules for these classes important to human life. The VSU-generated maps are validated against and corrected using an ancillary land cover dataset known as the Coastal Change Analysis Program (C-CAP), which the National Oceanic and Atmospheric Administration (NOAA) created for the years 2001, 2006, and 2011, as it uses similar land cover classes to the USGS NLCD. The NOAA C-CAP uses the same core land cover classes and criteria as the NLCD except for the cases of wetlands. In the NOAA C-CAP, additional land cover classes break out the wetlands into different types with respect to mixed environments corresponding to marine and grassland, shrub / scrub land, and forest vegetation.

As a result, the NLCD-equivalent codes of the C-CAP land cover classes replace the values of pixels in the VSU-generated maps that spatially correspond to the same pixels that belong to the urban, agricultural, and wetlands classes in the C-CAP. For example, if a pixel is classified as deciduous forest in the VSU-generated map and as low density developed in the C-CAP, then the pixel in the VSU-generated map is reclassified as low density developed. Similarly, if a pixel is classified as grassland in the VSU-generated map and as woody wetlands or cultivated crops in the C-CAP, then the pixel in the VSU-generated maps is reclassified as woody wetlands or cultivated crops. This reclassification is directly applicable only to the years for which there exist maps in both datasets. As a result, pixels replaced in the datasets for 2001 and 2006 are held constant for the time frames 2001 through 2005 and 2006 through 2010. This causes the urban, agricultural, and wetlands classes to experience quinquennial changes. These classes do experience change in the C-

CAP in the selected locales and result in identification of some trends in these classes.

Otherwise, no decision rules would result for the classes of developed, pasture / hay, cultivated crops, woody wetlands, and emergent herbaceous wetlands.

The analysis of the Moore and von Neumann cell neighborhoods assume a change in the land cover class per year occurs due to the land cover classes in its neighborhood in the past. The neighborhood analysis starts with year 2002 and proceeds through 2011 using each past year as the previous state resulting in the change to the next year, if any. In order to simplify and speed the neighborhood analysis, the maps corresponding to the current and prior years are differenced and only pixels with a non-zero difference are analyzed. Additionally, analysis of the Moore and von Neumann neighborhoods excludes the outermost ring of land cover in a map and starts at the pixel on the diagonal from the upper leftmost pixel. The Moore and von Neumann neighborhoods are defined as follows in Eq. 3.1 and 3.2, respectively:

$$N_M^1(i, j) = \{\sigma_{i,j}, \sigma_{i-1,j}, \sigma_{i-1,j-1}, \sigma_{i,j-1}, \sigma_{i+1,j-1}, \sigma_{i+1,j}, \sigma_{i+1,j+1}, \sigma_{i-1,j+1}\} \quad (3.1)$$

$$N_N^1(i, j) = \{\sigma_{i,j}, \sigma_{i-1,j}, \sigma_{i,j-1}, \sigma_{i+1,j}, \sigma_{i,j+1}\} \quad (3.2)$$

,where N_M^1 and N_N^1 are the Moore and von Neumann neighborhoods of radius one unit of pixel (i,j), respectively, and σ is the state at coordinates around pixel (i,j) [9]. Note, in a Moore neighborhood, the use of a radius of one unit around a target pixel assumes surrounding pixels at distances greater than 30 meters (the spatial resolution of Landsat

images) do not contribute to the annual change of the target. In the case of the von Neumann neighborhood, the use of a radius of one unit around a target pixel assumes surrounding pixels that are not orthogonal to the target or are at distances greater than 30 meters do not contribute to the annual change of the target.

During analysis, the previous state, the future state, and the neighborhood of the previous state are recorded per locale per change in year. Then, the locale data is recorded as a single set per change in year. In this thesis, annual changes in land cover are considered independent of the spatial orientation of the neighborhood of land cover classes. Thus, two neighborhoods are equivalent if they include the same land cover classes with equivalent amounts per class in each neighborhood. Due to this, prior to accumulative counting, the neighborhoods are converted to an array of the frequency of occurrence of each land cover class in the neighborhood for ease of counting. The extracted neighborhoods are counted to determine the total number of occurrences per neighborhood per possible transition per change in year. Using this data, rules are constructed in the form of if-else statements using the previous state and the neighboring land cover classes as the criteria for change in a basic reiterative model. Separate models were constructed for both neighborhoods and start with a large section of the NLCD 2001. Using the derived rules, both models were run from 2001 through 2011 with the intermittent maps recorded for comparison with the corresponding NLCD if their locales are included within the larger sample location of the model.

3.3 Results & Discussion

3.3.1 Rationale of Transition Exclusions

Analysis of the Moore and von Neumann neighborhoods of validated mapped pixels revealed several trends of interest, but some transitions were excluded. General nonsensical transitions may include transitions of land cover classes recorded during analysis that may arise due to mapping or validation errors, such as the case of a developed pixel surrounded by water pixels, and transitions that are not thought to occur naturally, such as downward transitions of high density developed (24) to deciduous forest (41). Additionally, neighborhoods in which all neighbors in a Moore neighborhood for transitions where the previous state does not equal the future state belong to the same class as the previous state i are excluded. This thesis assumes a neighborhood with eight members of the same class leads to continuance of the same state. Furthermore,

In the case of transitions from the developed classes to any of the forest, shrub land, woody wetlands, and cultivated crops, these transitions are not considered due to the amount of time needed to change urban areas into these heavily vegetated classes. For example, according to the NLCD criteria given in Table 2.6, forests and shrub land consist of plants greater than and less than 5 meters tall, respectively. Vegetation of such height is assumed to take more than one year to develop. Similarly, time is needed to change developed areas into barren land capable of sustaining mass agricultural activities. Those types of transitions will be considered and pursued in the future with repetition of this same or a similar methodology as applied to other regions and a larger time span. The

same rationale is used for excluding transitions of the type from barren land to any of the forest classes, including woody wetlands, and shrub land.

3.3.2 Results of Neighborhood Analysis

Tables 3.1 and 3.2 show the selected results and their frequencies, f , of occurrence for the von Neumann and Moore neighborhoods, respectively, for each transition from state i to state j given in each table. These neighborhoods were included in the model as decision rules. The results generally occur within the top five per encountered transitions for the von Neumann and Moore neighborhoods. Where resulting neighborhoods were few in number or the frequencies of occurrence were low in number (i.e., less than 15), the neighborhoods were selected based on their constituents.

In Tables 3.1 and 3.2, the frequencies of occurrence tend to be high for transitions between two vegetated classes and low for transitions between urban classes and between vegetated and urban classes. These cases result from the choice in mapped locations and the mapping correction. First, the chosen locations featured more vegetative classes as compared to urban classes and the vegetated pixels appear to remain vegetated temporally. Pixels of developed classes appear to remain stable temporally as well. Thus, fewer transitions appear for transitions from vegetated to developed classes and for transitions between developed classes, including the agricultural land use classes. Mapping additional urban locations and larger areas of developed land, such an area of developed land equal to the area of vegetated classes, would improve detection of urban transitions in further research using this method. Secondly, the mapping correction assumes no change in urban areas for years from 2001 to 2005 and from 2006 to 2010. Thus, very limited to moderate

levels of annual change in developed classes were expected. This study employed mapping correction in order to generate at least a few cellular automata rules for the developed and other land use classes. As for errors due to the mapping procedure, naturally, improving the classification scheme of the VSU output would improve detection of feature classes and, therefore, detection of change between land cover classes. Strategies for improving the VSU-mapping procedure are discussed in Section 2.3 and Section 4.1.

