Assessment of Remotely Sensed Image Processing Techniques for Unmanned Aerial System (Uas) Applications

Christopher Michael Zarzar
Assessment of remotely sensed image processing techniques for
Unmanned Aerial System (UAS) applications

By
Christopher M. Zarzar

A Dissertation
Submitted to the Faculty of
Mississippi State University
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
in Earth and Atmospheric Sciences
in the Department of Geosciences

Mississippi State, Mississippi

August 2017
Assessment of remotely sensed image processing techniques for
Unmanned Aerial System (UAS) applications

By

Christopher M. Zarzar

Approved:

________________________
Jamie L. Dyer
(Major Professor)

________________________
Padmanava Dash
(Committee Member)

________________________
Andrew E. Mercer
(Committee Member)

________________________
Robert J. Moorhead
(Committee Member)

________________________
Qingmin Meng
(Committee Member)

________________________
Renee M. Clary
(Graduate Coordinator)

________________________
Rick Travis
Interim Dean
College of Arts & Sciences
Unmanned Aerial Systems (UASs) offer a new era of local-scale environmental monitoring where access to invaluable aerial data no longer comes at a substantial cost. This provides the opportunity to vastly expand the ability to detect natural hazards impacts, observe environmental conditions, quantify restoration efforts, track species propagation, monitor land surface changes, cross-validate existing platforms, and identify hazardous situations. While UASs have the potential to accelerate understanding of natural processes, much of the research using UASs has applied current remote sensing image processing techniques without questioning the validity of these in UAS applications. With new scientific tools comes a need to affirm that previous techniques are still valid for the new systems. To this end, the objective of the current study is to provide an assessment regarding the use of current remote sensing image processing techniques in UAS applications. The research reported herein finds that atmospheric effects have a statistically significant impact on low altitude UAS imagery. Correcting for these external factors affecting the imagery was successful using an empirical line calibration (ELC) image correction technique and required little modification for use in a
complex UAS application. Finally, it was found that classification performance of UAS imagery was reliant on training sample size more than classification technique, and that training sample size requirements are larger than previous remote sensing studies suggest.
ACKNOWLEDGEMENTS

It was a once in a lifetime experience working under the guidance of Dr. Jamie L. Dyer. He is a one-of-a-kind advisor and mentor who has gifted me with opportunities and adventures ranging from Alaska to Louisiana, all beyond my imagination. I am grateful for the support and guidance of Dr. Robert J. Moorhead throughout my time at Mississippi State University. I also want to thank my committee, Dr. Robert J. Moorhead, Dr. Andrew E. Mercer, Dr. Padmanava Dash, and Dr. Qingmin Meng, for their help making this dissertation successful. This dissertation would not have been possible without the support of the Northern Gulf Institute. I want to acknowledge those at NGI who made this work possible through their valuable time, ideas, and support including Gray Turnage, Lee Hathcock, Sathish Samiappan, David Young, Sean Meacham, Louis Wasson, and John Cartwright.

A special thanks to my mother, Krista Zarzar, for her editorial expertise and commentary. I also want to thank my family and friends for all their love, support, and assistance throughout this dissertation process.
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS ................................................................................................................ ii

LIST OF TABLES ........................................................................................................................ v

LIST OF FIGURES ......................................................................................................................... vi

CHAPTER

I. INTRODUCTION ....................................................................................................................... 1

II. QUANTIFYING AND CORRECTING ATMOSPHERIC IMPACTS ON UNMANNED AERIAL SYSTEM IMAGERY .................................................................................. 6

   Literature Review ..................................................................................................................... 6
   North Farm Experiment ........................................................................................................... 11
   Image Correction ..................................................................................................................... 16
       Study Area ............................................................................................................................ 16
       Materials and Methods ........................................................................................................ 17
           Unmanned Aerial System ................................................................................................. 17
           Calibration Targets ........................................................................................................ 19
           Field Operations .............................................................................................................. 21
           Flight Data ....................................................................................................................... 23
           Empirical Line Calibration .............................................................................................. 23
   Results .................................................................................................................................... 26
       Image Correction ................................................................................................................ 26
   Discussion ............................................................................................................................... 29
   Conclusions ............................................................................................................................. 32

III. LAND SURFACE CLASSIFICATION OF UNMANNED AERIAL SYSTEM IMAGERY .................................................................................................................. 34

