2019

Interpreting Vegetation and Soil Anomalies in the Guarumen Area of Northwestern Venezuela Using Remote Sensing Applications

Tyler Edward Ricketts
University of Mississippi

INTERPRETING VEGETATION AND SOIL ANOMALIES IN THE GUARUMEN AREA OF NORTHWESTERN VENEZUELA USING REMOTE SENSING APPLICATIONS

A Thesis

Presented in partial fulfillment of requirements

For the degree of Master of Science

In the Department of Geology and Geological Engineering

The University of Mississippi

by

TYLER E. RICKETTS

December 2018

ABSTRACT

The Guarumen area of Venezuela is a tectonically active region that is approximately 1,640 mi^2 across the northern portions of the Barinas Basin and the foothills of the Mérida Andes. It is structurally influenced by the Caribbean plate to the north, the Nazca plate to the west, and the Maracaibo block against the Guyana Shield of the South American Plate. These result in an oblique boundary that gives rise to the fold-and-thrust belt of the Mérida Andes to the west, and the Caribbean Mountain system to the north, in concordance to the right-lateral shearing that is evidenced by the Boconó fault system. The goal of this research was to investigate the geological setting of northwestern Venezuela and further understand the geologic controls of the region, as it has become a region of interest for mineral, oil, and gas exploration. To achieve the goal, hyperspectral and multispectral data analysis were used to address land cover types by reducing hyperspectral and multispectral spectra to unique endmembers for use in classification. Then, provide an accurate land cover analysis using derived endmembers to characterize the outcomes concerning the influence of geological phenomena, and determine if microclimate analysis using satellite-based land surface temperature data can be effectively used to infer geologic structure or geomorphology, particularly soils and vegetation. Based on the hyperspectral data, an in-depth endmember analysis was conducted with image-derived spectra. These spectra were plotted in comparison with spectral libraries to identify the anomaly classification. It was determined that the natural vegetation make up of a specific region helped identify soil type. The Guarumen area was influenced by the sediment transport of the alluvial stream geomorphology of both the Merida Andes and the Caribbean Mountain System and both its respective geologies. The

microclimate analysis showed a land surface temperature comparison of two separate Landsat scenes. Both showed a similar mean temperature range due to Venezuela's tropical climate, but differed in other classifications. Results from this research show that remote sensing applications with limited field data can provide accurate land cover analysis concerning geological phenomena, but further field analysis is needed for more detailed classification.

DEDICATION

This thesis is dedicated to my family and the love of my life for always supporting me with unconditional love, encouragement, and prayers throughout this long rewarding process.

ACKNOWLEDGMENTS

I would like to thank my committee, Dr. Lance Yarbrough, Dr. Greg Easson, and Dr. Brian Platt for all the endless support they gave me throughout this research. In addition, I would like to thank Mississippi Mineral Resource Institute and America View for funding this research and making this thesis possible. I would also like to thank SpecTIR, PDVSA, and the people of Venezuela who all helped and supported this research throughout the preliminary processes. Lastly, I would like to thank all of the professors, friends, and staff within the University of Mississippi Department of Geology and Geological Engineering for their support and motivation throughout the years.

Tyler Ricketts

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Airborne acquisition of hyperspectral imagery has been used in the field of geology and geospatial applications since the early 1980's (Kruse et al., 2003). Hyperspectral applications offer substantial advantages over traditional remotely used sensors. This includes the increased spectral resolution that allows for more precise identification of specific vegetation indices, soil properties, minerals, and rock groups. Geological and vegetation applications archetypally use the bandwidth of 1–20 nm and, depending on its sensor, spectral windows from 64–400 different bands spanning from the visible-near infrared (VIR), and the short-wave infrared (SWIR) portions of the electromagnetic spectrum depending on the sensor (Harris et al., 2005). With this capability, hyperspectral image analysis can provide high-resolution spectra, along with spatial data. These spectra are used to determine necessary signatures from the spectral data provided by the specific sensor (Martinez et al., 2006). With the hyperspectral imagery and its more defined pixel data, hyperspectral applications may improve the discrimination of multiple components, particularly numerous vegetation types, even with low coverage (Jafari & Lewis, 2012).

With countries facing political and civil unrest, conducting field-based geological research is limited due to unsafe conditions for the researchers. During civil unrest, remote sensing becomes an integral tool of the research conducted on a specific field area considered a high region of interest. In the case of this research, Venezuela has become politically dangerous, and an economic threat for any continuous research on the ground. Due to the cancellation of the

field campaign, a detailed geologic background is important. Since fieldwork cannot be conducted, this makes previous field studies an integral part of this research.

The goal of this research was to investigate the geological setting of northwestern Venezuela and further the understanding of the region because it has become a region of interest for mineral, oil, and gas exploration. To achieve my goal, I used remotely sensed imagery and spectral data processing tools to address the following objectives: 1) given limited field data and in a variety of land cover types, reduce the hyperspectral and multispectral spectra to unique endmembers for use in classification; 2) provide an accurate land cover analysis using derived endmembers and characterize the outcomes with respect to influence of geological phenomena; and 3) determine if microclimate analysis using satellite-based land surface temperature data can be effectively used to infer geologic structure or geomorphology, particularly soils and vegetation.

When working with hyperspectral imagery, not all the bands used during the acquisition phase are useful in an analysis. Some hyperspectral bands contain excessive noise, while some select bands contain important responses of interest for specific research (Rasti, et al., 2018). A systematic approach to identifying these bands of interest is to take into account as many of the variables that may affect the region of interest, and for the user to understand that different materials respond at different bands (Martinez et al., 2006). Endmembers are specific hyperspectral signatures, or a specific response, i.e., reflectance, in those given wavelengths that are collected by the sensor that is seen as macroscopically pure (Bateson et al., 2000). These endmembers are extracted from the large amount of spectral data using a variety of different extraction techniques. These endmembers are then visualized to determine their vector based on their spectral composition. To conceptualize the endmember concept, the “spectrum

composition" model (Figure 1) (Martinez et al., 2006) demonstrates the endmember in the n-dimensional vector.

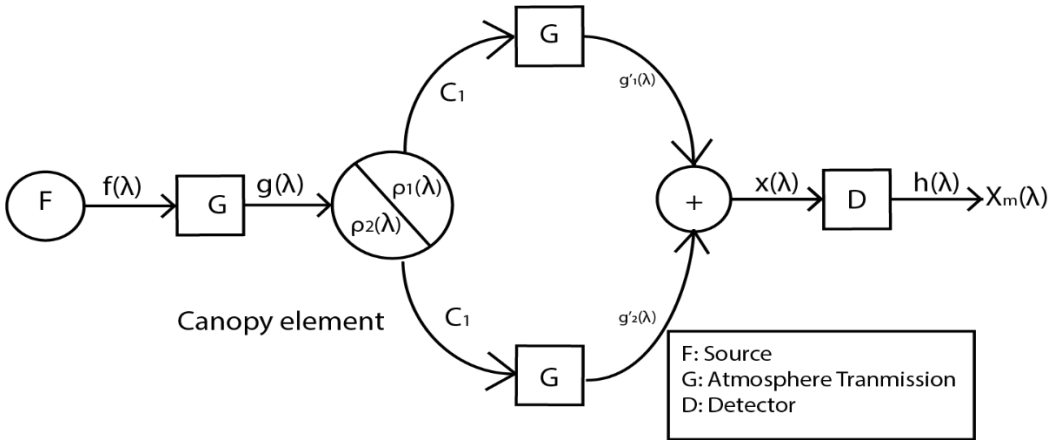

FIGURE 1. Spectrum Composition Model modified from Martinez et al., (2006).

The general workflow for extracting endmembers is in the form of 1) data preprocessing, 2) atmospheric correction of data to apparent reflectance, 3) determining data dimensionality to minimize noise of the reflectance data through linear transformation, 4) pixel purity index (Boardman et al., 1995), or N-FINDR (Winter, 1999) to locate the most spectrally pure pixels, 5) automated or manual extraction identification of endmember spectra, and 6) spatial mapping and specific endmember abundance estimations (Kruse et al., 2003). The main takeaway from this workflow is the minimization of data in both spectral and spatial dimensions to identify the fundamental spectra that can be used to analyze the hyperspectral data. Once the selection process is complete, the endmembers location and abundances can be mapped from the linearly transformed or the initial data (Kruse et al., 2003).

Hyperspectral and multispectral sensors differ in a variety of ways. Multispectral sensors collect data from three to six bands from the visible and near-IR region of the electromagnetic

spectrum. This is considered a crude spectral categorization of the emitted and reflected energy from the Earth, which is the primary limiting factor of multispectral sensors (Govender et al., 2007). Even though multispectral sensors are still used as an integral tool for remote sensing applications in various fields, hyperspectral remote sensing sensors collect several, narrow spectral bands from the visible (380–700 nm), near-IR (700 nm–1 mm), and short-wave infrared (1–2.5 mm) portions of the electromagnetic spectrum. These hyperspectral sensors collect 200 or more spectral bands, which allows the construction of an almost continuous spectral reflectance signature. With these narrow bandwidths, hyperspectral data allow a thorough examination of the Earth surface and its features which would be lost with the coarse bandwidths of multispectral data (Govender et al., 2008).

GEOLOGIC BACKGROUND

Regional Setting

The study site is a region in northwestern Venezuela just south of the city of Barquisimeto (10° 03′52″N 69° 20′1″E) (Figure 2) known as the Guarumen area. The site is a tectonically active region that is approximately 427,000 hectares (1,640 mi^2) across the northern portions of the Barinas Basin and the foothills of the Mérida Andes. It is structurally influenced by the Caribbean plate to the north, and the Nazca place to the west, as well as by the Maracaibo block against the Guyana Shield of the South American Plate. These result in an oblique boundary that gives rise to the fold-and-thrust belt of the Mérida Andes, in concordance to the right-lateral shearing that is evidenced by the Boconó fault system. The majority of the study refers to the Llanos region, which is geomorphologically classified as a Quaternary alluvial plain that borders the southeastern foothills of the Mérida Andes.

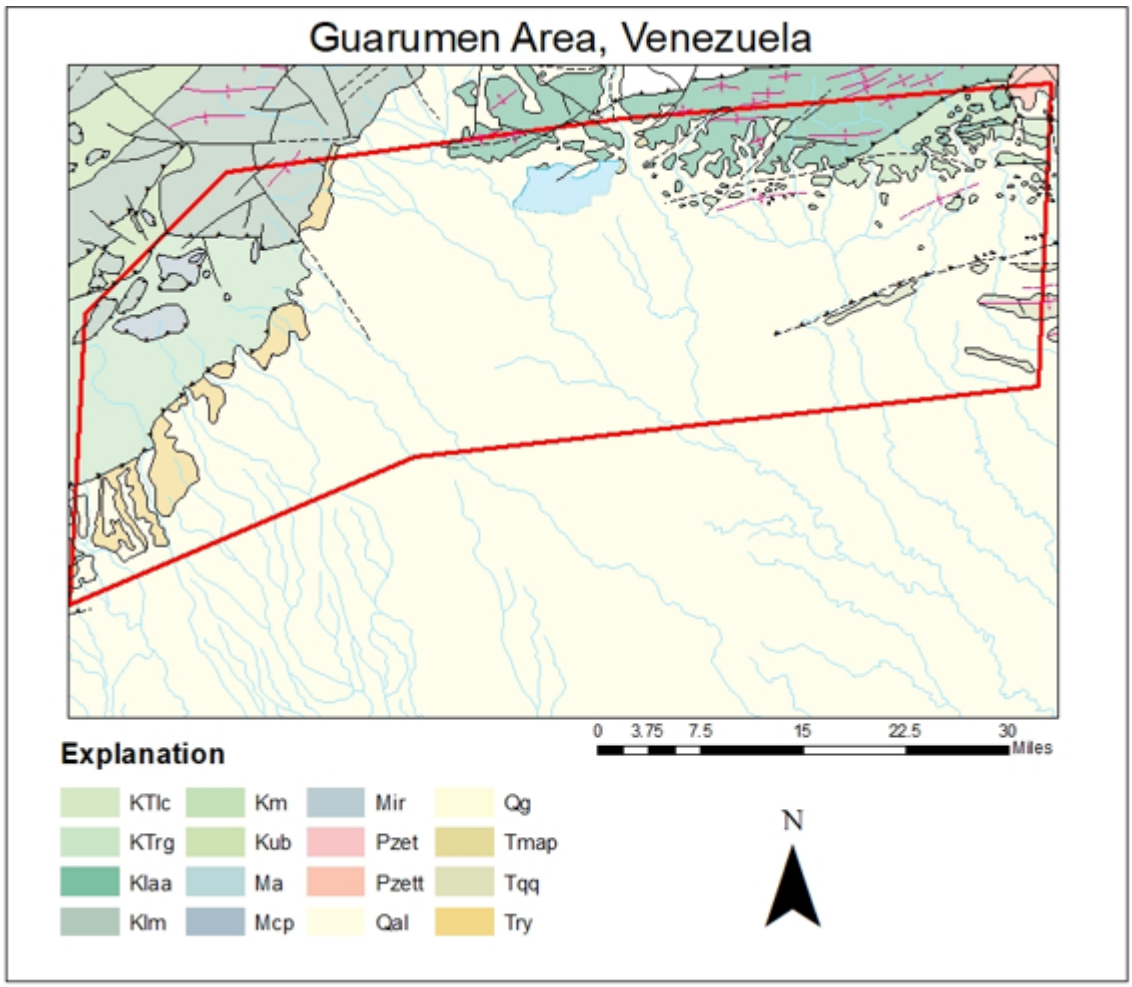

FIGURE 2: Geologic map showing structural and hydrologic features.

EON	ERA	PERIOD	FORMATION DESCRIPTIONS	Geology
Phanerozoic	Cenozoic	Quaternary	Alluvium	Sand, Gravel, Silt
			Guanapa Formation	Conglomerate, Sandstone, Claystone
		Tertiary	Rio Yuca Formation	Conglomerate, Sandstone
			Maporita Formation	Conglomerate, Sandstone, Limestone, Claystone, Marl
			Quebradón, Quiamare, and Naricual Formations	Shale, Coal, Siltstone, Sandstone
	Mesozoic	Cretaceous	Guárico Formation	Marl, Calcareous Shale, Mudstone
			Rio Guache Formation	Marl, Calcareous Shale, Mudstone
			Mucaria Formation	Carbonaceous Shale
			Barquisimeto Formation	Shale, Siltstone, Limestone
			Agua Blanca, Araure, Cojedes Formations	Limestone, Conglomerate
			Mamey, Yacambú, Colcancito Formations	Quartz Schist, Metaconglomerate, Shale, Metasandstone, Quartzite
		Jurassic	Cerro Pelón Ultramafics	Ultramafic Rocks
		Triassic		
	Paleozoic	Permian	El Tinaco Complex	Paragneiss, Amphibolite, Marble, Trondhjemite
		Pennsylvanian		

FIGURE 3: Stratigraphic column of Guarumen Area.