The neighborhoods seen in Tables 3.1 and 3.2 show patterns in the frequency of their components, most notably those with higher frequencies of occurrence. In the von Neumann results in Table 3.1, transitions from deciduous forest (41) to coniferous forest (42), from coniferous forest (42) to deciduous forest (41), from coniferous forest (42) to shrub land (52), and from woody wetlands (90) to coniferous forest (42) in addition to others have the same patterns. In the one case, the neighborhood of the target pixel comprises of three neighbors with the same class as the target and one neighbor corresponding to the future class. In the second case, the neighborhood consists of two neighbors of the same class of the target pixel and two neighbors of the future class.

Similar patterns arise in the Moore results in Table 3.2. There are many transitions resulting from neighborhoods of four cases: (1) where there are seven neighbors of the same class as that of the target pixel and one neighbor of the future class; (2) where there are six members of the class of the target pixel and two members of the future class; (3) where five members belong to the class of the target pixel and three members belong to the future class; and (4) where equal numbers of members belong to both the previous and future classes. Examples of these three cases include transitions from open space

developed (21) to pasture / hay (81), coniferous forest (42) to shrub land (52), coniferous forest (42) to woody wetlands (90), grassland (71) to deciduous forest (41), woody wetlands (90) to cultivated cropland (82), and emergent herbaceous wetlands (95) to open water (11).

These patterns suggest general rules based on such cases for both von Neumann and Moore neighborhoods may be written in these formats between any two classes. Though, this observation requires further exploration to be verified, such as through repetitions of the same method used in this thesis for other or larger locations. Based on the results in Table 3.1, these trends appear to become distinguishable with large frequencies of occurrence. Each type also needs to be further explored and tested beyond the scope of this thesis for accuracy and resulting rates in changes of cell counts per year for each class in a time series. In this thesis, all transitions in Table 3.1 and 3.2 were modeled.

There are exceptions to these patterns, including but not limited to such as transitions deciduous forest (41) to woody wetlands (90) and grassland (71) to shrub land (52) in Table 3.1 and open space developed (21) to barren land (31), low density developed (22) to 21, medium density developed (23) to high density developed (24), and 52 to 71 in Table 3.2, that contain neighborhoods with more than two classes. These generally occur with frequencies of occurrence less than 100, except in the case of the transition of 71 to 52; neighborhoods occurring with frequencies equal to one generally occurred in a list where all neighborhoods were encountered once in the entire time series of the corrected VSU-generated maps. These results may not be encountered in repetitions of this same basic methodology over larger sample areas where such transitions are encountered on an

annual basis. It should not be immediately concluded, however, that neighborhoods of two or more classes other than the future and previous states do not contribute to a change. All patterns discussed in this section need to be verified by repetition of this basic methodology for a much larger sample area over a longer time span and for additional states in the US.

Table 3.1 - Von Neumann Neighborhoods with Greatest Frequency of Occurrence Per Transition

<i>i</i>	<i>j</i>	<i>f</i>	Neighborhood	<i>i</i>	<i>j</i>	<i>f</i>	Neighborhood
21	22	1	[21, 21, 21, 21]	71	42	114	[71, 71, 42, 41]
		1	[21, 22, 22, 22]			80	[71, 71, 71, 42]
21	23	1	[22, 22, 22, 23]	71	52	625	[71, 71, 71, 52]
		1	[21, 21, 21, 22]			266	[71, 71, 52, 41]
22	21	1	[22, 22, 22, 23]	71	81	17	[71, 71, 81, 81]
		1	[21, 21, 21, 22]			16	[81, 71, 71, 71]
22	23	3	[21, 21, 22, 22]	71	82	5	[71, 71, 82, 82]
		3	[22, 22, 22, 23]			4	[82, 82, 71, 22]
22	24	3	[22, 22, 22, 22]	71	90	16	[71, 71, 90, 90]
		3	[22, 22, 23, 23]			15	[71, 71, 71, 90]
23	24	3	[23, 23, 24, 24]	81	21	12	[81, 81, 81, 22]
		2	[23, 24, 24, 24]			4	[81, 81, 81, 21]
31	71	61	[31, 31, 31, 71]	81	22	5	[81, 81, 81, 22]
		42	[31, 31, 71, 71]			2	[81, 81, 22, 22]
31	81	2	[31, 71, 81, 81]	81	41	115	[90, 81, 81, 81]
		1	[31, 31, 22, 81]			71	[81, 81, 22, 22]
41	42	3120	[41, 41, 41, 42]	81	42	218	[81, 81, 81, 42]
		3096	[41, 41, 42, 42]			121	[81, 81, 42, 42]
41	52	758	[41, 41, 41, 52]	81	71	162	[81, 81, 81, 71]
		434	[41, 41, 52, 52]			13	[81, 81, 71, 71]
41	71	545	[41, 41, 41, 71]	81	82	56	[81, 81, 81, 82]
		511	[41, 41, 71, 71]			20	[81, 81, 82, 82]
41	90	38	[41, 41, 90, 90]	81	95	5	[11, 81, 81, 81]
		33	[41, 41, 42, 90]	82	71	32	[82, 82, 82, 81]
42	41	2732	[41, 42, 42, 42]	82	81	39	[81, 81, 81, 81]
		1434	[41, 41, 42, 42]			3	[82, 82, 81, 71]
42	52	1921	[42, 42, 42, 52]	82	90	56	[82, 82, 82, 90]
		649	[42, 42, 52, 52]			26	[82, 82, 90, 90]
42	71	63	[42, 42, 42, 71]	90	11	4	[95, 95, 11, 11]
		18	[42, 42, 71, 71]			3	[11, 90, 90, 90]
42	81	2	[42, 42, 82, 82]*	90	41	130	[41, 90, 90, 90]
		2	[42, 42, 42, 82]*			71	[41, 41, 90, 90]
42	90	810	[42, 42, 42, 90]	90	42	794	[42, 90, 90, 90]
		537	[42, 42, 90, 90]			654	[42, 42, 90, 90]
43	90	115	[43, 43, 43, 95]	90	81	7	[81, 90, 90, 90]
		79	[43, 43, 95, 95]			7	[81, 81, 90, 90]

52	41	799	[52, 52, 52, 41]	90	82	2	[82, 82, 90, 90]
		793	[52, 52, 41, 41]	90	95	17	[71, 90, 90, 90]
52	42	3903	[42, 42, 42, 52]			5	[90, 90, 90, 95]
		2814	[42, 42, 52, 52]	95	11	1297	[11, 95, 95, 95]
52	71	231	[71, 52, 41, 41]			1005	[11, 11, 95, 95]
		208	[71, 52, 52, 41]	95	90	24	[90, 95, 95, 95]
71	41	918	[71, 71, 41, 41]			1	[41, 95, 95, 95]
		1654	[71, 71, 71, 41]			1	[42, 42, 95, 95]

* Modeled as 81 prior to correction to 82.