   Literature Review .................................................................................................................. 34
   Methods ................................................................................................................................. 39
       UAS Imagery ....................................................................................................................... 39
       Image Preprocessing ......................................................................................................... 41
       Image Classification ......................................................................................................... 41
       Classification Tuning ......................................................................................................... 44
   Results ................................................................................................................................. 45
   Discussion ............................................................................................................................. 57
# LIST OF TABLES

2.1 UAS Specifications ................................................................. 12
2.2 UAS Specifications ................................................................. 18
2.3 Laboratory-Measured Band Reflectance ................................. 20
2.4 UAS Flight Missions ............................................................... 23
2.5 Regression Results ................................................................. 27
3.2 UAS Specifications ................................................................. 40
3.3 UAS Flight Missions ............................................................... 41
3.4 Training and Verification Samples Breakdown ......................... 43
3.5 August Classification Median Accuracies .............................. 49
3.6 Mission Classification Median Accuracies .............................. 55
<table>
<thead>
<tr>
<th>Number</th>
<th>Figure Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Camera Spectral Response</td>
<td>12</td>
</tr>
<tr>
<td>2.2</td>
<td>GPS Response</td>
<td>14</td>
</tr>
<tr>
<td>2.3</td>
<td>Canon Plotted Confidence Intervals</td>
<td>14</td>
</tr>
<tr>
<td>2.4</td>
<td>MicaSense Plotted Confidence Intervals</td>
<td>15</td>
</tr>
<tr>
<td>2.5</td>
<td>Study Area</td>
<td>17</td>
</tr>
<tr>
<td>2.6</td>
<td>Canon Camera Spectral Response</td>
<td>18</td>
</tr>
<tr>
<td>2.7</td>
<td>Laboratory-Measured DHR</td>
<td>20</td>
</tr>
<tr>
<td>2.8</td>
<td>Calibration Tarp Deployment</td>
<td>21</td>
</tr>
<tr>
<td>2.9</td>
<td>Data Processing Workflow</td>
<td>22</td>
</tr>
<tr>
<td>2.10</td>
<td>Spectroradiometer-UAS Scatterplot</td>
<td>24</td>
</tr>
<tr>
<td>2.11</td>
<td>August Vegetation Spectral Response</td>
<td>28</td>
</tr>
<tr>
<td>2.12</td>
<td>Site Calibration Equation Comparison</td>
<td>30</td>
</tr>
<tr>
<td>2.13</td>
<td>Example of Image Extraction</td>
<td>32</td>
</tr>
<tr>
<td>3.1</td>
<td>Study Area</td>
<td>39</td>
</tr>
<tr>
<td>3.2</td>
<td>Random Forest Classifier Tuning</td>
<td>45</td>
</tr>
<tr>
<td>3.3</td>
<td>Classification Confidence Interval Plots</td>
<td>46</td>
</tr>
<tr>
<td>3.4</td>
<td>Classification Confidence Interval Plots</td>
<td>47</td>
</tr>
<tr>
<td>3.5</td>
<td>August Classification Accuracy Metrics</td>
<td>50</td>
</tr>
<tr>
<td>3.6</td>
<td>August Class Spectral Response</td>
<td>51</td>
</tr>
<tr>
<td>3.7</td>
<td>August Logistic Regression Classification</td>
<td>53</td>
</tr>
</tbody>
</table>
3.8 Mission Mosaic Classifications .................................................................54
3.9 Mission Classification Accuracy Metrics ...............................................56
3.10 Mission CIR Imagery ........................................................................57
3.11 Zoomed Classification View ..................................................................60
CHAPTER I
INTRODUCTION