Tectonic Basin History

Venezuela lies on the northern margin of South America facing the Caribbean Sea. The Caribbean Plate and South American Plate are the two major plates influencing the tectonics of this region. The South American plate, moving in a western direction, interacts with the Nazca Plate. The Nazca Plate is subducting underneath the western edge of the South American Plate which causes the collision of these two plates to be responsible for the uplift of the Andes mountains, along with transform and thrust faulting of the region (Deiros, 2000; and Pérez et al.,

2001). The Caribbean Plate is translated eastward approximately 22 mm per year in comparison to the neighboring South American Plate (Deiros, 2000; and Pérez et. al., 2001), while the Nazca plate is moving eastward at approximately 60 mm/yr, and subducting underneath Columbia (Freymueller et al., 1993). Most of the tectonic movement within Venezuela occurs between the Caribbean and South American plate and falls along the Boconó, El Pilar, and San Sebastián fault zones (Audemard et al., 1996).

The Precambrian Guayana Shield of Venezuela is rimmed by multiple foreland basins that all deepen in the direction of the Mérida Andes and the interior and coastal ranges. All of these basins contain Cretaceous and younger sediments above local rifts with Paleozoic and Jurassic sedimentary rocks (James, 2000). The basement arch, the El Baúl Uplift, separates the Barinas-Apure basin from the Eastern Venezuela Basin in the East, as seen in Figure 3. Close to the Colombian border in the west, the subsurface Arauca Arch separates the Barinas-Apure Basin from the Llanos Basin. The Eastern Venezuela Basin is crossed by the Anaco Fault that divides into the Guárico Sub-basin that lies in the west, and the Maturín Sub-basin to the east. (James, 2000).

FIGURE 4: Basement map of Venezuela (Ruiz & Adana, 2013)

Barinas-Apure Basin

The Barinas-Apure Basin is geographically located in the northwestern part of Venezuela that is southeast of the Mérida Andes. It is the third largest oil producing basin in Venezuela with an aerial extent of about 95,000 km^2 and a maximum sediment depth of approximately 5,000 m (Jácome & Chacín, 2008). To the west of the basin, the Mérida Andes, a NE-SW trending mountain chain, contains a large northeastern uplifted block approximately 420 km in length that reaches an elevation of about 5000 m (Gonzalez de Juan et al., 1980). This uplift is the result of Barinas-Apure Basin subsidence. However, the history of the subsidence in this basin in relation to tectonic evolution has never been quantitatively studied (Chacín et al., 2005). The Paleogene sedimentation in the Barinas Basin is missing lower Eocene and Paleocene strata. Rather, it begins with the upper and middle Eocene diagenesis of several formations grading from neritic shales to fluvial sandstones that establishes a single transgressive-regressive cycle (Pierce, 1960; Furrer, 1971; Osuna et al., 1995; Erikson et al., 2012; and Rincó et. al., 2016). Above the Barinas

Basins regional unconformity lie less than 500 m of Cretaceous marine sandstones, shales, and limestones (Erikson et al., 2012).

Barinas Outwash Plain

The Barinas Basin outwash plain is composed of material derived from igneous and metamorphic rocks. The Barinas alluvial fans are 15 kilometers in width from the foothills in the direction of the Los Llanos. These outwash plains are speculated to be correlated with the terrace deposits of Trujillo and Valera Formations, which are composed of outwash material that was deposited in a lake formed by the Motatán anticline (Liddle, 1946). The Barinas outwash plain is part of the larger network of the Venezuelan Los Llanos; this is the largest sedimentary basin of Quaternary origin in Venezuela (Vargas et al., 2002).

Venezuelan Andes

The Venezuelan Andes are a doubly vergent asymmetrical orogen that is bordered on the southeastern side of the range by the Barinas Basin. The Boconó strike-slip fault system obliquely transects this orogon (Erikson et al., 2012). The South American Andes have evolved throughout time in a multi-plated regime. The Venezuelan Andes show an uncovered core of Paleozoic granitic plutons, schists, and gneisses, while sides of the mountains are covered with Cenozoic and Mesozoic sedimentary rocks (Erikson et al., 2012). Previous studies have not identified intrusive autochthonous or extrusive igneous stratigraphy younger than the mid-Mesozoic (Erikson et al., 2012). The Mérida Andes, is in the Northeastern extension of the Eastern Cordillera of the Colombian Andes. This section of the range extends in a northeastern to

southwestern direction with its structural network following suit, as seen in Figure 5. This belongs to the main chain of the Andes that covers the Pacific coast of the continent.

A major structural factor that interacts with the Mérida range is the Bócono fault. This major fault trends NE-SW and slightly oblique to the chain axis of the regional mountain network joining the Caribbean Mountain System of the northern part of Venezuela. This region is characterized by a large variety of geomorphological features including 1 to 5-km-wide valleys, scarps, ridges, and depressions (Audemard & Audemard, 2002). On its eastern side, the Bócono fault is characterized by right lateral offsets of a variety of Quaternary features, such as its ridges and alluvial deposits (Schubert, 1982). These eastern features continue to become the foothills of the Mérida Andes range, gradually fanning out into the Llanos (Audemard & Audemard, 2002).

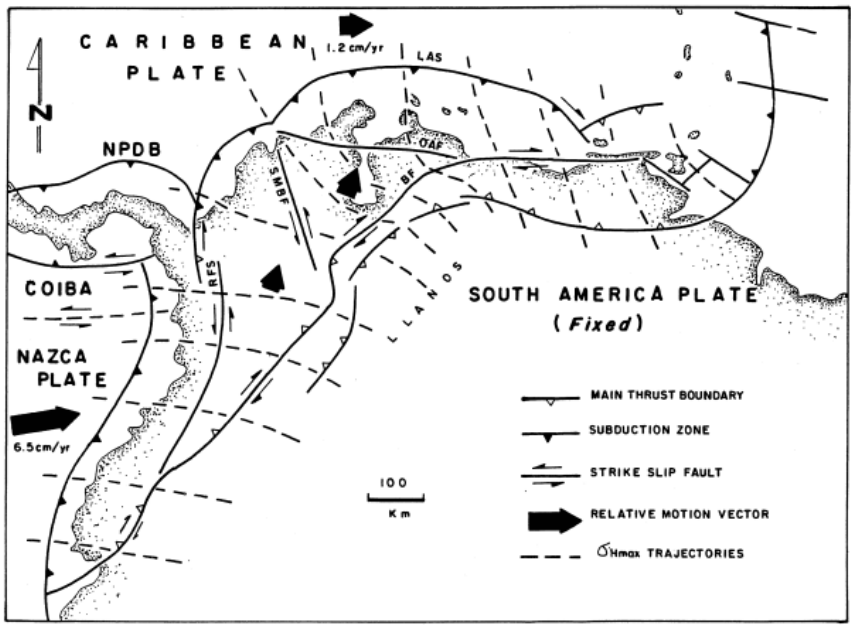

FIGURE 5: Structural Map of Merida Andes, Venezuela (Audemard & Audemard, 2002).

Caribbean Mountain System

On the northern coastal region of Venezuela, a northeastern extension of the Andes forms an elongated E–W trending belt known as the Caribbean Mountain System. This system Extends from the Sierra Nevada de Santa Marta and Guajira Peninsula in Columbia, and east to the island of Trinidad and Tobago. These ranges are separated on the west of the Venezuelan Andes by the Barquisimeto depression. The Caribbean Mountain System tectonically represented by the superposition of a series of several nappes. Geologically, this system ranges from early Paleozoic to mid-Cenozoic and consists of a mixture of all three rock types. The major tectonic events began in the Late Jurassic to the Early Cretaceous when the sealing of the oceanic location of lower Jurassic rocks transported north of the Guyana Shield. This compressive episode is related to the opening of the South Atlantic Ocean, which was accompanied by magmatic events. These events caused for a compressive phase during the Paleocene and Eocene, which resulted in the nappes being displaced and partially superimposed toward the south. The Barquisimeto depression relates the border between the thrusted nappes and the Andean Mountain range (Bellizzia & Dengo, 1990).

Structural Geology

The study area contains an exceptionally large number of alluvial plains. The plain is surrounded by mountain chains, including the Andes Mountains to the west and a smaller range of hills and mountains to the north, known as the El Baul Uplift (EBL). The structural features of these ranges predominantly are trending in a NE-SW orientation with some faults trending in a NW-SE orientation. Most of the NE-SW oriented faults are thrust faults that are thrusting to the NW due to horizontal compressional stresses (Davis et al., 2012).

Within the study area, there are a variety of different structural features. Some of the most dominant features are thrust faults and normal faults throughout the northern region. In the south are a variety of klippes, resulting from the regional thrust faulting. This continues south into the synclinal and anticlinal infused Piedmont and hill region that falls in elevation to the alluvial plain. There is a major thrust fault that is striking from NE-SW and is located in the alluvial plain that is separated from the rest of the structurally active region (Figure 4). The predominant structural pattern of the faulting area is NE-SW for the northern mountain range and NW-SE for the Merida mountain range.

Vegetation

The Guarumen area is located in two major vegetation regions in Venezuela. The majority of the region that is identified as Quaternary alluvium is part of the Los Llanos region, while the eastern part that begins to climb in elevation is part of the Mérida Andes region. Both regions have vegetation that is descendants of tropical-subtropical species, but there are two distinct regions within the Guarumen area that define the vegetation. The Merida Andes contain a variety of semi-evergreen forests that are described as a higher elevated, richer soil, and more saturated type forest. Only a certain portion of these semi-evergreen trees in the upper region of the mountain remains leafless for a short period of time. This upper region has a heavier rainfall approximately 1400–2000 mm, with a prolonged wet season of 8–9 months. Further east, the Piedmont of the Caribbean Mountain System consists of a tropical deciduous forest. This is the driest region of the Llanos, with an annual rainfall of approximately 800–1000 mm per year with a wet season ranging around 5–6 months (Sarmiento & Monasterio, 1969).

Los Llanos

The vast majority of the Venezuelan Llanos are comprised of flat or undulating lowlands. The lowland boundary is limited to the northern Coastal Cordillera mountain system, and the southern Precambrian mountain system of the Guayana Shield (Aymard, 2006). Los Llanos extends across a large SW-NE syncline occurring from the Coastal Cordillera and Guayana Shield. Due to these orogenic processes that were caused by the Andean uplift, the majority of the Llanos sits in a large depression that was originally flooded by marine waters. The Llanos was filled since the Upper Tertiary and during the Quaternary with sedimentation that dominates the geology of the region today. During the mid to Late-Neogene period, the depression went through tectonic processes that caused moderate uplift, which initiated erosion in the area and caused sediment accumulation in the region (Aymard, 2006). Due to these tectonic processes during the Neogene and Quaternary, four main types of geomorphological landscapes determine the areas soils and vegetation can be identified: eolian plains, alluvial plains, hills, and denuded surfaces. Some of the main features within the study area consist of water bodies, alluvial plains, alluvial fans and braids, lacustrine plains, piedmont, river valleys, foothill valleys, mountain valleys, high plateaus, hills, and mountains.

The Venezuelan Llanos contain a large variety of ecosystems with a large diversity of plant communities and different vegetation types. A complete list of the region's taxonomy of the vegetation types has not been fully written yet, but regional studies have been completed with different classification schemes. Some of the most important classifications were composed by Tamayo (1964), Ramia (1967), and Sarmiento (1984), but the classification of the area with the widest range of acceptance is the Beard (1953) classification (Aymard, 2006). Four major

regions are identified within the Llanos: the deciduous forest region, the semi-evergreen forest region at the Piedmont of the Andean Range, the swampy savannas in the floodplains of the Orinoco River, and the dry savanna region that is located in between the deciduous forest region and the Orinoco valley (Sarmiento & Monasterio, 1969). In terms of this research, the Llanos region located in the Guaramen area is primarily located in the semi-evergreen forest region, as seen in Figure 6 from Sarmiento & Monasterio (1969).

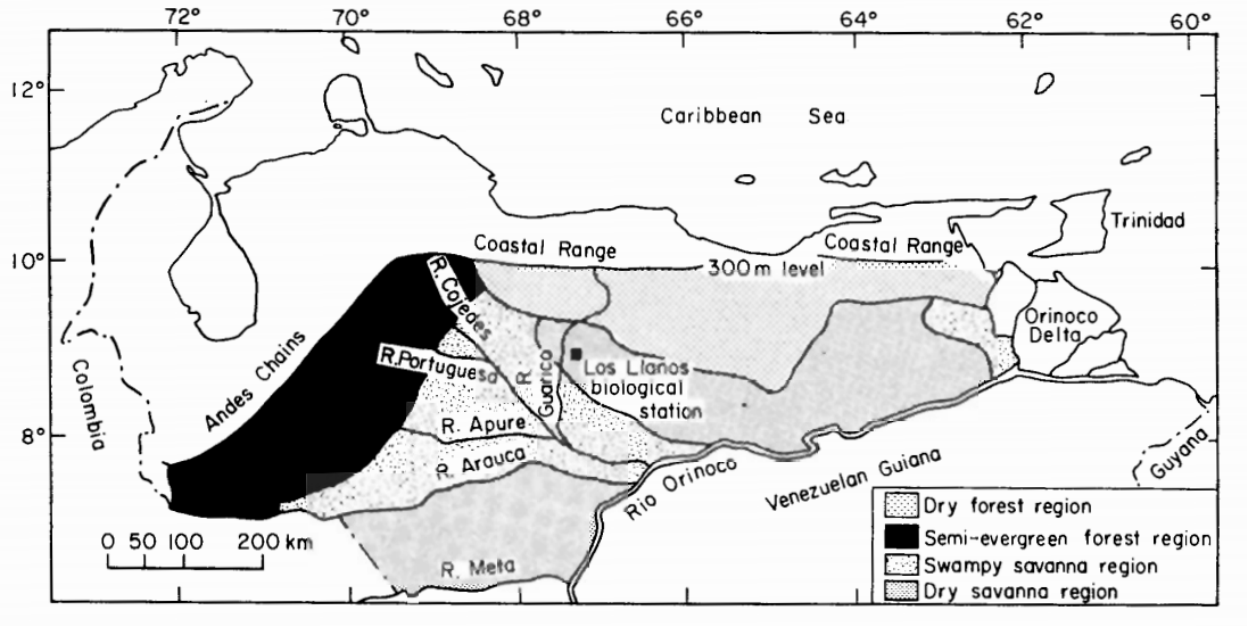

FIGURE 6: Map of Los Llanos Region of Northwestern Venezuela (Sarmiento & Monasterio, 1969)

METHODS

Data Acquisition

The data acquired for this research was provided by SpecTIR, a hyperspectral imagery company based out of Reno, Nevada. SpecTIR provided a series of hyperspectral imagery datasets of the Guaruman area. The hyperspectral imagery was acquired on January 19th, 2016 using a ProSpecTIR–VS2 airborne sensor with a flight altitude yielding 3-meter pixel data. This sensor was configured to collect from 390–2450 nm. This sensor consists of 360 bands with a 5 nm spectral resolution with a full-width half-maximum (FWHM) varying about 3.8 nm to 8.5 nm, with the full wavelength range of the sensor (Herweg et al., 2012). The data were atmospherically corrected from radiance to reflectance data with ATCOR 4 using version 6.2 MODTRAN. Other datasets used were acquired from USGS Earth Explorer. These were primarily Landsat 5,7, and 8 datasets with one set of ASTER digital elevation models. All imagery was placed in a database for ease of access.