Table 3.2 - Moore Neighborhoods with Greatest Frequency of Occurrence Per Transition

i	j	f	Neighborhood	i	j	f	Neighborhood
21	22	1	[21, 21, 21, 21, 22, 22, 22, 22]	42	90	387	[42, 42, 42, 42, 42, 90, 90, 90]
		1	[21, 21, 21, 22, 22, 22, 22, 22]			355	[42, 42, 42, 42, 42, 42, 90, 90]
21	23	1	[21, 21, 22, 41, 81, 81, 81, 81]	43	42	43	[42, 42, 42, 42, 43, 43, 43, 43]
		1	[21, 21, 21, 81, 81, 81, 81, 81]			31	[42, 42, 42, 43, 43, 43, 43, 43]
21	31	1	[21, 22, 22, 22, 22, 23, 24, 24]	43	90	65	[43, 43, 43, 43, 43, 95, 95, 95]*
		1	[21, 22, 22, 81, 81, 81, 81, 81]			74	[43, 43, 43, 43, 43, 43, 43, 95]*
21	81	7	[21, 21, 21, 21, 21, 21, 81, 81]	52	31	2	[41, 71, 71, 71, 71, 71, 71, 71]
		6	[21, 21, 81, 81, 81, 81, 81, 81]			4	[41, 41, 42, 42, 42, 71, 90, 90]
22	21	1	[21, 21, 21, 21, 21, 21, 22, 22]	52	41	228	[41, 41, 41, 52, 52, 52, 52, 52]
		1	[21, 22, 22, 22, 23, 41, 81, 81]			234	[41, 41, 52, 52, 52, 52, 52, 52]
22	23	2	[22, 22, 22, 22, 22, 22, 23]	52	42	1313	[42, 42, 42, 42, 42, 52, 52, 52]
		1	[21, 21, 21, 21, 22, 22, 22, 22]			1911	[42, 42, 42, 42, 42, 42, 52, 52]
22	24	2	[22, 22, 22, 22, 22, 23, 23, 24]	52	43	7	[52, 52, 95, 95, 95, 95, 95, 95]*
		2	[21, 22, 81, 81, 81, 81, 81, 81]			3	[52, 52, 52, 52, 95, 95, 95, 95]*
22	31	1	[22, 22, 22, 22, 23, 23, 23, 23]	52	71	71	[41, 41, 41, 52, 71, 71, 71, 71]
23	22	2	[23, 23, 23, 81, 81, 81, 81, 81]			66	[41, 41, 52, 52, 71, 71, 71, 71]
		6	[22, 22, 22, 22, 22, 23, 23, 23]	52	81	3	[41, 41, 52, 52, 52, 71, 71, 71]
23	24	1	[21, 21, 22, 22, 22, 24, 24, 81]			2	[41, 41, 41, 52, 52, 52, 52, 71]
		2	[22, 22, 22, 23, 23, 23, 23, 23]	71	31	8	[41, 52, 71, 71, 71, 71, 71, 71]
24	22	2	[22, 22, 22, 23, 23, 23, 24, 24]			7	[41, 52, 52, 71, 71, 71, 71, 71]
		1	[22, 22, 23, 23, 24, 81, 81, 81]	71	41	538	[41, 41, 71, 71, 71, 71, 71, 71]
24	23	1	[22, 22, 23, 23, 24, 81, 81, 81]			448	[41, 41, 41, 71, 71, 71, 71, 71]
		1	[21, 23, 23, 23, 23, 23, 23, 24]	71	42	53	[41, 41, 52, 71, 71, 71, 71, 71]
31	21	1	[23, 31, 52, 71, 71, 81, 81, 81]			42	[41, 42, 42, 42, 71, 71, 71, 71]
		1	[31, 71, 71, 71, 71, 81, 90, 90]	71	52	131	[41, 52, 71, 71, 71, 71, 71, 71]
31	71	25	[31, 31, 31, 31, 31, 71, 71, 71]			13	[41, 42, 42, 42, 71, 71, 71, 71]
		21	[31, 31, 31, 31, 31, 31, 71, 71]	71	81	17	[71, 71, 71, 71, 71, 71, 71, 81]
31	81	2	[31, 71, 81, 81, 81, 81, 81, 81]			7	[71, 71, 71, 81, 81, 81, 81, 81]
		1	[22, 22, 31, 31, 71, 81, 81, 81]	71	82	2	[71, 71, 71, 81, 82, 82, 82, 82]
41	21	2	[22, 22, 22, 71, 81, 81, 81, 81]			3	[71, 71, 71, 81, 81, 81, 81, 81]
		1	[21, 21, 21, 22, 22, 41, 43, 43]	71	90	8	[71, 81, 81, 81, 90, 90, 90, 90]
41	22	1	[22, 22, 22, 22, 41, 71, 71, 71]			6	[71, 71, 71, 71, 71, 71, 90, 90]
		1	[22, 22, 23, 41, 41, 52, 52, 90]	81	21	7	[21, 21, 21, 81, 81, 81, 81, 81]
41	31	6	[41, 41, 41, 41, 41, 42, 42, 52]			3	[21, 21, 22, 81, 81, 81, 81, 81]
		4	[41, 41, 41, 41, 41, 41, 41, 71]	81	42	88	[42, 42, 42, 81, 81, 81, 81, 81]
41	42	1198	[41, 41, 41, 41, 41, 42, 42, 42]			78	[42, 42, 81, 81, 81, 81, 81, 81]