The Unmanned Aerial System (UAS) offers a new era of local-scale environmental monitoring where access to invaluable aerial data no longer comes at a substantial cost. This provides the opportunity to vastly expand our ability to detect natural hazards impacts, observe environmental conditions, quantify restoration efforts, track species propagation, monitor land surface changes, cross-validate existing platforms, and identify hazardous situations (Aanstoos et al. 2010; Adams and Friedland 2011; Anderson and Gaston 2013; Barreiro et al. 2015; Casella et al. 2016; Chou et al. 2010; Erena et al. 2016; Feng et al. 2015; Gómez-Candón et al. 2014; Frew et al. 2012; Hardin and Jensen 2011; Hunt et al. 2010; Husson et al. 2016; Ollero et al. 2006; Torres-Sánchez et al. 2013; Zaman et al. 2011; Wallace et al. 2012; Martínez-de Dios et al. 2006; Li et al. 2012). UASs are a revolutionary research tool because they can fly with multispectral and hyperspectral imaging sensors at highly flexible temporal frequency (Lin et al. 2011; Wallace et al. 2012; Park 2015). Early UAS applications were mainly focused on the use of these systems as rapid response observation tools (Martínez-de Dios et al. 2006; Ollero et al. 2006; Changchun et al. 2012; Chou et al. 2010). Martínez-de Dios et al. (2006) experimented with the development of a methodology for the automated detection and monitoring of fire using UASs. Related to this is the possibility of using UASs as forest fire fighting tools to assess vegetation stress and risk-index of
fire. In addition, if a forest fire is occurring, the UAS can be used to fly in areas that are not safe for manned flight to monitor the fire evolution and estimate the burnt areas (Ollero et al. 2006). Chou et al. (2010) demonstrated how UAS imagery from Typhoon Morakot provided insight about the impacts of the typhoon. From the imagery, it was possible to estimate new collapsed lands and damaged structures, which is valuable information for emergency rescue efforts.

Image correction is not required for these basic rapid response operations, but image correction is required prior to using multitemporal data for classifications, or when accurate spectral information about features is required (Song et al. 2001). Therefore, it is important to understand to what degree current remote sensing assumptions and image correction techniques can be applied to UAS imagery. Some of the more advanced image correction techniques, like radiative transfer models, have been specifically developed for satellites. Current airborne image correction techniques often rely on the method of regressing image pixels with corresponding known ground truth data, termed the empirical line calibration (ELC) (Smith and Milton 1999; Moran et al. 2001b; Karpouzli and Malthus 2003; Kelcey and Lucieer 2012; Wang and Myint 2015). The ELC method has been successfully employed to correct UAS imagery (Kelcey and Lucieer 2012; Wang and Myint 2015); however, these studies were focused on building frameworks for UAS ELC image corrections under ideal conditions and were conducted at only a couple of altitudes (50 m & 100 m: Kelcey and Lucieer 2012; 10 m & 20 m: Wang and Myint 2015). Therefore, these studies failed to investigate whether the atmosphere impacted the UAS imagery to a degree that required image correction in the first place. In addition, the studies converted image values from digital numbers (DNs) directly to reflectance. This
limits the temporal application of the calibration equations developed because solar illumination properties are enveloped into the calibration equations. This means the development of new calibration equations would be required for separate times of the year because the seasonal variations in solar illumination cannot be corrected independent from the calibration equations. Thus, there is a need to first assess whether the atmosphere significantly impacts relatively low altitude UAS imagery. If so, then the proposed frameworks from Kelcey and Lucieer (2012) and Wang and Myint (2015) are valid for ideal conditions, but there is still the need to investigate the performance of ELC techniques in real-world UAS applications.

A more recent but natural application of high spatial resolution UAS imagery is for the classification of surface features. Authors have demonstrated that the use of UAS imagery in land surface classifications can produce classification accuracies comparable with manned aerial systems (MAS) and satellites (Casado et al. 2016; Casado et al. 2015; Laliberte et al. 2011; Laliberte and Rango 2009; Husson et al. 2016). Feng et al. (2015) demonstrated the effectiveness of using UASs to distinguish features in an urban area using UAS true color imagery in a Random Forest (RF) and a maximum likelihood classification technique. Husson et al. (2016) went a step further by successfully using UAS imagery to identify aquatic vegetation growth-form and dominant-taxon in a complex coastal estuary using both a threshold and RF classification techniques. With UAS imagery there is also the opportunity for greater classification precision and accuracy because the higher resolution imagery decreases spectral mixing (Casado et al. 2016). Using an Artificial Neural Network (ANN) classification technique, Casado et al. (2016) determined that accuracy increased with higher resolution (2.5 cm) compared to
lower resolution (5 cm and 10 cm) images. Possibly most important is the finding by Casado et al. (2016) that even in a small study area (a 1.4 km river reach), processing time was substantially increased with increased image resolution. This offers insight into the balance between classification accuracy and practical analysis time. However, similar to all previously mentioned UAS studies, the Casado et al. (2016) study is forcing UAS into current remote sensing image processing techniques while comparing accuracies from a single classification instance. This does not take into account the range of equally likely outcomes associated with slight variations to variables such as image preprocessing technique, training sample configuration, and classification technique. Therefore, there is a need for an investigation of how current remote sensing classification techniques perform in UAS image classification.