Region of Interest Selection

Processing data of the whole study area would be ideal, but due to technological limitations and time constraints, the data were separated into multiple Regions of Interest (ROI) to process the data in a timely manner. Four ROI were selected based on their variety of location, geology, geomorphology, and structural characteristics. Each location was used to identify a

specific region of the larger study area and was used to make correlations with vegetation and soil anomalies.

ROI (Figure 7) was located in the southwestern region of the study area, bordering the foothills of the Merida Andes. It lies within five different geologic formations; these include Rio Guache, Rio Yuca, Guanapa Formations, and Quaternary alluvium formations. Geomorphologically, it has a variety of alluvial fans and fluvial braids, hills, and piedmont. There are no structural features in the ROI observed on the surface, but just south is a thrust fault thrusting upward in a northern direction. The ROI contains a mixture of sedimentary rocks ranging between calcareous shales, conglomerates, limestone, marl, sand, silts, and clays.

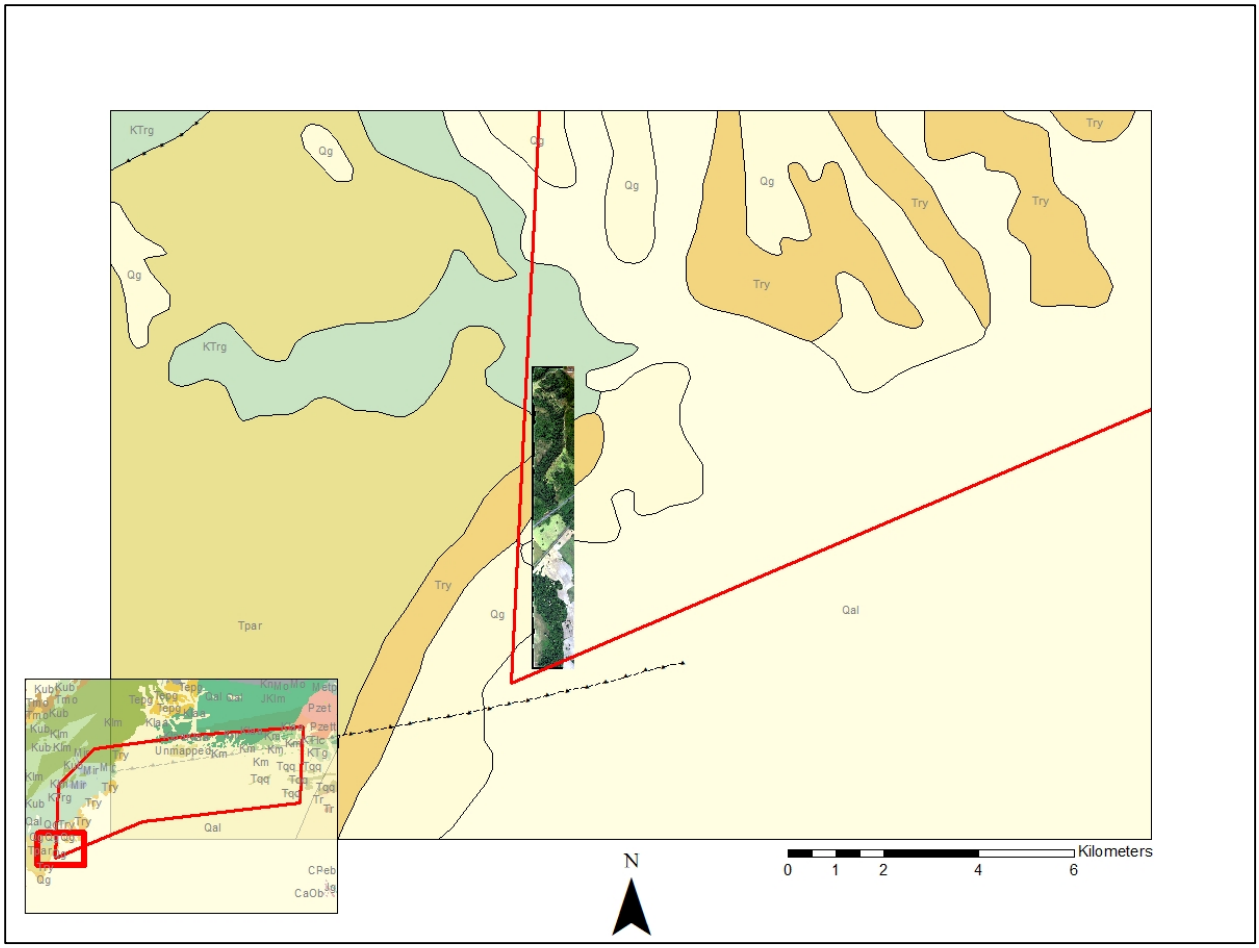

FIGURE 7: Regional Map of ROI 1.

ROI 2 was located in the north central region of the study area (Figure 8). This location is set within the Quaternary alluvium deposits and is located in a highly agricultural region, with a mixture of healthy vegetative agricultural plots, and some that are in the process of controlled burning. The northernmost part of the ROI is the Las Majaguas Reservoir; this is being used as a control for water with the endmember analysis.

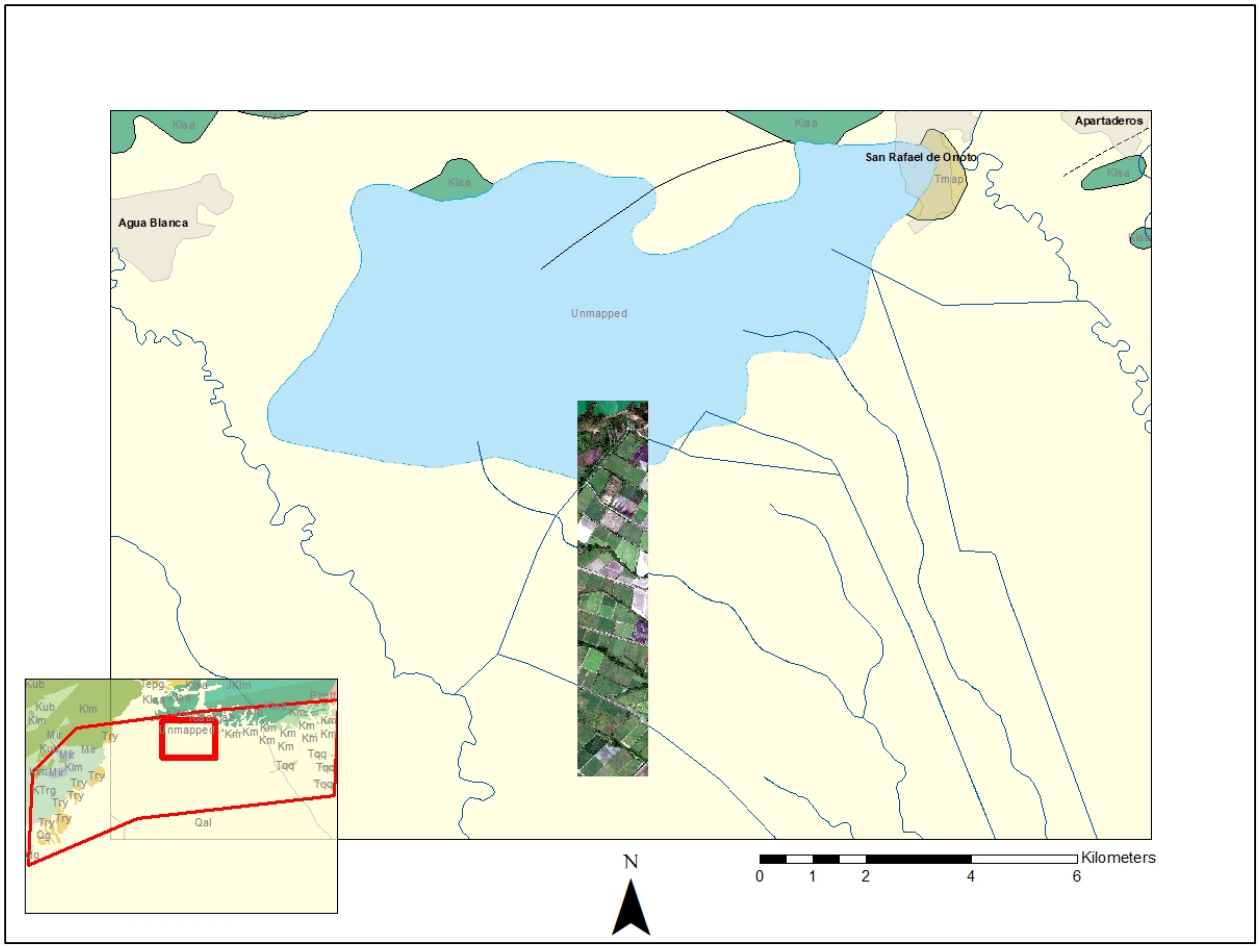

FIGURE 8: Regional map of ROI Two.

ROI 3 is located in the northeastern part of the study area (Figure 9). This region was selected within a highly active structural environment with one unclassified fault running in a southeastern direction, and a syncline transecting the center. ROI 3 lies within the Quaternary alluvium, containing sand, gravel, and silt. The Murcaria Formation contains carbonaceous shale, and the Agua Blanca, Araure, and Cojedes Formations contain limestone and conglomerate.

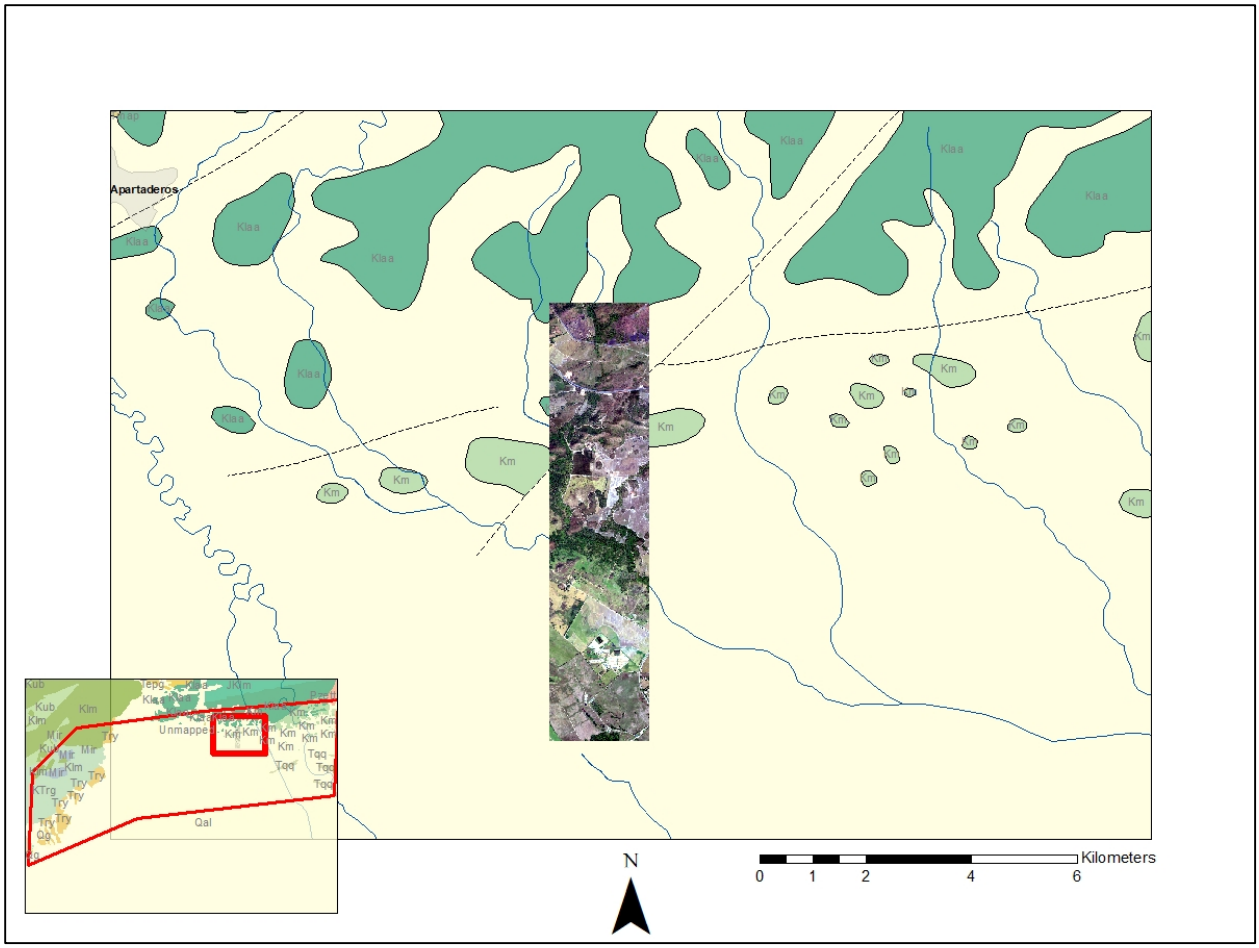

FIGURE 9: Regional Map of ROI 3.

ROI 4 is located northwest of ROI 1 and sits in the Quaternary alluvium, and the Rio Yuca Formation (Figure 10). These formations consist of unconsolidated sands, sandstone, silts, gravels, and conglomerates. In the southern part of the region, there is a major stream crossing in a northwest to southeastern direction. This stream deposits a high sediment load of mafics, marl, calcareous shales, and mudstones from the Rio Guache Formation.

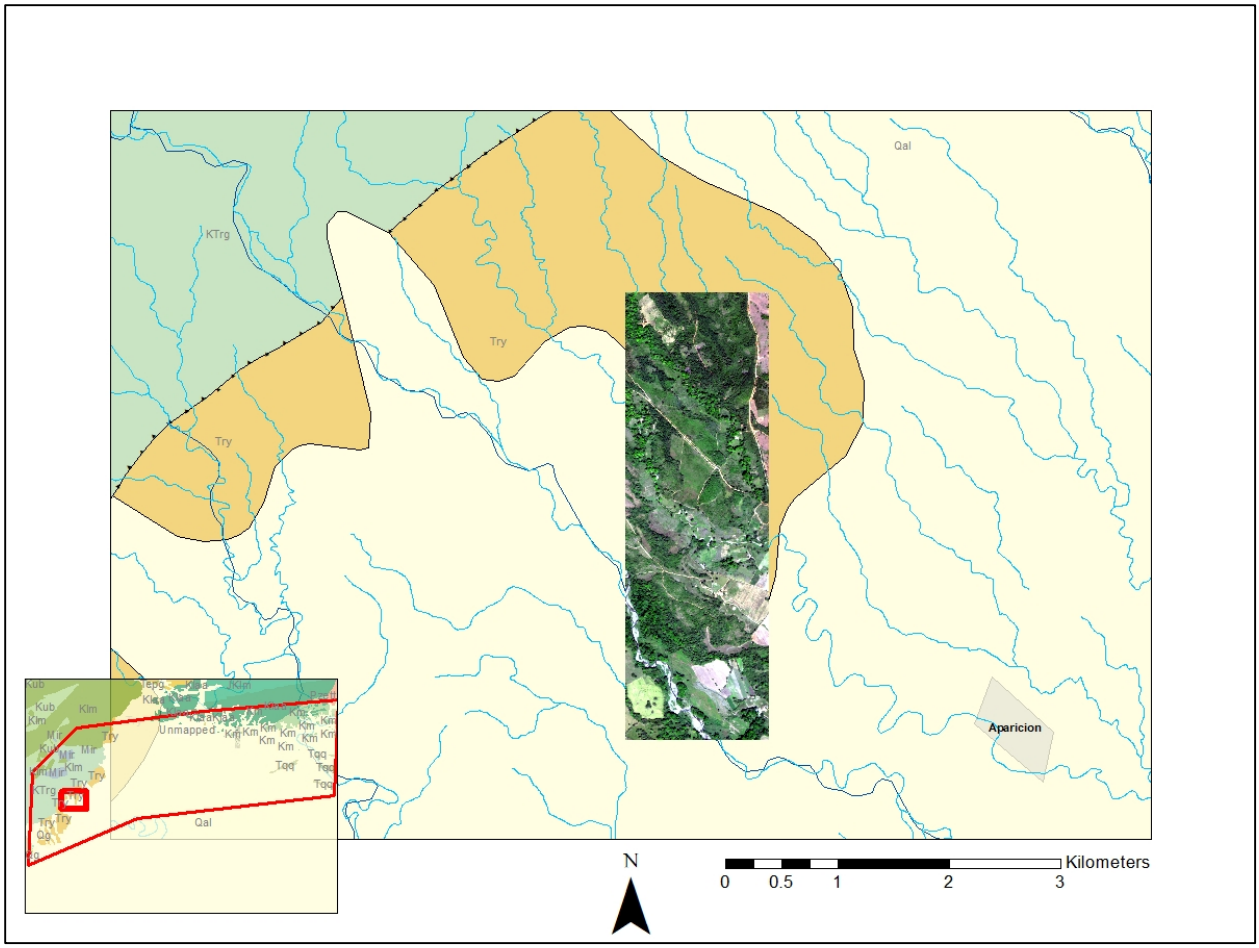

FIGURE 10: Regional map of ROI 4.