	1108	[41, 41, 41, 41, 42, 42, 42, 42]	81	82	15	[81, 81, 81, 81, 82, 82, 82, 82]	
41	43	3	[42, 43, 43, 43, 43, 43, 52, 52]		17	[81, 81, 81, 81, 81, 81, 82, 82]	
		1	[43, 43, 43, 90, 90, 90, 90, 90]	82	81	6	[81, 81, 81, 82, 82, 82, 82, 82]
41	52	233	[41, 41, 41, 41, 41, 41, 52, 52]			4	[81, 81, 82, 82, 82, 82, 82, 82]
		293	[41, 41, 41, 41, 41, 41, 41, 52]	82	90	13	[82, 82, 82, 82, 82, 82, 90, 90]
41	71	175	[41, 41, 71, 71, 71, 71, 71, 71]			18	[82, 82, 82, 82, 90, 90, 90, 90]
		147	[41, 41, 41, 71, 71, 71, 71, 71]	90	11	3	[11, 11, 11, 90, 90, 90, 90, 90]
41	81	5	[41, 41, 41, 41, 41, 71, 71, 71]			2	[11, 11, 11, 11, 95, 95, 95, 95]
		3	[41, 41, 41, 41, 41, 42, 52, 52]	90	41	39	[41, 41, 90, 90, 90, 90, 90, 90]
41	82	2	[41, 41, 81, 81, 81, 90, 90, 90]			26	[41, 41, 41, 90, 90, 90, 90, 90]
		2	[41, 81, 81, 81, 90, 90, 90, 90]	90	42	335	[42, 42, 42, 90, 90, 90, 90, 90]
41	90	24	[41, 41, 42, 42, 42, 42, 42, 42]			330	[42, 42, 42, 42, 90, 90, 90, 90]
		13	[41, 41, 41, 42, 42, 42, 42, 90]	90	71	31	[81, 81, 90, 90, 90, 90, 90, 90]
42	21	4	[21, 42, 42, 90, 90, 90, 90, 90]			23	[71, 71, 90, 90, 90, 90, 90, 90]
		4	[21, 42, 90, 90, 90, 90, 90, 90]	90	81	5	[81, 81, 81, 90, 90, 90, 90, 90]
42	41	822	[41, 41, 42, 42, 42, 42, 42, 42]			3	[81, 81, 81, 81, 90, 90, 90, 90]
		820	[41, 41, 41, 42, 42, 42, 42, 42]			3	[81, 81, 81, 81, 90, 90, 90, 90]
42	43	23	[42, 42, 42, 42, 42, 42, 43, 43]	90	82	2	[82, 82, 82, 90, 90, 90, 90, 90]
		4	[42, 42, 43, 43, 43, 81, 81, 81]			1	[82, 82, 82, 82, 82, 90, 90, 90]
42	52	732	[42, 42, 42, 42, 42, 42, 52, 52]	95	11	679	[11, 11, 11, 95, 95, 95, 95, 95]
		436	[42, 42, 42, 42, 42, 52, 52, 52]			672	[11, 11, 95, 95, 95, 95, 95, 95]

* Values of 95 were mistyped as 90 in the model prior to correction of the model.

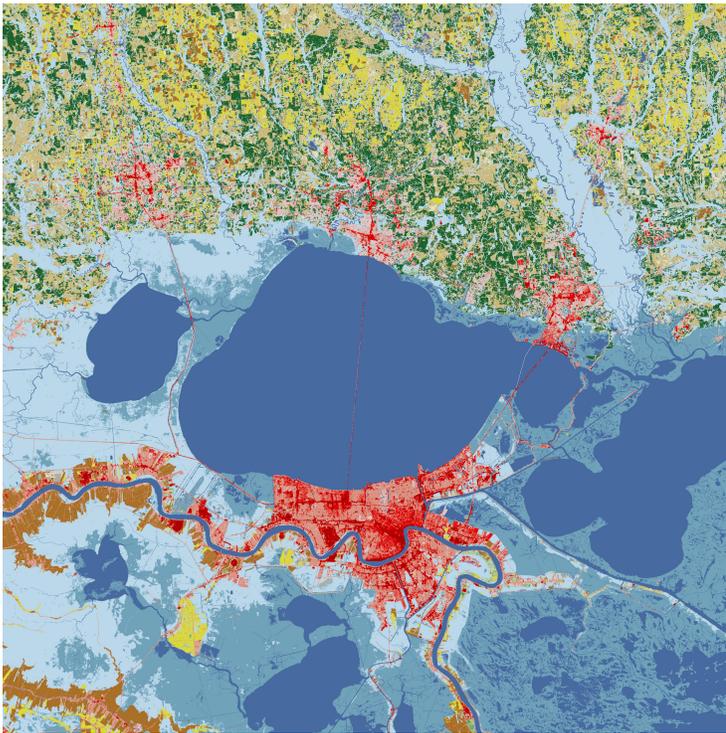
3.3.3 Modeling Results

Two models were created using the neighborhoods given in Tables 3.1 and 3.2. The decision rules were written as if-else statements using the previous state of the target pixel and the count of each class included within the given neighborhood of the target pixel as the criteria to matching it to any of the selected neighborhoods of the corresponding transitions. If the previous state and the neighborhood of the target pixel in the previous state do not match these criteria for change, then the pixel remains in the same state — the same land cover class as the previous state. These rules were separated into two models according to the type of neighborhood used in deriving the rules: the von Neumann and Moore models. All neighborhoods per transition as seen in Tables 3.1 and 3.2 were included into each corresponding model. Each model started on year 2001 using the NLCD 2001 and ran for 11 iterations through year 2011. Modeled results of year 2011 are compared to the NLCD 2011 for validation in the same fashion.

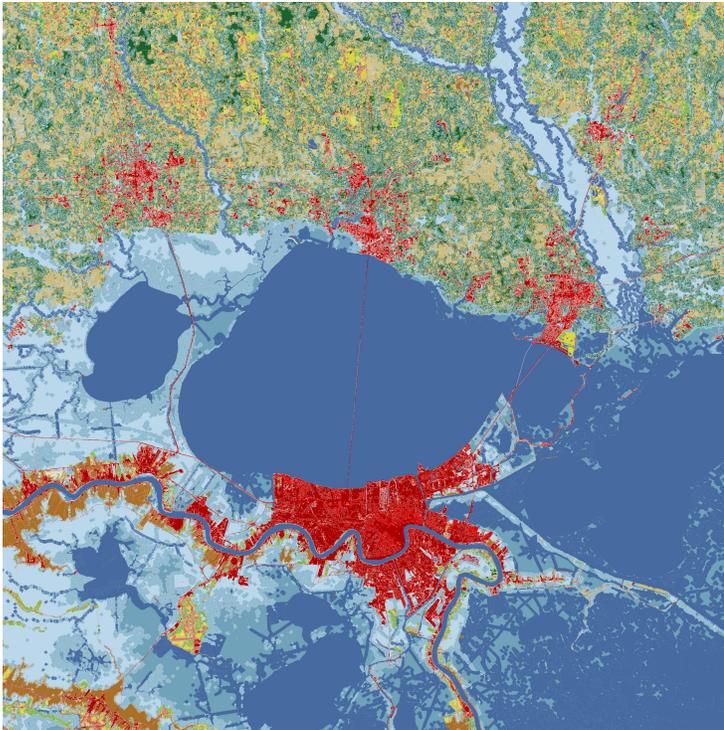
The chosen location corresponds to a slice of the NLCD 2011 covering the image coordinates of [2000:6000, 2000:6000] and features New Orleans. The following forecast maps use the color scheme used in in the NLCD 2011 legend, as seen in Figure 2.1.

Von Neumann Forecasting

Figure 3.1 shows the NLCD 2011 (A) clipped to the same coordinates used in the model and the forecast of 2011, as captured by the rules using the von Neumann neighborhoods in Table 3.1. The von Neumann model relies on a neighborhood defined by Eq. (3.2). Immediately noticeable is the greater amount of the open water land cover in the forecast than in the NLCD 2011 and the pixelated appearance of the wetlands and other vegetated classes. On the other hand, the collective rules produced similar visual patterns in the developed classes and pasture / hay. Straight lines corresponding to artificial structures such as roadways and bridges are largely preserved in the forecast image.