Where collection of remote sensing imagery is often conducted during near ideal conditions (i.e. near solar zenith and minimal clouds), UASs offer the opportunity to collect data during inclement weather conditions. There is valuable data to be collected during these conditions, such as the extent of flooding soon after a precipitation event (Di Baldassarre et al. 2009; Schumann et al. 2009; Matgen et al. 2011; Pulvirenti et al. 2011; Hostache et al. 2012; Schumann et al. 2011). However, the benefits of the rapid deployment capability bring new challenges because many UAS applications will not occur during the ideal conditions under which many remote sensing image processing techniques were developed. This will likely require modification to image processing techniques to account for the challenges that will be faced in these UAS applications.

While UASs have the potential to accelerate understanding of natural processes, it is important to determine whether current remote sensing image processing techniques
can be applied to UAS application. Much of the research using UASs has applied standard remote sensing image processing techniques without questioning the validity of these techniques in various UAS applications. With new scientific tools comes a need to affirm that previous techniques are still valid for the new systems. To this end, the goal of the current study is to provide an assessment of whether current remote sensing image processing techniques can be applied to UAS imagery. With this understanding, future research can work to build new techniques where necessary, or confidently continue to apply remote sensing image processing techniques that are valid in UAS applications.
CHAPTER II
QUANTIFYING AND CORRECTING ATMOSPHERIC IMPACTS ON UNMANNED AERIAL SYSTEM IMAGERY

Literature Review

Unmanned Aerial Systems (UASs) provide a more practical and cost-effective approach to data collection in field operations. UASs do not face the same challenges associated with satellite or manned aircraft because UASs can be flown as needed with the ability to fly under the cloud deck; this is especially important in quick response imagery applications. An important question is whether UAS imagery is bound by the same environmental factors that require atmospheric correction and radiometric calibration of satellite and manned aircraft system imagery (Che and Price 1992; Roberts et al. 1986; Farrand et al. 1994; Clark et al. 1995; Dwyer et al. 1995; Ferrier and Trahir 1995; Ferrier and Wadge 1996; Smith and Milton 1999; Moran et al. 2001b; Karpouzli and Malthus 2003; Kelcey and Lucieer 2012; Wang and Myint 2015; Gordon 1997; Hu et al. 2001; Melack and Gastil 2001; Hu 2002; Dash et al. 2012). If so, then an equally important question is whether current atmospheric image correction techniques can be applied to UAS imagery.

Two environmental factors that need to be accounted for in remotely sensed imagery include (1) absorption and scattering of light by atmospheric constituents, and (2) surface illumination geometry (Hu 2002). Atmospheric correction and radiometric
calibration techniques are methods used to account for these environmental impacts on remotely sensed imagery. Atmospheric correction and radiometric calibration methods have evolved from the more rudimentary flat field and image average methods (Roberts et al. 1986; Kruse 1988) to the more rigorous use of radiative transfer models (Farrand et al. 1994; Clark et al. 1995; Green 1998; Melack and Gastil 2001; Clark et al. 2002; Guanter et al. 2006). A detailed review of the numerous correction and calibration methods is given by Gao et al. (2009), so the current review will focus on four of the most common correction and/or calibration techniques.

An advantage of the flat field and image average calibration methods is they require only information already available in the image (e.g. roads and concrete). This ease of application comes with the downside of relatively poor results in depicting the true pixel value of a surface feature. The empirical line calibration (ELC) method adds more skill to the previous methods by developing a regression from in situ spectroradiometer measurements and airborne sensor measurements of calibration targets (Smith and Milton 1999; Moran et al. 2001b; Karpouzli and Malthus 2003; Kelcey and Lucieer 2012; Wang and Myint 2015). This regression is applied to the remaining pixels to nudge the data to look more like the controlled and often more advanced spectroradiometer data. A benefit of this ELC method is that it can combine the atmospheric correction and radiometric calibration into a single calibration equation.

More complex atmospheric correction and radiometric calibration methods have since been developed. One such method uses a systematic process of subtracting the contributions of atmospheric constituents and water Fresnel reflection properties (Gordon 1997; Hu et al. 2001; Melack and Gastil 2001; Hu 2002; Dash et al. 2012). A limitation
of this method is that the resulting corrections are sensor-specific. This method is also limited to only providing atmospheric correction of the imagery, so a calibration of the sensor imagery may still be required to normalize the data for illumination (Filippi et al. 2006; Dash et al. 2012).