Image Processing

The hyperspectral, multispectral, and DEM imagery used was processed with ENVI 5.3. The geospatial processing was conducted with ESRI 10.4.1. This included mapping and metadata of geology, geomorphology, elevation, and structural maps. The hyperspectral imagery was composed of 360 bands with ground sampling distance (GSD) at 3 meters in both the X and Y directions. The imagery was smoothed with a Savitsky-Golay convolution filter, and CO_2 mitigation applied. Again, with the large study area, and limited field campaign, the hyperspectral imagery was subset into four ROIs. These four ROIs were identified and selected

based on limited field data collected from previous literature, and differing geologic, geomorphologic, and structural locations throughout the research area.

Image Analysis

The objective of the analysis was to investigate the geological setting of northwestern Venezuela and further the understanding of the region because it has become a region of interest for economic natural resources. Hyperspectral and multispectral data and spectral data processing tools were used to address unique derived endmembers for classification of land cover, and to analyze the classifications to characterize the outcomes with respect to the influence of geological phenomena. This was done to identify how these anomalies related to the regional geomorphology and structural geology and identify any form of correlation within the study area. The primary method used for this analysis was to conduct an endmember analysis with the hyperspectral data by using the Spectral Hourglass Workflow. This tool was provided through ENVI 5.2, with the capabilities to limit the number of bands to the primary bands that are purest in terms of its pixels. With the ability to spectrally analyze the image derived endmembers for specific spectral signatures, and compare them to spectral libraries, makes it possible to find anomalies in both vegetation and soil properties.

The secondary method for this analysis consists of identifying the same anomalies mentioned before, but by mapping the regional emissivity and land surface temperature to identify abnormality, and correlate them to the previous method.

Image Derived Endmembers

With a high spectral resolution of hyperspectral imagery, interpretation strategies are taken into consideration based on the data, and spectral unmixing is facilitated. Mixed pixels are common in hyperspectral data due to a poor spatial resolution of the imaging spectrometer, or intimate mixing effects. This mixture problem has both a macroscopic and microscopic dilemma in terms of mixtures on a large scale with grass, trees, and soil, in a 3x3 meter-pixel, or on a small scale with minerals intimately mixed in a 1-meter pixel (Mauro et al., 2014). A key dilemma with unmixing is deciding whether to use field derived spectra observations, or image-derived from the dataset. The benefit of image-derived endmembers advances itself with automation and does not require the need for expensive field spectrometers or the need for field visits. Image-derived endmembers may, however, neglect to incorporate minor landscape components that never attain abundance levels that they dominate a single, *pure*, pixel (Dehaan & Taylor, 2003).

Spectral Hourglass Workflow

The method used to conduct this bad band elimination is known as the “Spectral Hourglass Processing Workflow” (Figure 11). This workflow leads to finding and mapping spectral endmembers by allowing sub-pixel target detections, unambiguous spectral mixing, and material identification. Beginning from the top of Figure 11, this method normally begins with atmospherically correcting the imagery with FLAASH, but due to the preprocessed acquisition through ATCOR4 MODTRAN, this step was already completed.

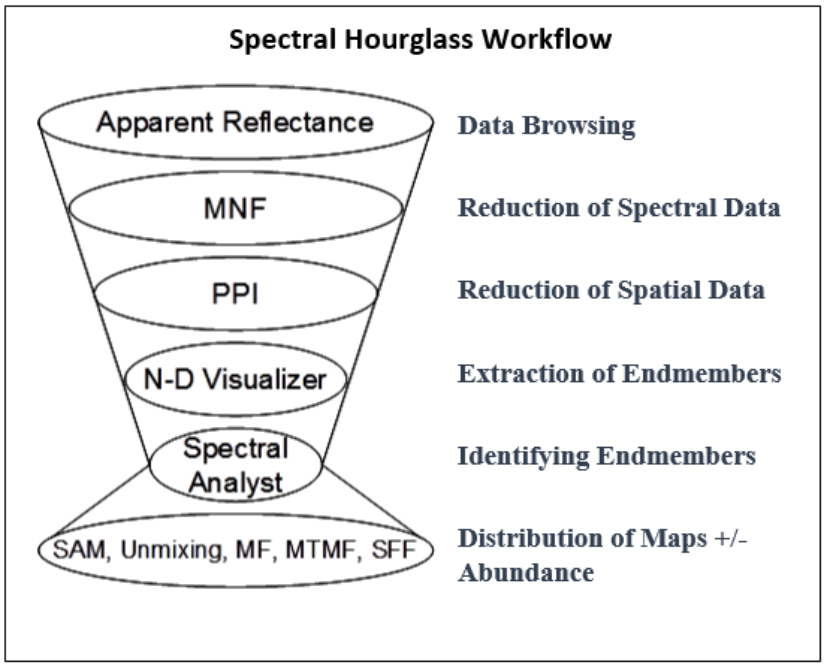

FIGURE 11: Spectral Hourglass Workflow (Modified from Harris Geospatial, 2018).

The next step of the hourglass was computing the minimum noise fraction (MNF). The MNF is used for reducing large multiband datasets into a smaller number of components that contain the majority of the spectral variance within the dataset (Harris et al., 2005). The formal definition, which was provided by Boardman and Kruse (1994) states that MNF transforms are used to determine the intrinsic dimensionality of the image data and to separate the noise in the data while reducing the computational requirements for further processing. The MNF transform is modified from Green et al. (1988), and was implemented in ENVI. It involves two sequences of principal component analysis (PCA). The first rotation uses the noise covariance matrix principal components to deconstruct and rescale the noise in the data. This noise also known as *noise whitening*. The results of the first rotation transformed data that have no band-to-band

correlations, but do contain unit variance. The second rotation uses the derivatives of the principal components from the original image data after the noise-whitened process has taken place and rescaled by the noise standard deviation. These processes are determined by examining the final eigenvalues of the image file. For this research, a forward transform MNF was used to help determine which bands contained coherent images and to estimate the noise statistics from the data (Green et al., 1988).

The next processes on the hourglass workflow were pixel purity index (PPI) and the n-dimensional visualizer. The PPI is used to define the most spectrally-pure pixels in both hyperspectral and multispectral data. The PPI is computed by random vector selection through the n-dimensional data cloud, where each individual vector is constrained to pass through the center point (the mean value) of the n-dimensional cloud and are plotted on the visualizer. The n-dimensional visualizer is a multi-dimensional scatter plot that labels the purest pixels on a random unit vector. It is used to identify, locate, and cluster the purest pixels and the extreme spectral endmembers in the n-dimensional space. This is done by defining the PPI threshold and locating the tail end of the pixels plotted on the random unit vector (Boardman et al., 1995).

The PPI threshold is defined by the user. Usually, the threshold is defined by using the same units as the pixel values, but when performing a MNF, as in this study, the threshold is defined by the noise standard deviation units. The lower the value of the threshold, the fewer pixels will be identified as *pure*. This tool helps to visualize the shape of the data cloud by plotting image data in spectral space (Harris Geospatial, 2018). The purest pixels are distinguished by recurrently projecting n-dimensional scatter plots onto a random unit vector. The purest pixels in every projection and the number of times the pixel is marked extreme are

recorded (Boardman et al., 1995) (Figure 12). These pure pixels are expected to represent spectrally homogenous samples of the mixing endmembers.

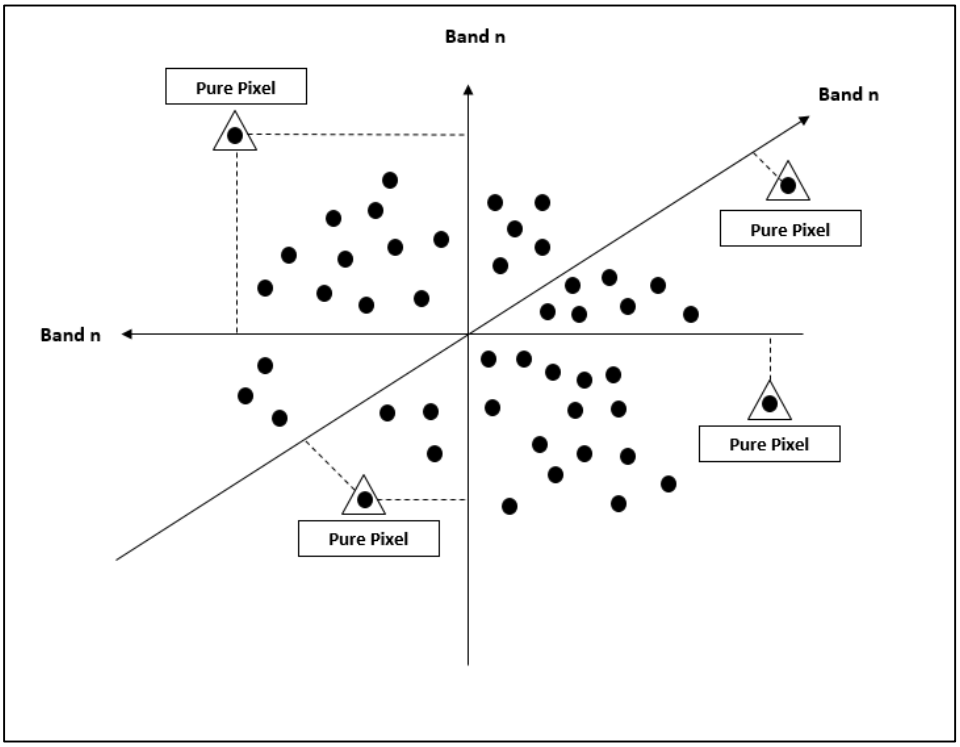

FIGURE 12: n-dimensional pixel purity model (Modified from González et al., 2018).

The ability to visualize the algorithmic procedure of the PPI is shown by the hyperspectral data **F** with n dimensions; the number of bands that are generated through the process, **K**; and a limited threshold value, $\mathbf{t_v}$, which identifies as final endmembers based on the number of times those pure pixels have been selected as extreme. At least $\mathbf{t_v}$ amount of times in the PPI step of the hourglass workflow (González et al., 2010). The PPI algorithm can be seen as follows:

(1) *Band generation*. Randomly generate sets of bands **K** as unit vectors.

$$\{\mathbf{Band}_j\}\binom{K}{j} = 1 \qquad \text{Equation 1}$$

(2) *Extreme estimations.* Every **Band** $_j$ ={1,…., K}, with each pixel vector $\mathbf{f}_i$ in its original dataset **F** is projected with each band using pixel products of $|\mathbf{f}_i * \mathbf{Band}_j|$ to identify sample vectors in its minimum and maximum projections (extreme). This is the extrema for each $\mathbf{Band}_j$, termed as $S_{Extrema}(\mathbf{Band}_j)$. Even though different bands create different extreme pixel sets, it is probable that some vectors may appear in more than one extreme pixel set. To justify this, it is defined that the indicator function of a set **F**, is denoted by ℓ_S (x), to denote the relationship of an element **x** to the particular set is as follows:

$$\ell_S(x)=\begin{cases}1 \; if \; x \; \in S_{Extrema} \\ 0 \; if \; x \notin S_{Extrema}\end{cases} \qquad \text{Equation 2}$$

If x is identified as an *extrema*, then ℓ_S is populated with "1", which is an indicator member of the subset. If x is not identified as an *extrema*, then ℓ_S is populated with "0", which indicates that it is not a member of the subset.

(3) *Calculation of PPI.* With the indicator function above in use, the PPI score was calculated based on the pixel vector $\mathbf{f}_i$ (The number of times a given pixel is selected as extreme) and use the following equation:

$$N_{PPI}(\boldsymbol{f_i}) = \sum_{\boldsymbol{j=1}}^{\boldsymbol{K}} ls_{extrema}\big(\boldsymbol{Band_j}\big)(\mathbf{f}_i) \qquad \text{Equation 3}$$

(4) *Endmember Selection.* In the selection process, the n-dimensional visualizer is used to locate pixel vectors with scores of $N_{PPI}(\boldsymbol{f_i})$ that are above $\boldsymbol{t_v}$ and label them in the ROI tool as spectral endmembers that are assigned a specific color attribute (González et al., 2010).

After the endmembers were identified, the major work can be started identifying endmembers. This will ultimately lead to distribution mapping, the penultimate step in the

hourglass workflow. The spectral analyst tool was used to identify specific endmembers, classify, and compare them to specific spectral libraries. The analyst tool used does not identify the spectra; it only provides likely recommendations for spectra identification. The actual identification of the spectra is based on the user interpretation. The spectral analyst tool is best used to identify materials based on their spectra. The spectral libraries used were the provided library database defaulted by ENVI 5.2.

Using the spectral analyst tool through ENVI, a weight needed to be input for both of the similarity methods. By default, the spectral feature fitting uses a weight of 1.0, which is the weight that was used during the process. The minimum and maximum values were set at a scale from 0.0 to 1.0 for both the Spectral Feature Fitting, and the Binary Encoding, but for the Spectral Angle Mapper, a Min-Max was set to 0 to 0.78540, which was the default setting. Each of these different method descriptions was read with different values. For example, the Spectral Angle Mapper values are read in radians, the Spectral Feature Fitting is read in RMS error units, and the Binary Encoding is read in the percentage of the bands. Each of these values were plotted in a spreadsheet with the recommended spectral identity, the value representing the precision of both spectra from the library, and the data (Harris Geospatial, 2018).