(A)



(B)

Figure 3.1 - Von Neumann Forecasting Results vs. NLCD 2011: (A) NLCD 2011; (B) 2011 forecasted results of the von Neumann model.

Table 3.3 shows numerical results comparing the forecasted map of 2011 and the NLCD 2011. Trends visual in Figure 3.1 are shown by data in Table 3.3 For instance, emergent herbaceous and woody wetlands decreased in cell counts from 2006 to 2011. Similarly, medium and high intensity developed classes experience an increase in cell counts, as occurs in the NLCD 2011. The rates of change in the cell counts per year are largely similar between the the von Neumann forecast and the NLCD, as shown in Table 3.3 with many of the classes have the same signs indicating growth and decline. Some classes do not follow this patter, such as open space developed (21) and shrub land (52). This may occur due to annual changes in neighborhoods permitting transitions. This is further implicated by percent differences which are different between the 2006 and 2011 comparisons, such as the percent differences for classes open space developed (21), barren

land (31), deciduous forest (41), etc. In particular, classes barren land (31) and deciduous forest (41) experience very large variance in their percent differences. Behavior such as this also indicates the large changes incurred by the von Neumann neighborhoods in addition to using two neighborhoods per rule in general. The von Neumann neighborhood consists of only four neighbors and the lower number of components for comparison leads to an increase in the inaccuracy of the forecasting results. These observations may be validated by replicating the underlying methodology of this thesis for larger areas or more areas interspersed around the globe.

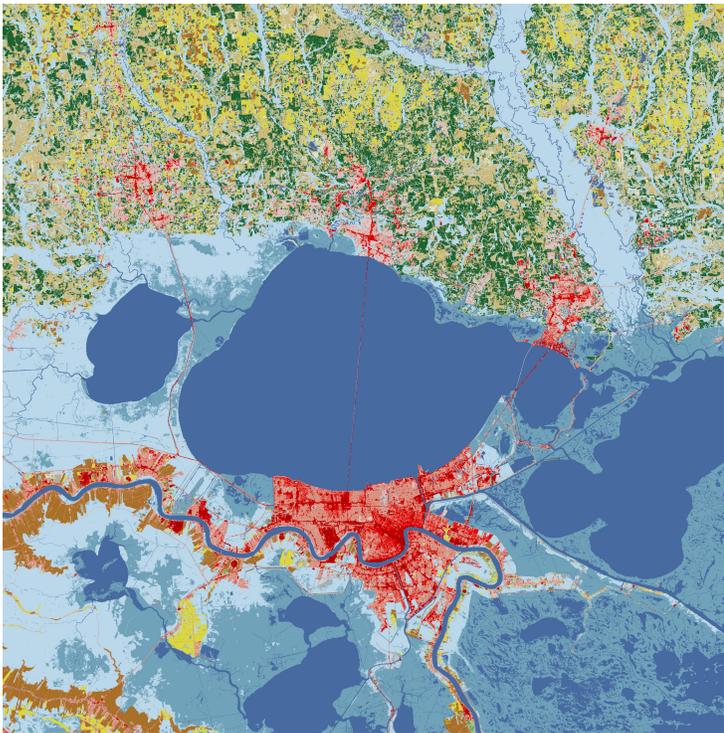
Overall, the model performed satisfactorily. The 2011 von Neumann land cover forecast agrees approximately 66.38% with the NLCD 2011 with over 10 million matched pixels. This lower accuracy is not surprising when compared to use of the Moore neighborhood method.

Table 3.3 - Comparison of von Neumann Forecasts with NLCD - The abbreviation (VN) corresponds to results of the von Neumann model.

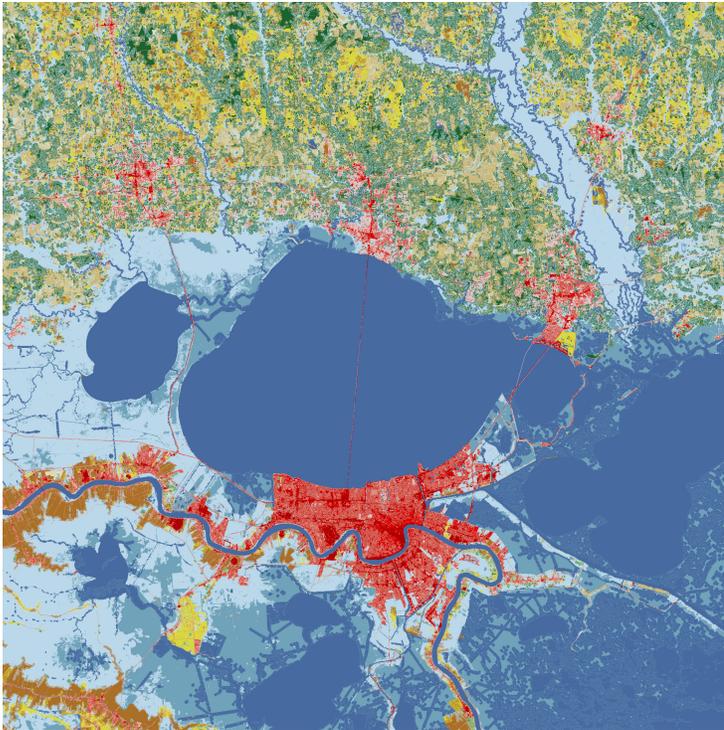
Class	VN '06 Counts	NLCD '06 Counts	% Difference	VN '11 Counts	NLCD '11 Counts	% Difference	VN Slope	NLCD Slope
11	5,214,502	4,164,424	-25.22	5,735,395	4,280,815	-25.36	104,178.6	23,378.2
21	523,652	541,206	3.24	537,860	533,816	-0.75	2,841.6	-1,478
22	283,341	657,902	56.93	282,017	655,162	132.31	-264.8	-548
23	271,602	210,095	-29.28	279,119	226,229	-18.95	1,503.4	3,226.8
24	521,48	112,136	-365.05	526,699	122,585	-76.73	1,042	2,089.8
31	27,677	32,247	14.17	26,235	77,695	196.15	-288.4	9,089.6
41	153,125	17,787	-760.88	206,233	16,999	-91.76	10,619.6	-157.6
42	1,391,513	1,455,258	4.38	1,315,191	1,285,433	-2.26	-15,264.4	-33,965
43	62,703	77,554	19.14	61,388	72,386	17.92	-263	-1,031.6
52	1,262,427	1,156,212	-9.19	1,348,726	1,374,366	-18.91	17,259.8	43,630.8
71	219,336	256,527	14.50	237,462	192,553	-18.91	3,625.2	-12,794.8
81	440,581	587,672	25.03	388,822	561,526	44.42	-10,351.8	-5,229.2
82	340,560	413,360	17.62	298,758	400,980	34.22	-8360.4	-2,476
90	3,358,376	3,622,337	7.29	3,161,296	3,610,859	14.22	-39,416	-2,295.6
95	1,929,116	2,695,293	28.43	1,594,809	2,588,596	62.32	-66,861.4	-21,339.4

Moore Forecasting

Figure 3.2 shows (A) the NLCD 2011 and (B) the 2011 forecasting results of the Moore model. The classification color legend is shown in Figure 2.1. Noticeable differences exist in the visual results. The forecast for 2011 shows much degradation in emergent herbaceous wetlands (95) and corresponding large growth in water, suggesting the rules for emergent herbaceous wetlands and other classes with transitions to water classes should be limited to one neighborhood. These results are further demonstrated in Table 3.4. Vegetated classes of pasture / hay, cultivated crops, shrub land, and woody wetlands look largely similar whereas urban areas show more areas of high intensity developed.