The final stage of the hourglass is the post classification stage. This includes: performing a spectral angle mapper (SAM) classification that uses the n-dimensional angle to match the pixels to the spectra (Kruse et al., 1993); Linear Spectral Unmixing (LSU), that shows the relative abundance of materials that are based on specific spectral characteristics and that the reflectance of each pixel is at a linear grouping of each endmember is present within the pixel (Dehaan & Taylor, 2003); Matched Filtering (MF), to locate abundances for user-defined endmembers by using a partial mixing (Boardman et al., 1995); Mixture Tuned Matched

Filtering (MTMF) to add an infeasibility image to the results; and Spectral Feature Fitting to compare reference spectra to the fit of image spectra using a least-squares technique (Harris Geospatial, 2018).

Kappa Coefficient

The kappa statistic, or Cohen's kappa, is a statistical analysis used widely in remote sensing applications to test interrater reliability. Interrater reliability is important because it signifies data collected in the study area is accounted as a correct representation of the variables measured. The data collectors (raters) assign the same score to its matching variable; this is called the interrater reliability score. Cohen suggested that the kappa scores be interpreted as follows (Table1) (McHugh, 2012).

Value of Kappa	Level of Agreement	% of Data that are Reliable
0-0–20	None	0–4%
0.21–0.39	Minimal	4–15%
0.40–0.59	Weak	15–35%
0.60–0.79	Moderate	35–63%
0.80–0.90	Strong	64–81%
Above 0.90	Almost Perfect	82–100%

TABLE 1: Interpretation of Cohen's kappa

The kappa assessment shows that a coefficient equal to 1 means a perfect agreement whereas a value that is close to 0 means that the agreement is no better than it would be expected by chance (Rwanga & Ndambuki, 2017). This statistic is calculated through hyperspectral imagery by creating a random point generator where a select number of random points are generated on a classification image. This is then plotted on a confusion matrix where it is compared against the

observed accuracy of the classification. This observed accuracy can be determined by the diagonal in the error matrix. This is computed by the following equation:

$$K = \frac{\text{Observed Accuracy} - \text{Chance Agreement}}{1 - \text{Chance Agreement}} \qquad \text{Equation 4}$$

Microclimate Analysis

Regional climate change with interactions between vegetation, environment, and climate in the long-term, and a lack of, or excess of precipitation, groundwater, and surface water in the short-term, can create a window for anomalous vegetation development in spectral properties that can be observed through satellite and airborne imagery. Both emissivity and land surface temperature (LST) have been studied thoroughly as a powerful tool in land surface properties (Sun et al., 2012). LST is defined as the temperature interaction where the surface is touched with the hand or body temperature on the ground. This method has proven itself successful in terms of climate change, agricultural processes, and land cover. LST was originally generated using MODIS and ASTER data of the same date and time and matched to create a LST output. This was a successful attempt for calculating the LST, but recently an updated algorithm was created using Landsat 8, which computed a result with better accuracy (Rajeshwari, 2014).

For my study, a microclimate relief map was computed from the LST of the study area to detect possible anomalies related to microclimate conditions based on two separate time frames of the same location. This process used Landsat 8 data computed on ENVI 5.3 using both the band math and radiometric calibration tools. The algorithm process consisted of five major processes to achieve LST as follows (Avdan &Jovanovska, 2016):

(1) . *Radiometric Calibration*. Once the bands are converted to reflection, the TIRS band data is converted from spectral radiance to brightness temperature (BT), which is calculated from the thermal constants (K_1 and K_2) from the metadata. The following equation is used in the band math tool through ENVI to achieve this process.

$$BT = \frac{K_2}{\ln(\frac{K_1}{L\lambda}+1)} - 273.15 \qquad \text{Equation 5}$$

To obtain these results in Celsius, the absolute zero constant (-273.15 ˚C) was used to convert the final values from Kelvin to Celsius.

(2) Calculating NDVI. The visible and near-infrared bands were used to calculate normalized difference vegetation index (NDVI). Landsat 8 bands 4 and 5 were used in the equation below to calculate the vegetation index of the Landsat scene. The calculation of the NDVI is crucial to this algorithm process due to the later calculations of proportion of vegetation and emissivity. In this equation, the NIR represents the near-infrared (band 5) and R represents the red band (Band 4).

$$NDVI = \frac{\text{NIR (band 5)} - \text{R(band 4)}}{\text{NIR (band 5)} + R \text{ (band 4)}} \qquad \text{Equation 6}$$

(3) Proportion of Vegetation. This was conducted by calculating the min/max range of the calculated NDVI from the previous equation.

$$P_v = \left(\frac{(NDVI - NDVI_{min})}{(NDVI_{max} - NDVI_{min})}\right)^2 \qquad \text{Equation 7}$$

(4) Land Surface Emissivity. In order to calculate LST, land surface emissivity must be known. Due to the land surface emissivity being a proportionate factor of Planck's Law to predict emitted radiance, the ground emissivity is calculated as follows:

$$e = \varepsilon_{v\lambda} P_v + \varepsilon_{s\lambda}(1 - P_v) + C_\lambda \quad \text{Equation 8}$$

Where ε_v and ε_s are the mean vegetation and soil emissivity values, and C represents the surface roughness.

(5) Land Surface Temperature. The final phase of the algorithm is to calculate the LST in Celsius. This equation was calculated in the Band math tool as follows (Stathopoulou & Caralis, 2007):

$$LST = \frac{\text{BT}}{\left\{1+\left[\left(\frac{W}{P}\right)\ln(e)\right]\right\}} \quad \text{Equation 9}$$

Where W is equal to the wavelength of the emitted radiance, P is equal to Planck's constant multiplied by the velocity of light divided by Boltzmann's constant (Weng et al., 2004).

Upon Completion, the LST map was converted from grayscale to a multicolored image. This provided clearer visualization of the high and low LST of the Landsat image. The LST data were then plotted on a histogram to give a representation of how many pixels were collected in each temperature range. The cursor value tool was used to identify specific locations from the ROI to measure the surface temperature of a location and compared for quality. These surface temperatures were then classified within a color range to be projected in map view. I then spatially analyzed and interpreted the surface temperature data anomalies located in the microclimate map, then compared to the regional geological features within the region to find potential correlations.

RESULTS

Endmembers and Spectral Data

Four ROI were selected throughout the study area to give a variety of geology, vegetation, geomorphology, and structural characteristics. Due to the study area being large, this multiple analysis was necessary due to processing capabilities. For each ROI, the endmembers were extracted and plotted as spectral reflectance in reference to one another (Figure 12). Anomalies, or similarities were observed within specific regions of the study area to understand how the soil and vegetation affect the geology, geomorphology, and structural geology of the region. The endmembers for each ROI were identified through the "purest pixels" in the n-dimensional space with wavelengths ranging from 0.01 to 2.45536 micrometers. Each of these endmembers appeared to be associated with a variety of soils, vegetation, water, and urban characteristics. These endmembers were derived through a multiple phase process to ensure the analysis of these endmembers were conducted as accurately as possible without the use of field data. The endmember analysis was an efficient classification method to extract the purest pixels of the hyperspectral imagery and locate any trends within the ROI to address the research objectives.

Each of the four ROI spectra revealed specific characteristics to the larger study area region. These characteristics varied based on their geological and geomorphological location and showed specific regional features that could be further investigated based on soil and vegetation anomalies.

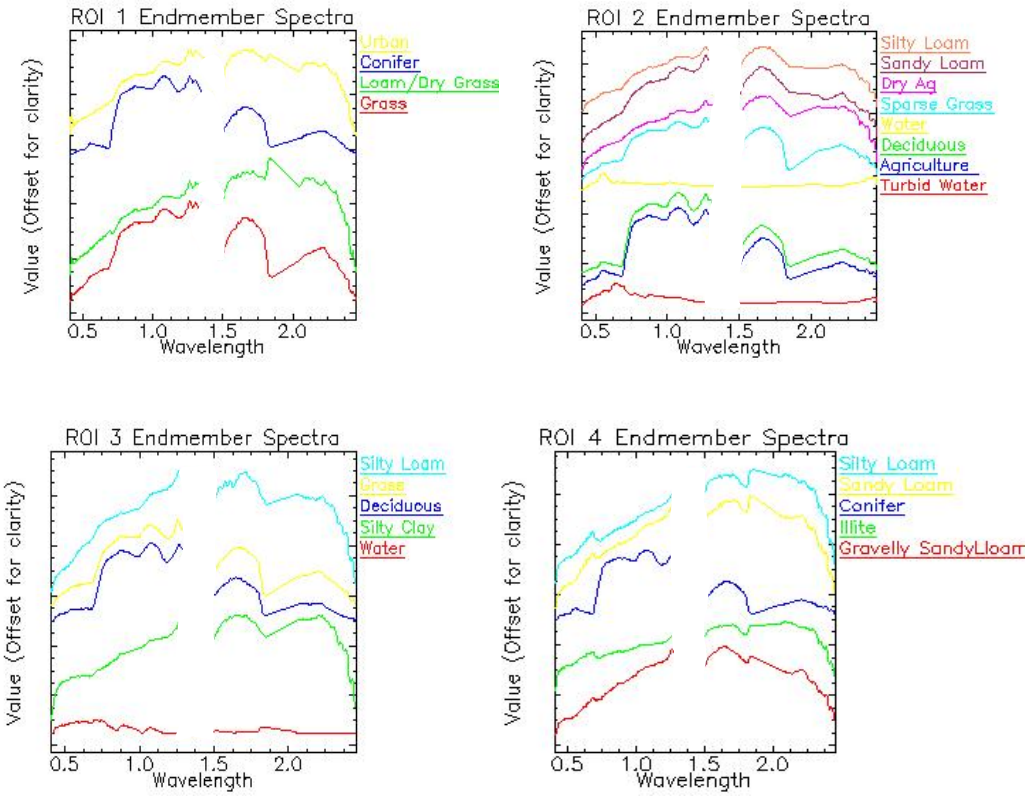

FIGURE 13: Endmember spectra of four ROI locations. (Wavelengths in μm).

Spectrally distinct endmembers were extracted by using the spectral analyst tool to its best ability, and if failed, the spectra were determined by examining the image distribution, spectra, and abundance of the ROI. The n-dimensional classes were roughly characterized as belonging to some form of soil, which includes mineralogy and petrologic properties, and vegetation, which differentiated on a macro scale from trees, and flat ground cover ranging from grassland to agriculture. Each of the ROI spectra was compared to one another to identify any anomaly that may be present within the region.

ROI 1

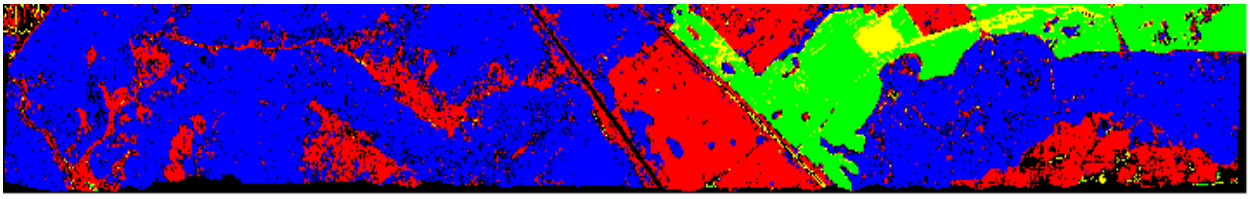

FIGURE 14: SAM image classification of ROI 1

ROI 1 (Figure 13), was a highly forested area with limited to no agriculture. The majority of the vegetation make up consisted of coniferous trees, while the remaining region consisted of healthy grass, urbanization, and dry grass with an exposed loamy soil. When measuring the agreement and accuracy of ROI 1 and its classification, the ROI received a Kappa Coefficient of 0.52 (Table 1) and an overall accuracy assessment of 68.11%.

Class	Coniferous	Grass	Dry Grass	Urban	Total
Coniferous	1923	0	0	0	1923
Grass	6	2590	0	0	2596
Dry Grass	0	0	5434	0	5434
Urban	0	0	622	1	623
Unclassified	42	18	0	165	225
Total	1971	2608	6056	166	10801

TABLE 2: Kappa Coefficient and accuracy assessment of ground truth pixels of ROI 1

The first endmember of this ROI, “Urban” (yellow) consisted of a small percentage of the area. Within the Kappa statistic, the value registered in the majority of dry grass. This was due to the albedo effect, and the high reflectance of particular areas of the dry grassy/loam locations registered as urban. This can be seen in the dry grass/loam location.

The next endmember identified was a coniferous tree network, represented as blue in (Figure 13). This endmember covered the majority of the ROI, and throughout the Piedmont and

hill regions of the Merida Andes as well. There was slight chlorophyll absorption at the 500–550 nm, and a high reflection in the near infrared ranging from 800–1200 nm. The water absorption bands of this endmember responded around 1500–2400 nm.

The dry grass/loam endmember was separated from the healthy grass based on its high albedo. This endmember was contained to a single region of the ROI where the lack of vegetation health and bare soil were exposed due to a lack of water, or land management. Based on this endmember located in the Quaternary alluvium plain, the soil of this exposed surface more than likely contained a mixture of sand and silt along with the dry vegetation within the pixels. The water absorption features of this endmember give a range of 1,400–2,300 nm and reduce drastically at 2,400 nm, showing a possible increase in soil moisture.

The last endmember recorded with ROI 1 had the characteristics of healthy grass vegetation. This endmember's spectra had a high chlorophyll absorption around 600 nm with a defined cell structure in the near infrared at approximately 800–1,200 nm, and a slightly better absorption feature at 1,800 nm. This endmember occurs both in an agricultural and natural state within the ROI.

ROI 2

ROI 2, (Figure 14), contained multiple variations of soil classifications with two different types of water bodies, and a deciduous tree network rooted within a highly agricultural area. The soil types varied in this region with a higher concentration of silts and sands that were exposed due to exposed agricultural activities, or a decrease of surface water due to reservoir damming. This location is shown to be the most agriculturally active out of the entire study area. The analysis of this ROI received a Kappa Coefficient of 0.64 and overall accuracy of 72% (Table 2).

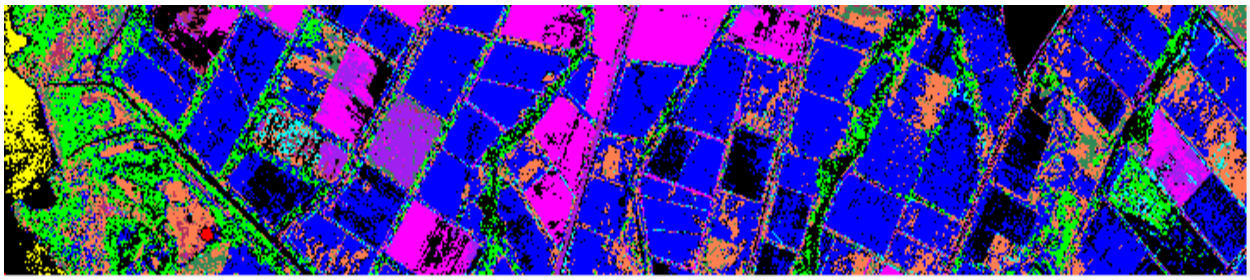

FIGURE 15: SAM image classification of ROI 2

Class	unclassified	Turbid water	Deciduous	Agriculture	Water	Sparse Grass	Dry Ag	Sandy Loam	Silty Loam	Total
Unclassified	24342	1	1174	3520	209	133	247	81	40	29747
Turbid Water	2	49	0	0	0	0	0	0	0	51
Deciduous	811	0	11161	937	0	1	0	0	5	12915
Agriculture	1589	0	1049	52492	0	20	0	1	8	55158
Water	207	0	0	0	1340	0	0	0	0	1547
Sparse Grass	271	0	12	242	0	1402	47	56	79	2213
Dry Ag	153	0	0	0	0	12	11135	1	0	11536
Sandy Loam	59	0	12	22	0	64	13	2812	157	3139
Silty Loam	212	0	0	14	0	74	345	74	4893	5612
Total	51988	51	14582	60747	1758	1839	12034	3058	5676	151733

TABLE 3: Kappa Coefficient and accuracy assessment of ground truth pixels of ROI 2.