(A)



(B)

Figure 3.2 - Moore 2011 Forecasting vs. NLCD 2011: (A) NLCD 2011; (B) 2011 forecasting results of the Moore model.

Table 3.4 shows the numerical results of comparisons between the forecasted land cover of 2011 using the Moore model and the actual NLCD 2011. This table provides the cell counts per class for the modeled years of 2006 and 2011, the percent difference between each dataset per year, and the slopes of the two datasets for the duration from 2006 to 2011. As seen in Table 3.4, the modeling results vary by class and year. The differences in cell counts vary by year and class. For 2006, the smallest percent difference between the Moore and NLCD class counts occur for classes open space developed (21), coniferous forest (42), shrub land (52), pasture / hay (81), cultivated cropland (82), and woody wetlands (90) while the largest percent differences occur for classes barren land (31) and deciduous forest (41). The data for 2011 show the smallest percent differences occur for same classes, except 52, and the largest percent differences occur for classes 41

and emergent herbaceous wetlands (95). These results seem to suggest that alterations in cell counts per class may change annually. From a modeling perspective, this may occur due to annual changes in cell neighborhoods per year. Further modeling is required to fully explore this.

These results show that the Moore rules perform differently by class as well. Although all changes in neighborhoods by year may be partially to blame, these results suggest some of the transition rules are ineffective for the classes with the largest percent differences, namely 41 deciduous forest, and those with large variance between the cell counts per year, such as 31 and 95. These rules must be further tested to determine which perform best. Additionally, these rules should be validated individually and then tested in a model using only one neighborhood per decision rule, as using two neighborhoods per rule may account for the increased percent differences of these classes. After all, not all transitions were modeled and some transitions have more rules than others, as seen in Table 3.2.

The slopes were calculated as the difference between the cell counts per 2006 and 2011 over a span of 5 years. The closest slopes correspond to classes medium intensity developed (23), mixed forest (43), and pasture / hay (81), indicating similar patterns of growth between the forecast and the NLCD 2011, which supports previous statements. Similarly, the slopes farthest apart occur for classes 11, 22, 31, and etc., including classes with high percent differences, and indicates poor performance.

Table 3.4 - Comparison of Moore Forecasts with NLCD - The prefixes M06, N06, M11, and N11 correspond to abbreviations of the terms Moore 2006, NLCD 2006, Moore 2011, and NLCD 2011 to indicate results of the results of the predictions for 2006 and 2011 by the Moore model and the NLCD 2006 and 2011, respectively.

Class	M06 Counts	N06 Counts	% Difference	M11 Counts	N11 Counts	% Difference	M Slope (counts / year)	N Slope (counts / year)
11	4,999,382	4,164,424	-20.05	5,467,872	4,280,815	-21.71	93,698	23,278.2
21	515,871	541,206	4.68	513,974	533,816	3.86	-379.4	-1,478.0
22	532,711	541,206	19.03	498,591	655,162	31.40	-6,824	3,226.8
23	279,079	210,095	-32.83	296,229	226,229	-23.82	3,575	3,226.8
24	137,070	112,136	-22.24	142,847	122,585	-14.18	1,155.4	2,089.8
31	72,423	32,247	-124.59	84,465	77,695	-8.01	2,408.4	-9,089.6
41	34,130	17,787	-91.88	47,990	16,999	-64.58	2,772	-157.6
42	1,398,856	1,455,258	3.88	1,354,479	1,285,433	-5.10	-8,875.4	-33,965
43	109,977	77,544	-41.83	113,490	72,386	-36.22	702.6	-1031.6
52	1,163,289	1,156,212	-0.61	1,233,956	1,374,366	11.38	14,133.4	43,630.8
71	223,642	256,527	12.82	246,076	192,553	-21.75	4,486.8	-12,794.8
81	604,548	587,672	-2.97	583,347	561,526	-3.74	-4,240.2	-5,229.2
82	417,517	413,360	-1.01	415,649	400,980	-3.53	-373.6	-2476
90	3,591,269	3,622,337	0.86	3,502,997	3,610,859	3.08	-17,654.4	-2,295.6
95	1,920,236	2,695,293	28.76	1,497,313	1,354,912	72.88	-84,584.6	-21,339.4

Overall, the model performs well as compared to the NLCD 2011. The land cover forecast for 2011 contains over 12 million matched pixels, equating to a percent agreement of approximately 76.85% with the NLCD 2011. This is greater than the agreement of the von Neumann forecast for land cover in 2011 with the NLCD. This is expected, considering how the Moore rules use a greater number of values for comparison of neighborhoods. Despite using a mash-up of neighborhoods, the Moore model performs well with a good agreement of approximately 77%, demonstrating how decision rules based on Moore neighborhoods and the method of deriving those rules may be used to develop an accurate model.

4. Conclusion & Future Work

4.1 Land Cover Mapping Conclusions & Applications

In general, the mapping results showed predominantly vegetation classes of forest, shrub land, and grassland. The state of Louisiana hosts large amounts of vegetation on its land surface, promoting the spectral dominance of the endmembers of these classes. This trumps the detection of urban areas, causing the vegetation endmembers contributing to a reflectance measurement to have large fractional compositions compared to those of urban endmembers. Thus, the classification algorithm chooses the vegetation classes over the urban classes due to the high cumulative sums of the vegetation types. This may be remedied using different classification methods as adapted for the NLCD classification scheme. In the future, classification methods such as maximum likelihood, fuzzy, or a score-based classification should be tested with this method of spectral mixture analysis.

Another source of error is the spectral similarity between the analyzed Landsat scene and the spectral library. The spectral library was sampled where the relative spectral response of each Landsat band approached or equaled unity. This attempted to heighten the spectral similarity between Landsat 5 scenes and the spectral endmembers of urban materials, because it resulted in urban endmembers occurring in the sub-pixel maps. Obviously, these sub-pixels classified as urban materials did not have large enough compositional weights to be classified as urban classes. There are two approaches that should be tested with the mapping method incorporating VSU. First, in the preprocessing of each Landsat scene, techniques for computing surface reflectance should be tested with

this method. Secondly, the spectral library could be resampled to full-width half maximum and convolved with the relative spectral response of the Landsat 5 TM sensor and adjusted for simulated atmosphere to simulate at-satellite reflectance. This latter method can be challenging due to lack of atmospheric data for each day on which a Landsat scene was acquired.

Another potential source of error that can be tested in the future includes the use of the MNDWI to extract water, as it appears to lead to added water being extracted. The generally used threshold of the MNDWI is 0.0. However, various researchers suggest adjusting the threshold based on lighting conditions between images in a time series. Increasing the threshold to 0.2 or 0.3 decreases the amount of pixels carried over as noise but also decreases the number of pixels extracted as water. The MNDWI could be eliminated as a means of quickly extracting water features, but eliminating this step would be accompanied by an increase in processing time. Further analysis into the effects of thresholding the MNDWI on the mapping results is needed to determine on which route to proceed.

Although there are significant sources of error in the procedure, this mapping method using VSU demonstrates it is feasible to map land cover using spectral mixture analysis techniques, provided the right classification system is used. In small-scale testing, the single-pass-per-pixel VSU algorithm analyzed a sample size of 100 by 100 with 6 bands of data in approximately 500 seconds on average with only two criteria used in spectral matching. In comparison, when implemented in Python, the VMESMA algorithm — without reiteration through the subset of selected endmembers per pixel to minimize the root mean

square error of the least-squares solution — with three criteria generally requires 630 seconds on average to analyze the same sample size. When accounting for the reiterative-per-pixel nature of VMESMA, the temporal difference between these two performance durations is expected to increase. The method using VSU should be further analyzed for practical applications, including land cover mapping as attempted in this thesis.

Spectral mixture analysis techniques are applicable to mapping many types of physical surface features, including sea ice, snow, forest and other dominant vegetation types, minerals and other earth types, algae, water, urban materials, and other substances. For example, mapping minerals and other soils may be used to identify lands suitable for mining and farming. Mapping forests and other vegetation in addition to minerals permits monitoring of natural environments, such as the amount of land they cover, expanse of fires or burn damage, or mineral runoff into waterways near mining or drilling operations. Such applications may then be used to investigate other phenomena. For example, mapping land cover and then differencing annual land cover datasets leads to the mapping of land cover change.

The applications of mapping physical surface features other than land cover using VSU are nearly boundless. Such products may be used in decision making processes, such as in the cases of identifying where to drill for desired earthen substances or where to plant to yield the most crops in relation to a viable water source. As with any method, errors appearing during mapping lead to propagation of errors in further data products, which can mislead stakeholders in decision making.

Consider a mining company looking for surface characteristics signifying the

abundance of a valuable mineral (such as oil, coal, or diamonds). The mining company acquires maps created using a new method with inherent errors and the maps contain false detection of the surface characteristics signifying a great abundance of their target mineral. Based on their maps, the company spends millions to mine or drill the indicated site only to find their site contains limited quantities of their mineral of interest. As a result, the company experiences no return on their investment and the map makers lose business. Such errors, which are fixable in the method prior to application, cause large losses financially and temporally and must be avoided in the real world wherever possible. Therefore, errors in VSU and classification must be minimized prior to implementation for any application. Similarly, the errors encountered in this method require correction prior to large scale application to land cover mapping.

4.2 LCCM Conclusions

In Section 3, land cover changes were modeled using decision rules derived from analysis of the Moore and von Neumann neighborhoods of a time series of maps covering a single decade. The results of the neighborhood analysis identified general cases for decision rules for both von Neumann and Moore neighborhoods. These general neighborhoods typically included neighbors belonging only to the classes equivalent to the previous state of the target pixel and the future class with differing numbers of the two classes, forming a general ratio by type. As these general cases often possess high frequencies of occurrence, this suggests general rules may be written in the format of these general neighborhoods for any pair of classes. This observation requires further exploration and testing prior to actual application. Future research could approach this task by using the methods covered in this

thesis to map larger locations interspersed throughout the US, or even around the globe. Future research could also determine if these rules are specific only to the areas studied in this thesis or are compatible with only similar regions (i.e. environmentally, geographically, etc.).

The von Neumann land cover forecast for 2011 agreed with the NLCD 2011 by approximately 66%, a satisfactory result for this model and its general incorporation of decision rules constructed from von Neumann neighborhoods. The forecast showed significant degradation of emergent herbaceous wetlands and growth in high intensity developed areas, indicating potentially overly prolific decision rules. Other land use classes of pasture / hay and cultivated crops experienced less growth or decline, similar to a few other vegetated classes.

The Moore land cover forecast for 2011 produced an agreement of approximately 77%. Such a moderately high agreement demonstrates the efficacy of decision rules using a Moore neighborhood. The forecast also showed large decline in cell counts of emergent herbaceous wetlands with a corresponding growth in open water cell counts, similar to the von Neumann model. Other land use classes experienced patterns of growth and decline similar to those shown in the NLCD. Though, the percent differences between cell counts of the Moore model were often less than those resulting in the von Neumann model when comparing their forecasts to the NLCD 2006 and 2011.

Both the von Neumann and Moore models performed satisfactorily, but the Moore forecast clearly out performs the von Neumann model. This is expected due to the larger number of elements in the Moore neighborhood, increasing the accuracy of correctly

transitioning during each iteration of the Moore model as compared to the von Neumann model. Both models displayed similar errors of cell count trends and of spatial agreement of cells in the forecast and NLCD maps per year in forecasting land cover changes for particular land classes, including emergent herbaceous wetlands. In both models, the trends of growth of open water and of decline of emergent herbaceous wetlands were more exaggerated than the trends of both classes encountered in the NLCD, as shown by the lesser loss of emergent herbaceous wetlands cell counts in the NLCD. Such behavior indicates the inadequacy of the decision rules used for those classes and may be improved in the future by mapping more areas where these land cover classes occur and change in relation to other classes. Additionally, patterns of error in the measures of percent difference between cell counts per class by year may indicate that the neighborhoods causing annual land cover changes may themselves change annually or with a different periodicity. This is further stipulated by changes in the NLCD cell counts from 2006 and 2011. These observations require further validation for confirmation.

The agreements between the forecast and NLCD maps for the year 2011 reflect the spatial agreement between the cells of the two maps per year mapped. The agreements of 66% and 77% of the von Neumann and Moore forecasts, respectively, with the NLCD 2011 demonstrate that 34% and 23% of the cells in the von Neumann and Moore forecasts, respectively, do not contain the same land cover class code as the NLCD 2011. For example, there are areas of coniferous forest in the NLCD 2011 which are mixed forest in either or both of the forecasts and urban areas could possess a land cover code higher than an urban area in the NLCD 2011. Such errors between similar classes as this must be taken into

account for calculating the agreement between the two datasets.

This method of deriving rules from classified imagery for implementation in a forecast model possesses wide applications, including those mentioned in the previous section, but the following brief discussion will be applied to land cover change. Forecasting may be used as a tool for decision making processes in such applications as wetlands restoration. Consider the case of companies in or near New Orleans, Louisiana, that possess facilities and employ local citizens of the area. As discussed in the introduction of this thesis, Louisiana may suffer increased damages from hurricanes in the future due to wetlands loss and degradation. Local companies could restore wetlands at the cost of thousands or millions of dollars in order to protect their facilities and their employees, but they need to know which potential locations of restoration would best minimize their future damages and the costs of restoration. Forecasting future wetlands degradation using a rigorous model produce results that, when coupled with other analyses, may pinpoint areas best suited for restoration while minimizing future hurricane-related damage and the cost of restoration.