The first endmember was identified as turbid water due to the presence of suspended soils. This can be seen in the spectra 650–700 nm range where it has a higher response in reflectance compared to the water spectra from the reservoir. The spectra water from the reservoir showed relatively little absorption in the wavelengths less than 600 nm, this was shown by a higher transmittance in the blue and green portion of the spectrum. In this ROI, a deciduous tree network was identified based on spectral reflectance of the classified endmember as a visual interpretation of the imagery. These trees are closely associated with a type of broadleaf tree, based on regional studies. The deciduous endmember spectra contained a moderate chlorophyll

absorption at 550 nm with high reflectance in the near infrared from 650–1,300 nm range. The absorption bands of this spectra were low in reflectance due to high water absorption rates at the 1,500–2,000 nm range. The agricultural endmember of ROI 2 was similar in spectral reflectance as the deciduous endmember, but, with minor differences in reflectance, the analysis was able to detect the differences between the two endmembers. The stressed agriculture was easily detected through the analysis phase due to the high reflectance rate in the spectra. By spatially deriving the image, it is shown that some of the dried, stressed region is due to possibly controlled burning, or a change in crop for next planting season. The majority of the exposed soils identified in the analysis were sandy and silty loams. These two endmembers can be distinguished based on the reflectance where sandy loam has a lower reflectance than silty loam.

ROI 3

ROI 3 (Figure 15) was located in a highly structural region with high soil exposure in the Northern part of the study area. This location is south of the Caribbean mountains in a deciduous tree network with a soil mixture identified as silty clays and silty loams. This ROI received a Kappa Coefficient of 0.97 and overall accuracy of 98.62% (Table 3).

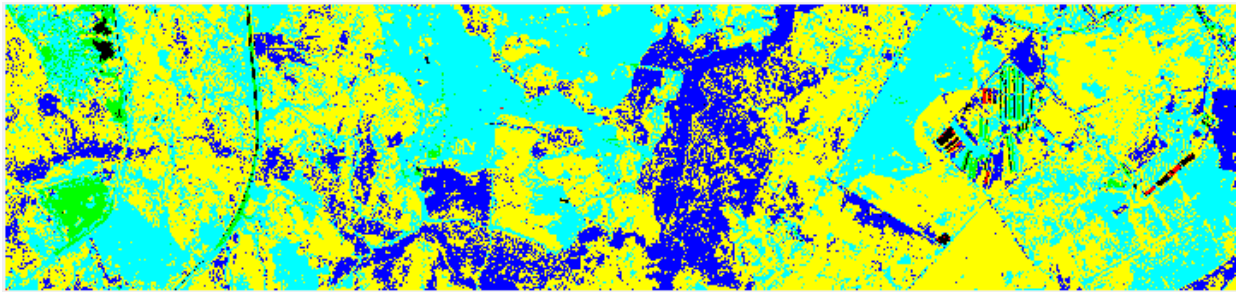

FIGURE 16: SAM image classification of ROI 3

Class	Water	Silty Clay	Deciduous	Grass	Pale Brown Silt
Water	75	0	0	0	0
Silty Clay	0	50	0	0	0
Deciduous	0	0	1881	0	0
Grass	0	0	0	69	0
Brown Silt	0	0	0	0	0
Unclassified	2778	0	0	0	0
Total	2853	50	1881	69	0

TABLE 4: Kappa Coefficient and accuracy assessment of ground truth pixels of ROI 3

The first endmember classified in the ROI 3 spectra was silty loam. This is located throughout the ROI and covers most of the structural features, (e.g., klippes). During the spectral analysis, the silty loam also showed traces of carbonaceous shale, which is a prominent feature in the Mucaria Formation, located in ROI 3. Silty clay is located in the ROI; this was determined based on the difference of the reflectance in the spectra. The silty clay endmembers showed a more prominent absorption feature at the 1,600–2,200 nm range than the silty loam endmember spectra due to clays' impermeable features, but the silty loam showed a higher reflectance overall since water was more easily absorbed.

ROI 4

ROI 4 (Figure 16) is located at the western part of the study region covering the Piedmont, hills, and the braided streams of the Merida Andes. The ROI is geologically located in both the Rio Yuca Formation and the Quaternary alluvium outwash plains. These formations contain a mix of unconsolidated sands, silts, gravels, sandstones and conglomerates, and the

braided stream contains a mixture of illite and possible igneous rock transported from the Cerro Pelon Formation upstream. This ROI endmember classification received a Kappa Coefficient of 0.47 and overall accuracy of 64.09% (Table 4).

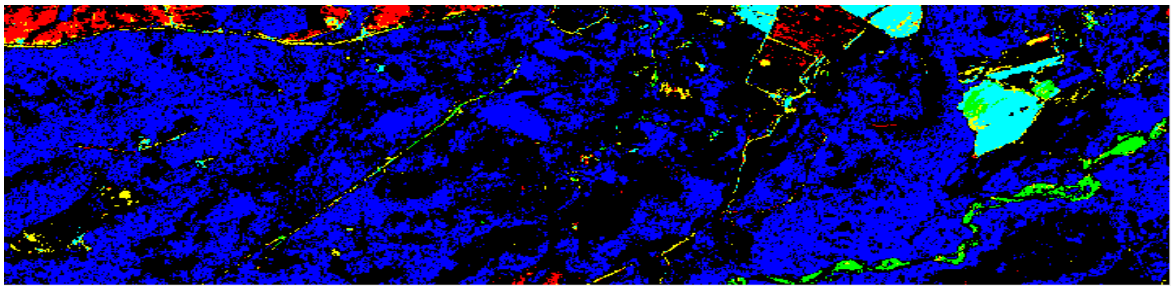

FIGURE 17: SAM image classification of ROI 4

Class	Silty Loam	Sandy Lam	Conifer	Illite	Gravelly Sandy Loam
Silty Loam	75	0	0	0	0
Sandy Loam	0	50	0	0	0
Conifer	0	0	1881	0	0
Illite	0	0	0	69	0
Gravelly Sandy Loam	0	0	0	0	0
Unclassified	0	0	0	0	2778
Total	75	50	1881	69	2778

TABLE 5: Kappa Coefficient and accuracy assessment of ground truth pixels of ROI 4.

The first endmember, gravelly sandy loam, was identified on the eastern side of the ROI as an exposed surface within the Rio Yuca Formation. This endmember showed a shallow absorption feature at 500 nm due to a possible concentration of iron oxides. The next endmember identified relates to the braided stream flowing in a southeastern direction from the Merida Andes mountain. This endmember primarily matched with gray illite, or smectite bearing clays. These clays originate from the igneous bearing Cerro Pelon Formation that consists of both

mafic and felsic petrology. The presence of high water content, due to the endmembers' geomorphological setting, is represented in the absorption range of the spectra where there is a high drop off in the 2,300–2,400 nm range. This ROI is located in a coniferous forest region, in relation to ROI 1, this forest shares a similar reflectance in the chlorophyll a absorption range about 650–690 nm, and a near identical cell structure in the near infrared, with moderate water absorption at 1,400–1,900 nm. The last two endmembers identified in ROI 4 were sandy loam and silty loam. These two endmembers were located at exposed surfaces throughout the region where the silty loam was located prominently in the agricultural location, while the sandy loam was located along dirt roads and naturally exposed surfaces.

Reflectance Profile

A reflectance profile was constructed from a NDVI, (See Figure 17), to represent reflectance bands in comparison to vegetation health (Figure 18). The brighter the value, the healthier the vegetation in the NDVI image; the darker the value, the less healthy the vegetation will be in the image. The NDVI is showing a higher vegetation value in the Merida Andes mountain range, the irrigated reservoir agricultural network, and the Caribbean mountain system to the north. The Quaternary alluvium shows a lower signature of vegetation, while areas that have outwash from one of the two mountain ranges show a healthier NDVI. This profile shows an overall trend of a high reflectance in the Merida Andes mountain range, a slight decrease in vegetation health at the contact between the Merida Andes and the Quaternary alluvium at the 500 nm, a rise in vegetation health in agricultural locations supplied by reservoir irrigational support at the 1,200–1,600 nm range, and a subtle decrease in vegetation health in the northern Caribbean mountain system.

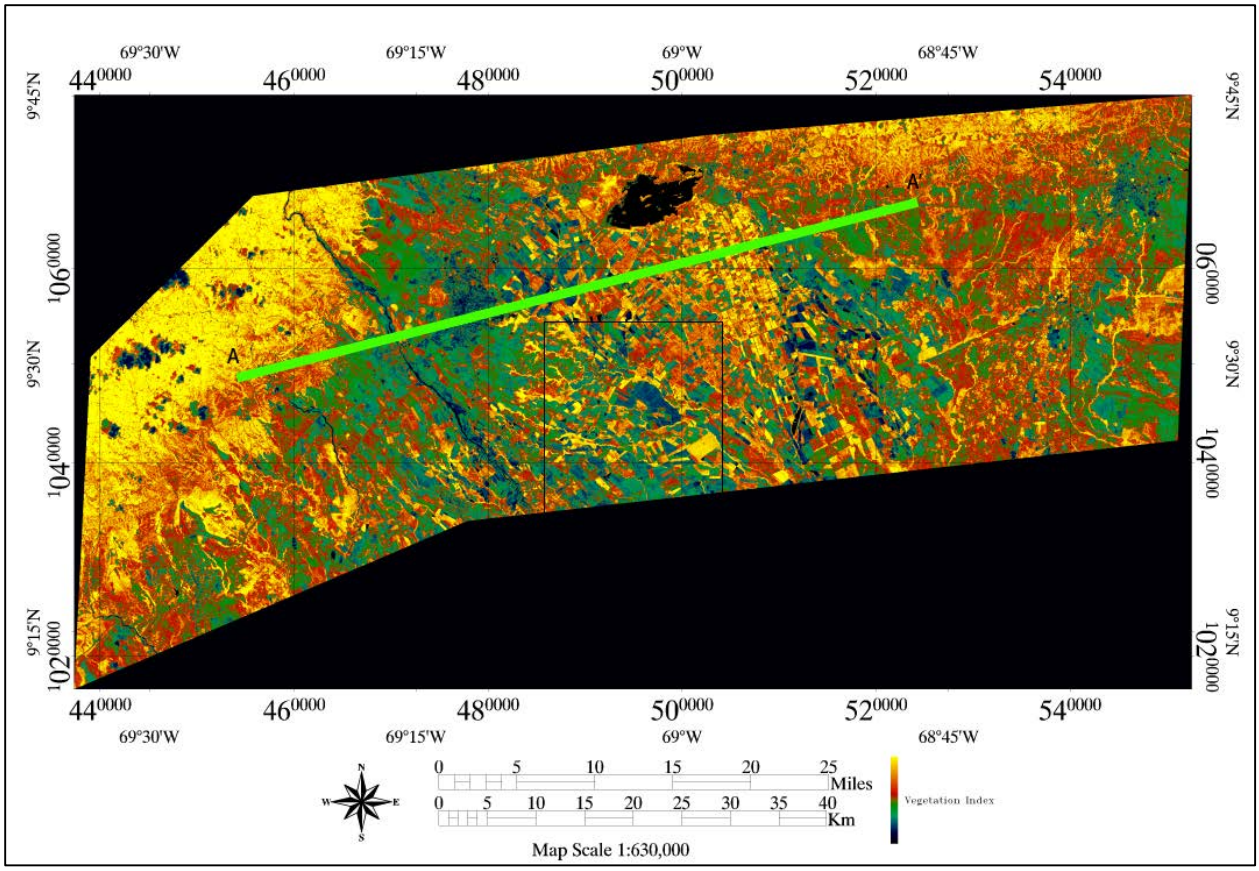

FIGURE 18: NDVI image of study area with A-A′ transect

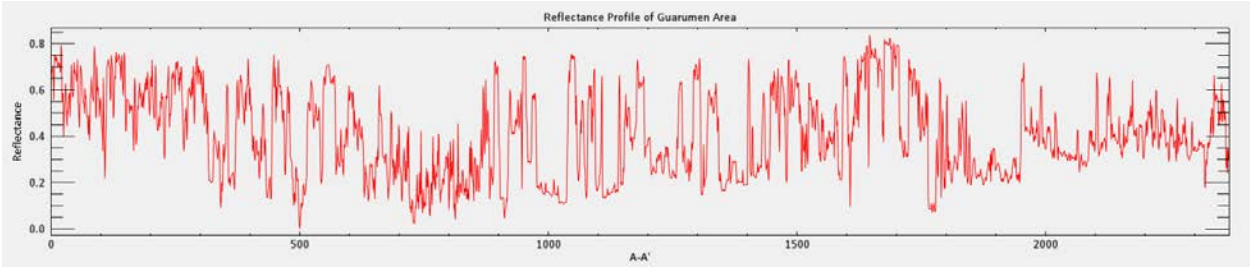

FIGURE 19: Reflectance profile from A-A’ of vegetation health

Microclimate Analysis

The algorithm was run for two identical Landsat 8 scenes from two different time frames. Each run of the algorithm with the two Landsat images was conducted uniformly for accuracy. The first scene was acquired on December 17, 2017 and the second scene was acquired on July 28, 2018. These two scenes were chosen based on climate, cloud coverage, and rainfall. Both Landsat scenes had limited cloud cover so a more accurate analysis could take place. The two scenes were also taken seven months and eleven days apart from each other to show reference for the different climates over the year. Lastly, the scenes were taken based on rainfall. In Venezuela, July has the highest amount of rainfall, while December is one of the more moderate to low precipitation months in the year (Lyon, 2002).

The December LST map showed a uniform temperature range throughout a majority of the scene. The scene had a mean accumulation of LST at 26 ˚C, a minimum of 7 ˚C, and max of 36 ˚C, with a standard deviation of 3.26 ˚C (Figure 21). Areas that were higher in elevation showed a lower LST over locations lower in elevation besides several temperature anomalies within the Quaternary alluvium showed a large decrease in temperature compared to its surrounding locations (Figure 20).

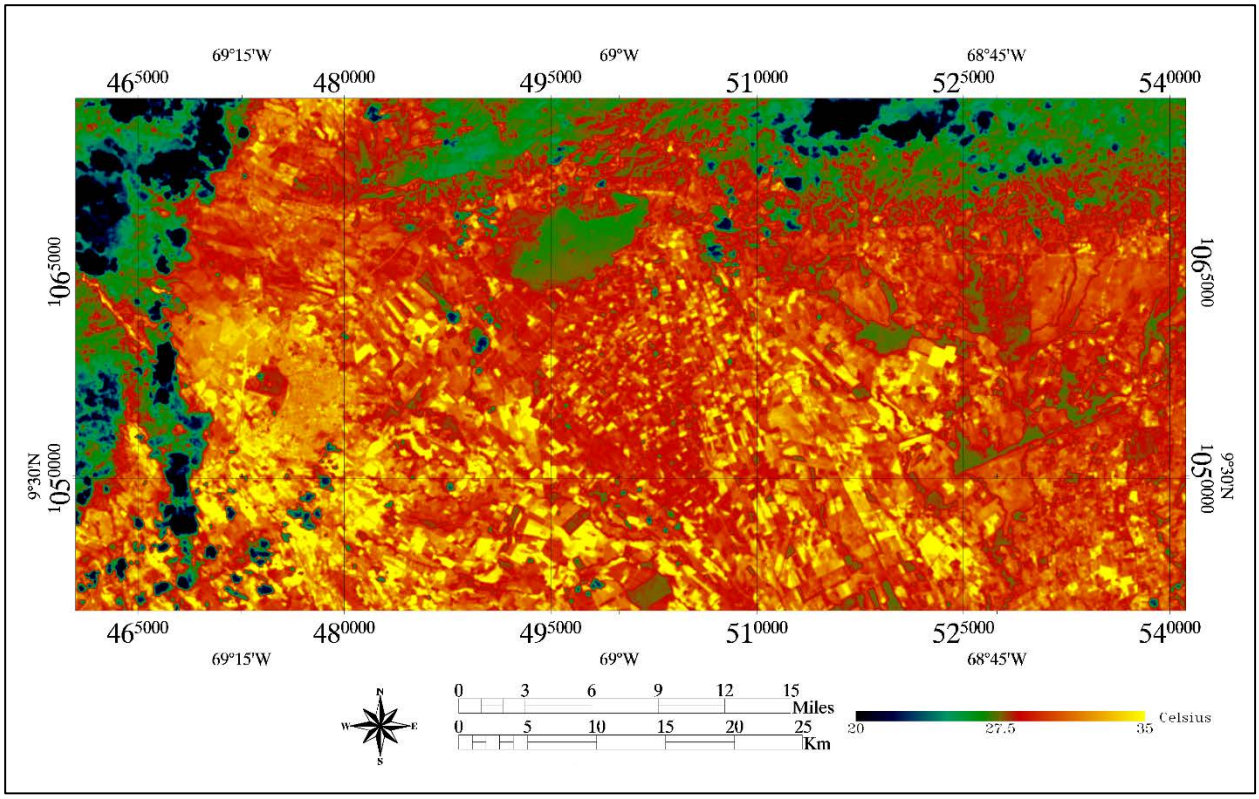

FIGURE 20: Land Surface Temperature map December 17, 2017

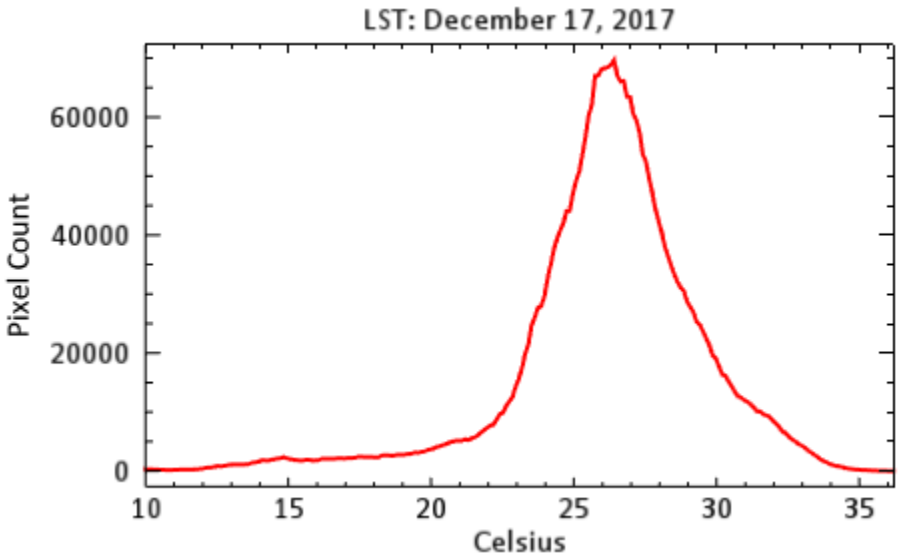

FIGURE 21: Histogram of temperature gradient compared to pixel data.

The July LST analysis showed a more spread out temperature range throughout the Landsat scene with a mean temperature of 28 ˚C, minimum temperature of 10 ˚C, maximum temperature of 70 ˚C, and a standard deviation of 3.26 ˚C. The LST map (Figure 22) represents a diverse distribution in temperature at both high and low elevations. The reservoir's agricultural irrigation network shows a decrease in LST, while the surrounding locations in the Quaternary alluvium have a higher temperature reading.

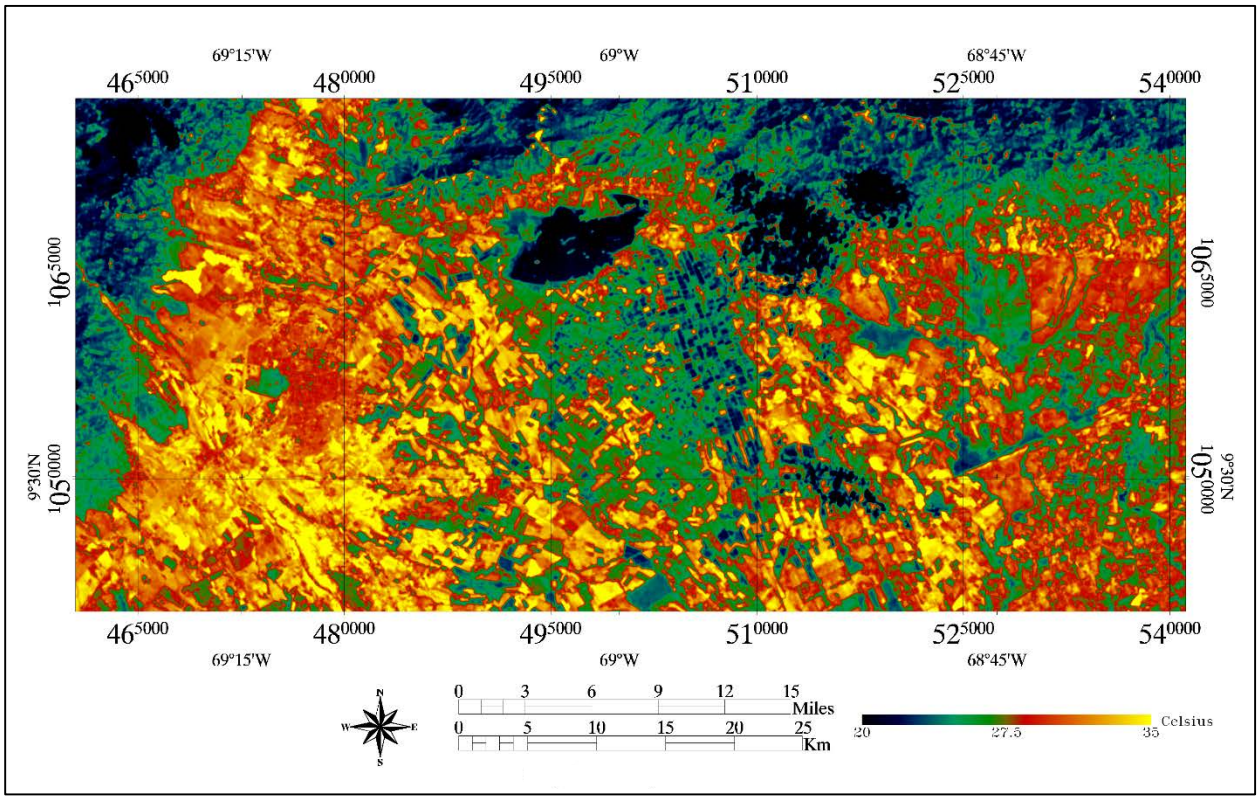

FIGURE 22: Land surface temperature map July 28, 2018.

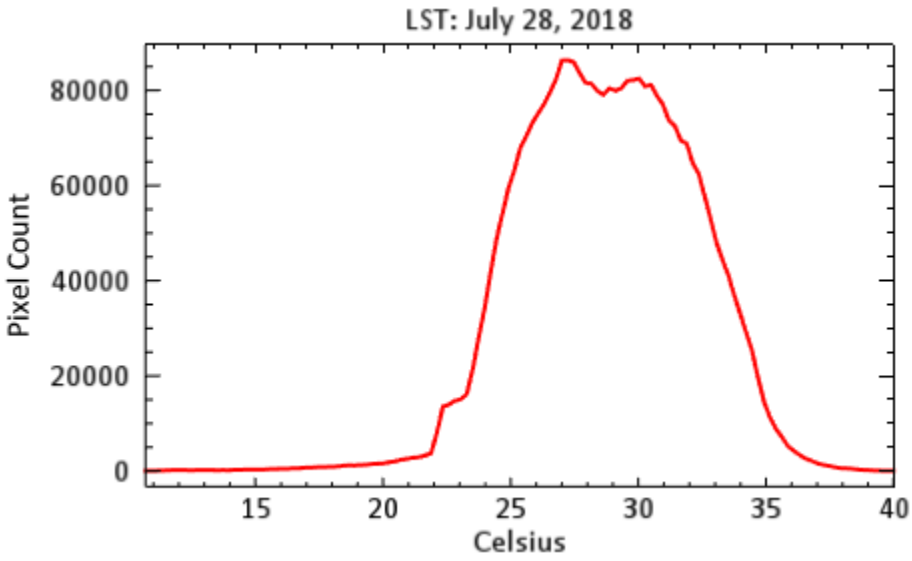

FIGURE 23: Histogram of temperature gradient compared to pixel data.

DISCUSSION & CONCLUSIONS

The high spectral resolution of hyperspectral imagery exceeds the suitability for extracting information in multivariate environments compared to a multispectral approach. Even though multispectral imagery has its many benefits, hyperspectral analysis is the most detailed approach to remotely sensed analysis. The higher spectral resolution of hyperspectral imagery improves the discrimination ratio of different physical components such as vegetation, water, soil, urbanization, and geological features. This research integrated both hyperspectral and multispectral imagery to conduct an in-depth analysis of how soil and vegetation anomalies interact with geologic applications. The Guarumen area of northwestern Venezuela is a structurally active region with sixteen different geologic formations located in the region. Its structural influence comes from the Merida Andes mountain system to the west and the Caribbean mountain system to the north. These two mountain systems both transport sediments to the alluvial basin, containing major cities such as Acarigua, Ospino, and Agua Blanca, which is the country's largest source of agricultural activity. Due to the high concentration of agricultural areas, it is difficult to identify the underlining geology and comprehend what is occurring without some form of in-depth analysis. The goal of this research was to use hyperspectral endmember analysis, normalized difference vegetation index reflectance profiling with Landsat imagery, and microclimate analysis with Landsat imagery. These methods were conducted to compose an in-depth quantitative analysis of the research area and identify how the soil and vegetation anomalies are interacting with the local geology, geomorphology, and

structural geology. By conducting this research, it is possible to further understand the capabilities of remote sensing applied to the geological sciences by utilizing hyperspectral and multispectral analyses without the accessibility of conducting a field study.

The endmember analysis was the primary method used for this research. The objective of this analysis was to identify endmembers in each of the four ROI and relate each to the respective region. Since the geological analysis scaled to the whole region of the study area, the endmembers classified within each ROI were an acceptable representation for the regions each represented. The most significant discovery from this analysis was how the fluvial geomorphology from the Merida Andes and the Caribbean mountain system affected the agricultural region in the Quaternary alluvium plain. Based on the spectra collected from the endmember analysis, ROI that identified in these geomorphologic regions showed an increase in vegetational health. This anomaly could be caused by the transport of micronutrients from sediment flows originated from either igneous or metamorphic formations in the mountainous regions surrounding the basin. This can be seen in the Merida Andes where several extrusive and intrusive igneous formations come in contact or originate with the stream patterns that are flowing into the basin. The Merida Andes sediment transport locations in the ROI were identified to be composed of gray illite/smectite bearing soils within the spectral readings, but traces of basalt were identified with the spectral analyst that led to speculation that basalt and/or granite could have been mixed in with this clayey transport. This can be supported by the presence of coniferous trees within the Merida Andes. In ROI 1 and 4, coniferous trees were spectrally referenced through the spectral analyst tool. Coniferous trees prefer to grow under acidic soil conditions (Martikainen & Boer, 1992). This helps support the spectral identification

of the sediment transport containing igneous bearing clay minerals due to the area's vegetation makeup and geologic history.

In the Caribbean mountain system region, the Mucaria, Agua Blanca, Araure, and Cojedes formations are the primary geologic components of the study area. These formations consist of Cretaceous carbonaceous shales, limestones, and conglomerates. At the southern portion of the formation, they begin to mix with the Quaternary alluvium basin that contains sands, gravels, and silts. This mixing of sediment is caused by a large deposit of an alluvial plain pushing sediments down elevation into the Quaternary basin. Identification of the composition of the soil was helped by the forest network that was located in the ROI. ROI 2 and 3 both consisted of a deciduous tree network that was identified through the endmember analysis. Both of the spectral reflectance bands shared close characteristics while varying in a slight difference in the 700–1,300 nm with the grass and coniferous spectral responses in the other ROI. These three types of spectral responses were best differentiated from image derivation and physically identifying the vegetation from the imagery. Deciduous trees prefer to grow under alkaline soil conditions due to their base soil contents. This can be seen in the ROI 3 spectra where the soils were spectrally identified as silty loam and clay loam. These two soil endmembers were identified with the spectral analyst tool and were matched to the selected soil types based on the comparison of the spectral libraries contents.

A high concentration of structural activity can be seen in ROI 3 through the GIS data; it shows a series of thrust faulting in an east-west direction with a northern uplift that has a transecting southwest-northeast inferred fault that is not visible in true color imagery. Minimal structural activity could be seen throughout the hyperspectral imagery even with convolution filters applied. Specific areas that were higher in elevation above the Quaternary alluvium

showed some structural significance in the imagery, but not enough to correlate a definite answer. This is primarily due to the heavy concentration of agriculture covering the majority of the study area.

The reflectance profile was an attempt to characterize vegetation health in the form of a "topographic profile." The results from this analysis came back favorable, with an interesting look at using a cross-sectional polyline to measure reflectance and correlate it with specific data. In this analysis, the reflectance was measuring vegetation health from A–A′. The cross-section showed a trend that supplemented the endmember analysis. It showed that the Merida Andes had a higher vegetation index value compared to the Caribbean Mountain System, which could support the idea of soil properties of both regions and how they affect vegetation. The central region also showed a high value, but this was supplemented to the reservoir irrigation system supplying water to the regional agriculture.