Consider the same scenario using the forecasts given in this thesis, which featured accelerated degradation of emergent herbaceous wetlands. The accelerated trend would mislead a company as to future degradation and cause the company to invest too heavily in wetlands restoration, potentially minimizing future damages while maximizing cost. If a decelerated trend were encountered, the opposite could occur and the company would invest too little in wetlands restoration, potentially increasing future damages and minimizing cost. Such errors are not permissible in a forecast to be used in such a large

scale endeavor. Thus, prior to application, the neighborhoods and decision rules derived in this thesis require further validation and testing prior to large scale application.

The satisfactory and good agreements resulting from the von Neumann and Moore forecasts, respectively, demonstrate how land cover change may be simulated with at moderately high accuracy using neighborhoods rather than stochastic models.

Furthermore, these results demonstrate how land cover trends may be approximated by deriving decision rules from analysis of neighborhoods of pixels in a time series of maps.

These results could be improved by mapping other locations where these land cover classes intermix and a greater time interval should be attempted, as forecasting results are known to improve with increases in sample sizes.

A. Bibliography

- [1] de Almeida, C. M., M. Batty, A. M. V. Monteiro, G. Camara, B. S. Soares-Filho, G. C. Cerqueira, and C. L. Pennachin. "Stochastic cellular automata modeling of urban land use dynamics: empirical development and estimation." *Computers, Environment, and Urban Systems*. Vol. 27, No. 5 (2003): 481-509. Web. 10 July 2015.
- [2] Baldrige, A.M., S.J. Hook, C.I. Grove, and G. Rivera. "The ASTER spectral library version 2.0." *Remote Sensing of Environment* 1 13 (2009): 711-715. Web. 10 Sep. 2014.
- [3] Batty, Mathew. *Cities and Complexity: Understanding Cities with Cellular Automata, Agent-Based Models, and Fractals*. The MIT Press, 2007. Print.
- [4] Clark, R. N., et al. *USGS Digital Spectral Library splib06a*. Denver: U.S. Geological Survey, Data Series 231, 2007.
- [5] Crooks, A., C. Castle, and M. Batty. "Key challenges in agent-based modeling for geo-spatial simulation." *Computers, Environment, and Urban Systems*. Vol. 32, No. 6 (2008): 417-430. Web. 10 July 2015.
- [6] Fry, J., G. Xian, S. Jin, J. Dewitz, C. Homer, L. Yang, C. Barnes, N. Herold, and J. Wickham. "Completion of the 2006 National Land Cover Database for the Conterminous United States." *Photogrammetric Engineering and Remote Sensing*. Vol. 77. No. 9 (2011): 858 - 864. Web. 5 May 2015.
- [7] Gao, B., "NDWI - A Normalized Difference Water Index for Remote Sensing of Vegetation Liquid Water from Space." *Remote Sensing of Environment*. Vol. 58 (1996): 257-266. Web. 5 May 2015.
- [8] Garcia-Haro, F.J., S. Sommer, and T. Kemper. "A new tool for variable multiple endmember spectral mixture analysis (VMESMA)." *International Journal of Remote Sensing*. Vol. 26(10) (2005): 2135-2162. Web. 1 Apr. 2015.
- [9] Hoekstra, A. G., J. Kroc, and P. M. A. Sloot (2010). "Introduction to Modeling Complex Systems Using Cellular Automata." Ed. Hoekstra, A. G., J. Kroc, and P. M. A. Sloot. *Simulating Complex Systems by Cellular Automata*. New York: Springer, 2010. Print.
- [10] Homer, C., J. Dewitz, J. Fry, M. Coan, N. Hossain, C. Larson, N. Herold, A. McKerrow, J.N. VanDriel, and J. Wickham. "Completion of the 2001 National Land Cover Database for the Conterminous United States." *Photogrammetric Engineering and Remote Sensing*. Vol. 73, No. 4 (2007): 337 - 341. Web. 5 May 2015.
- [11] Homer, C.G., J.A. Dewitz, L. Yang, S. Jin, P. Danielson, G. Xian, J. Coulston, N.D. Herold, J.D. Wickham, and K. Megown. "Completion of the 2011 National Land Cover Database for the conterminous United States - Representing a decade of land cover change information." *Photogrammetric Engineering and Remote Sensing*. Vol. 81, No. 5 (2015): 345-354. Web. 5 May 2015.
- [12] Huete, A.R. "A Soil-Adjusted Vegetation Index (SAVI)." *Remote Sensing of Environment*. Vol. 25 (1988): 295-309. Web. 5 May 2015.
- [13] de Jong, Steven M. And Freek D. Van der Meer. *Remote Sensing Image Analysis: Including the Spatial Domain*. Vol. 5. USA: Springer Science+Business Media Dordrecht, 2004.
- [14] Lillesand, T. M., R. W. Kiefer, and J. W. Chipman. *Remote Sensing and Image Interpretation*. 6th ed. Hoboken: John Wiley & Sons, Inc., 2008. Print.
- [15] Louisiana. Coastal Protection and Restoration Authority of Louisiana (CPRAL). *Louisiana's Comprehensive Master Plan for a Sustainable Coast (LCMPSC)*. Baton Rouge: CPRAL, 2015. Web. 7 July 2015.
- [16] "Mean Sea Level Trend 8761724 Grand Isle, Louisiana." *NOAA Tides & Currents*. National Oceanic and Atmospheric Administration (NOAA), 10 October 2013.

- Web. 10 July 2015.
- [17] van der Meek, Freek D. And Steven M. De Jong. *Imaging Spectroscopy: Basic Principles & Prospective Applications*. Vol. 4: Dordrecht: Springer Science+Business Media Dordrecht, 2001.
- [18] National Aeronautics and Space Administration (NASA). *Landsat 7 Scientific Users Handbook*. Greenbelt: NASA, 2011. Web.
- [19] Roberts, D.A., M. Gardner, R. Church, S. Ustin, G. Scheer, and R.O. Green. "Mapping Chaparral in the Santa Monica Mountains Using Multiple Endmember Spectral Mixture Models." *Remote Sensing of Environment*. Vol. 65 (1998): 267 - 279. Web. 1 Apr. 2015.
- [20] Settle, J. J. And N. A. Drake. "Linear mixing and the estimation of ground cover proportions." *International Journal of Remote Sensing*. Vol. 14, No. 6 (1993): 1159-1177. Web. 5 Jul. 2015.
- [21] United States Geological Survey (USGS). *MRLC 2001 Image Processing Procedure*. Sioux Falls: USGS, 2006.
- [22] Xu, Hanqiu. "Modification of normalised difference water index (NDWI) to Enhance open water features in remotely sensed imagery." *International Journal of Remote Sensing*. Vol. 27(14) (2006): 3025 - 3033. Web. 5 May 2015.
- [23] Zha, Y., J. Gao, and S. Ni. "Use of noralized difference built-up index in automatically mapping urban areas from TM imagery." *International Journal of Remote Sensing*. Vol. 24(3) (2003): 583-594. Web. 5 May 2015.