Measuring LST of the two Landsat scenes posed interesting finds of both data sets. Both sets of data shared similar results in terms of temperature. The 2017 dataset had a mean temperature of 28 ˚C and the 2018 data with a mean of 26 ˚C. This close range in temperature is most likely due to the tropical climate that Venezuela faces year-round where temperature changes minimally fluctuate throughout the year. Even though the majority of the region remains a consistent temperature year-round, the higher elevations of the Merida Andes and the Caribbean Mountains show a decrease in temperature. From the data, it is believed that elevation plays a key role in controlling temperature throughout the region. Both LST maps show a gradual decrease in temperature the lower the elevation, minus the anomalies in the lower elevation region that is contrast to the regional temperature. It is believed that the key changes from the 2017 and 2018 LST data are the amount of rainfall and how it affects the surface

temperature of the region. July is shown to have the highest precipitation on average, while December is shown to have a more moderate to low precipitation percentage in Venezuela (Lyon, 2002). It is shown that heavier precipitation cycles decrease surface temperature throughout the region. This can also be seen in the two histograms of both dates where the 2018 date has a greater range of temperatures due to chances of high precipitation rates.

One other factor that has an effect of the microclimate of the region is soil type. Different soil composition affects microclimates based on the amount of water retained or how quickly it evaporates. Soils that show a high concentration of clay retain more moisture compared to soils with a concentration of sands. This retention can affect humidity and temperature of a region (Ismangil et al., 2016). With this knowledge of soil properties, it is possible to identify soil classifications based on LST in a generalized classification scheme of water retention based on pore size, and thermal conductivity and heat capacity based on mineralogy.

Due to the region being highly agricultural, it was difficult to interpret soil or vegetation classifications. It was safely assumed that any region identified as agricultural could have been supplemented with groundwater pumping, controlled burning, foreign plant species, or distribution of foreign soil. This was taken into consideration during the image deriving process during the endmember analysis. This brought the challenge of identifying specific vegetation trends that could be identified as natural, along with exposed soil surfaces that could be identified as a “natural” occurrence. Most of these “natural” occurrences took place in higher elevation geomorphological regions such as the mountainous regions of both the Merida Andes and Caribbean Mountains, and a slight decline in elevation to the hills, alluvial valleys, and

piedmont regions. There were also exposed regions of unconsolidated sediments along rivers and streams in the area, especially ones that came down in elevation from both mountain systems.

Another limitation that occurred during the analysis was the cancellation of the field campaign. This was cancelled due to the political and economic turmoil that the country of Venezuela is facing. Too many safety issues arose weeks prior to the field campaign, and it was mutually decided to postpone the field campaign indefinitely. The rest of the analysis was to be conducted with remote sensing applications. With the limitation of field data in remote sensing applications, limitations are set on what can and cannot be done. Regions that have had minimal research conducted pose the problem of ground accuracy assessment due to the lack of physical evidence of specific plant species, and soil characterizations. Even with this limitation, this research gives insight on how limited field data, with a enough remotely sensed data, does make it possible to conduct an analysis remotely while still getting sufficient results to solve the original research objective.

To further answer the questions of the research objectives, a proper field campaign with ground truth data from a field spectrometer would help quantify the hyperspectral data by comparing the observed data versus the data collected by the sensor. This would help differentiate between geological and vegetation components through spectral mixture analysis. Another benefit would be to gather both soil and vegetation samples for in-situ field analysis, and analysis in a lab setting. This would further the precision of the classification and make it possible to classify to the level of plant species and mineralogical makeup of collected soil samples.

Nevertheless, the analysis completed did show that the endmembers, and LST, did correspond to real soil and vegetation components on the ground, and was able to provide a

quantitative “macro” scaled analysis of the Guarumen area by identifying large scale patterns in the data by correlating them to the geologic makeup of the area. This was possible given the limited field data by reducing the hyperspectral spectra to unique endmembers for classification, and conducting LST data to analyze the surface temperature to identify different microclimate properties affecting the area. It can be concluded that, though only limited ground truth information was available in the study area, it is possible to correlate geological phenomena through remote sensing applications.

LIST OF REFERENCES

Alirio Bellizzia, G. D. (1990). The Caribbean Mountain System, Northern South America; A Summary. In *The Geology of North America* (pp. 167–175). The Geological Society of America

Audemard M., Franck A.; Singer P., A. (1996). Active fault recognition in northwestern Venezuela and its seismogenic characterization: Neotectonic and paleoseismic approach. *Geofisica Internacional*, *35*(3), 245–255.

Audemard, F. E., & Audemard, F. A. (2002). Structure of the Merida Andes, Venezuela: Relations with the South America-Caribbean Geodynamic Interaction. *Tectonophysics*, *345*, 299–327.

Avdan, U., & Jovanovska, G. (2016). Algorithm for Automated Mapping of Land Surface Temperature Using LANDSAT 8 Satellite Data. *Journal of Sensors*, *2016*, 8.

Ball, J., & Student, G. (2006). Accuracy Analysis of Hyperspectral Imagery Classification Using Level Sets. Reno, Nevada: ASPRS.

Bateson, C. A., Asner, G. P., & Wessman, C. A. (2000). Endmember Bundles: A New Approach to Incorporating Endmember Variability into Spectral Mixture Analysis. *IEEE Transactions on Geoscience and Remote Sensing*, *38*(2), 1083–1094

Bellizzia, A., & Dengo, G. (1990). The Caribbean mountain system, northern South America; a summary. *The Caribbean region. The Geological Society of America, Boulder, CO*, 167-175.

Boardman, J.W., Kruse, E. A. (1994). Automated Spectral Analysis: A Geological Example Using AVIRIS Data. *Procedings of the Tenth Thematic Conference on Geologic Remote Sensing: Exploration, Environment, and Engineering*, 1407–1418.

Case, J. E. (1980). Crustal Setting of Mafic and Ultramafic Rocks and Associated Ore Deposits of the Caribbean Region. Menlo Park, California. Retrieved from

David Ismangil, Daniel Wiegant, Eyasu Hagos, Frank Van Steenbergen, Matthijs Kool, Francesco Sambalino, Giulio Castelli, Eleca Bresci, F. H. (2016). *Managing the Microclimate*.

Dehaan, R. (2014). Image-derived spectral endmember as indicators of salinization. *International Journal of Remote Sensing*, *24*(November 2003), 775–794. https://doi.org/10.1080/01431160110107635

Ellis, J. M., Davis, H. H., & Quinn, M. B. (2005). Airborne Hyperspectral Imagery for the Petroleum Industry (pp. 6–8). Fourteenth International Conference on Applied Geologic Remote Sensing.

Eoardmanl, J. W., Kruselj, F. A., & Green, R. O. (1993). *Mapping Target Signatures Via Partial Unmixing of AVIRIS Data*. Boulder, Colorado.

Erikson, P. Johan, Kelley, Shari A., Osmolovsky, Peter, Verosub, K. L. (2012). Linked Basin Sedimentation and Orogenic Uplift: The Neogene Barinas Basin Sediments Derived from the Venezuelan Andes. *Journal of South American Earth Sciences*, *39*, 138–156.

Fabiola Ruiz, M. A. (2013). Characterization of a heavy oil reservoir combining multiattribute analysis and spectral decomposition for density prediction, Maturín Sub-basin, Venezuela (Vol. 38). Canadian Society of Exploration Geophysicists.

Geudtner, D. (2014). Implementation of the TOPS Mode on RADARSAT-2 in Support of the Copernicus Sentinel-1 Mission. RADARSAT-2 TOPS SAR Interferometry (InSAR) Scene Pair Data Acquisitions. *ESA Tecnical Note*, *1*(25 February 2015), 1–17.

Ghulam, A., & Louis, S. (2010). Calculating surface temperature using Landsat thermal imagery. Saint Louis University.

González, Carlos, Resano, Javier, Mozos, Daniel, Plaza, Antonio, Valencia, D. (2010). FPGA Implementation of the Pixel Purity Index Algorithm for Remotely Sensed Hyperspectral Image Analysis. *EURASIP Journal on Advances in Signal Processing*, 13.

Govender, M., Chetty, K., Naiken, V., & Bulcock, H. (2008). A comparison of satellite hyperspectral and multispectral remote sensing imagery for improved classification and mapping of vegetation. *South African Water Research Commission*, *34*(2), 147–154.

Govender, M., & Chetty, K. T. (2007). A Review of Hyperspectral Remote Sensing and its Application in Vegetation and Water Resource Studies. *Water S.A.*, (January 2015). https://doi.org/10.4314/wsa.v33i2.49049

Green, R. O. (1998). Summaries of the Seventh JPL Airborne Earth Science Workshop. NASA (Vol. 1–409). Pasadena, California.

Hamzeh, S., Naseri, A. A., Alavipanah, S. K., Mojaradi, B., Bartholomeus, H. M., Clevers, J. G. P. W., & Behzad, M. (2012). Estimating salinity stress in sugarcane fields with spaceborne hyperspectral vegetation indices. *International Journal of Applied Earth Observations and Geoinformation*, (April). https://doi.org/10.1016/j.jag.2012.07.002

Harris, J. R., Rogge, D., Hitchcock, R., Ijewliw, O., & Wright, D. (2006). Mapping lithology in Canada's Arctic: application of hyperspectral data using the minimum noise fraction transformation and matched filtering 1. *NRC Canada*, *2193*(2005), 2173–2193. https://doi.org/10.1139/E05-064

Herweg, J. A., Kerekes, J. P., Weatherbee, O., Messinger, D., Aardt, J. Van, Ientilucci, E., … Raque, N. (2012). SpecTIR Hyperspectral Airborne Rochester Experiment Data Collection

Campaign. *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery*, *8390*, 1–10. https://doi.org/10.1117/12.919268

Hytla, P. C. (2009). Anomaly detection in hyperspectral imagery: comparison of methods using diurnal and seasonal data. *Journal of Applied Remote Sensing*, *3*(1), 33546. https://doi.org/10.1117/1.3236689

Jácome, M., & Chacín, L. (2008). Subsidence Analysis of the Barinas-Apure Basin : Western Venezuela (Vol. 10718, pp. 682–686). CSPG CSEG CWLS Convention.

James, K. H. (2000). The Venezuelan Hydrocarbon Habitat, Part 2: Hydrocarbon Occurrences and Generated-Accumulated Volumes. *Journal of Petroleum Geology*, *23*(April), 133–164.

Jia, L., Wang, W., Zhou, J., Menenti, M. (2010). Detecting Short-and Long-Term Vegetation Anomalies in Response to Drought and Climate Change by Satellite Observations of Land Surface. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science*, *XXXVIII* (8), 201–202.

Johan P. Erikson, Shari A. Kelley, Peter Osmolovsky, K. L. V. (2012). Linked basin sedimentation and orogenic uplift: The Neogene Barinas basin sediments derived from the Venezuelan Andes. *Journal of South American Earth Sciences*, *39*, 138–156.

John B. Adams, Milton O. SMith, P. E. J. (1986). Spectral mixture modeling: A new analysis of rock and soil types at the Viking Lander 1 Site. *Journal of Geophysical*

Joseph W. Boardman, Fred A. Kruse, R. O. G. (1993). Mapping Target Signatures Via Partial Unmixing of AVIRIS data.

Karim Saheb Ettabaa, M. B. S. (2016). Anomaly Detection in Hyperspectral Imagery: An Overview. In I. Global (Ed.), *Handbook of Research on Emerging Perspectives in Intelligent Pattern Recognition, Analysis, and Image Processing* (p. 127). Handbook of Research on Emerging Perspectives in Intelligent Pattern Recognition, Analysis, and Image Processing.

Kruse, F. A., Boardman, J. W., & Huntington, J. F. (2003). Comparison of Airborne Hyperspectral Data and EO-1 Hyperion for Mineral Mapping. *IEEE Transactions on Geoscience and Remote Sensing*, *41*(6), 1388–1400.

Lili Zhang, C. Z. (2017). Hyperspectral anomaly detection based on spectral–spatial background joint sparse representation. *European Journal of Remote Sensing*, *50*(1), 362–376.

Lyon, B. (2002). Enhanced Seasonal Rainfall in Northern Venezuela and the Extreme Events of December 1999. *Journal of Climate*, *16*.

Marina Stathopoulou, C. C. (2006). Daytime Urban Heat Islands from Landsat ETM+ and Corine Land Cover Data: An Applicaiton to Major Cities in Greece. *Solar Energy*, *81*(3), 358–368.

Martikainen, Pertii, De Boer, W. (1992). Nitrous Oxide Production and Nitrification in Acidic Soil from a Dutch Coniferous Forest. *Soil Biology and Biochemistry*, *25*(3), 343–347.

Martínez, P. J., Pérez, R. M., Plaza, A., Aguilar, P. L., Cantero, M. C., & Plaza, J. (2006). Endmember extraction algorithms from hyperspectral images. *Annals of Geophysics*, *49*(February), 93–101.

McHugh, M. L. (2012). Interrater Reliability: The Kappa Statistic. *Biochemia Medica*, *22*, 276–282.

Mura, M. D., Chanussot, J., Plaza, A., Mura, M. D., & Persyval, L. (2014). *An Overview on Hyperspectral Unmixing*.

Rajeshwari, A., & Mani, N. D. (2014). Estimation of Land Surface Temperature of Dindigul District Using Landsat 8 Data. *International Journal of Research in Engineering and Technology*, *3*(4), 122–126

Rasti, B., Scheunders, P., & Ghamisi, P. (2018). Noise Reduction in Hyperspectral Imagery: Overview and Application. Remote Sensing, 3, 1–28.

Rwanga, S. S., & Ndambuki, J. M. (2017). Accuracy Assessment of Land Use / Land Cover Classification Using Remote Sensing and GIS. *International Journal of Geosciences*, 611–622. https://doi.org/10.4236/ijg.2017.84033

Sarmiento, G., Monasterio, M. (1969). Studies on the savanna vegetation of the Venezuelan Llanos. I. The use of association analysis. *Journal of Ecology*, *57*, 579–598.

Schubert, C. (1982). Neotectonics of Boconó Fault, Western Venezuela. *Tectonophysics*, *85*(3–4), 205–220. https://doi.org/https://doi.org/10.1016/0040-1951(82)90103-2

Sun, Qinqin, Wu, Zhifeng, Tan, J. (2012). The Relationship Between Land Surface Temperature and Land Use/Land Cover in Guangzhou, China. *Environmental Earth Sciences*, 1687–1694.

Tian, Q. (2012). Study on Oil-Gas Reservoir Detecting Methods Using Hyperspectral Remote Sensing. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XXII ISPRS Congress, XXXIX-B7*(September), 157–162. https://doi.org/10.5194/isprsarchives-XXXIX-B7-157-2012

Weng, Q., Lu, D., & Schubring, J. (2004). Estimation of land surface temperature – vegetation abundance relationship for urban heat island studies. *Remote Sensing of Environment*, *89*, 467–483. https://doi.org/10.1016/j.rse.2003.11.005

Winter, M. E. (1999). N-FINDR: An Algorithm for Fast Autonomous Spectral End-Member Determination in Hyperspectral Data. SPIE.

VITA

Tyler Ricketts received a Bachelor's Degree in Geology from the University of Mississippi in May, 2017. He worked as the department's Assistant Administrative Coordinator from August 2013 to May 2017, a drill site geologist for EP Minerals contracted through Mississippi Mineral Resource Institute from November 2016 to March 2017 and an undergraduate teaching assistant from August 2013 to May 2017. He also served as the Universities Vice President of the Association of Environmental and Engineering Geologists from August 2014 to May 2015, and President from August 2015 to May 2016. He was a finalist for the Department of Energy Geothermal Energy Student Competition in October, 2014. He has presented and published research at Geological Society of America, Association of Environmental and Engineering Geologists, and PECORA 20 conferences.