Another robust and widely used correction method uses radiative transfer models to simulate the interactions between the atmosphere and radiation. Results from radiative transfer models prove to be most accurate when the atmospheric-radiation parameters at the time of the imagery can be accurately quantified (Roberts et al. 1986; Farrand et al. 1994; Clark et al. 1995; Dwyer et al. 1995; Ferrier and Trahair 1995; Ferrier and Wadge 1996). While monitoring systems aid in the quantification of the radiative transfer parameters (Holben et al. 1998), the effort put into this method does not always translate to substantial improvements over simpler methods. Therefore, the ELC method has proved to be a compromise of these common correction and calibration methods by providing a rather straightforward process that can render accurate image values.

UASs are typically flown at relatively low altitudes (<350 m). This reduces the need for radiative transfer models because of the relatively small column of atmosphere through which reflected light travels before reaching the sensor. As a result, studies have relied on ELCs for the correction and calibration of the UAS imagery (Kelcey and Lucieer 2012; Wang and Myint 2015). The foundation of an ELC is the development of calibration equations from the regression of *in situ* measurements and remotely sensed data. An ELC requires one or more calibration targets to be present in the imagery. Ground spectroradiometer measurements of these calibration targets are taken in tandem with the sensor flyover. The sensor spectral response is applied to the ground
spectroradiometer data, thereby converting the ground spectroradiometer data into ground reference data specific to that sensor flown. For each sensor wave band, a regression between spectroradiometer ground reference data and the sensor data produce the calibration equations. These equations are applied to the remaining imagery to predict sensor radiance or reflectance values, depending on whether adjustments for illumination are included in the process (radiance) or not (reflectance).

Arguably, the most important component of the ELC is the proper selection of calibration targets. There are some requirements that must be met when selecting calibration targets: (1) the target should be large enough to fill the sensor instantaneous field of view (IFOV), (2) the target should be orthogonal to the sensor, (3) the target should have Lambertian reflectance properties, and (4) the target should be free from vegetation (Che and Price 1992). Chemically-treated and laboratory-calibrated tarps make for great calibration targets in an ELC (Moran et al. 2001a); however, these can be impractical depending on the size of the sensor IFOV. Some studies have relied on targets such as black asphalt and concrete for use with remotely-sensed satellite and manned aircraft systems (Smith and Milton 1999; Karpouzli and Malthus 2003). There are problems with using these man-made surfaces because of the lack of control that may surface as noise in the data, such as light-colored compounds in certain asphalts (Smith and Milton 1999).

Another consideration is how the number of calibration targets impacts the accuracy of the ELC. An assumption of early ELC was that a linear relationship existed between remotely sensed radiance and ground reflectance. This required that at least two calibration targets be used; however, numerous authors have found greater accuracy
comes from increasing the number of calibration targets (Farrand et al. 1994; Price et al. 1995; Smith and Milton 1999; Karpouzli and Malthus 2003; Xu and Haung 2008; Kelcey and Lucieer 2012; Wang and Myint 2015). In a case study over Thorney Island, England, Smith and Milton (1999) found that the use of three calibration targets greatly improved the accuracy of the ELC compared to the use of two calibration targets. While they found the regression of their sample was linear, the use of three calibration targets allowed the assessment of potential non-linearity. This was also the case with the UAS study by Wang and Myint (2015). Using nine calibration targets, Wang and Myint (2015) found an exponential relationship between the sensor-recorded DNs and surface spectroradiometer reflectance. Thus, two calibration targets are essential to an ELC, but three or more calibration targets are necessary to account for any non-linearity that may exist.

For UAS applications, the size of the calibration tarp should not be an issue because of the relatively high resolution and low flight altitude of the systems (Kelcey and Lucieer 2012; Wang and Myint 2015). More important is the cost of the calibration target because a benefit of using UAS in the first place is the relatively low operating costs compared to manned and satellite systems. Kelcey and Lucieer (2012) developed an ELC that used three calibration targets constructed from white Tyvek, grey fabric, and black fabric. In an attempt to develop a more cost-effective calibration target, Wang and Myint (2015) designed and tested 10 different materials for use as calibration targets. They found that Masonite hardboard painted with a nine-level gray gradient worked best. While both studies found promising results, there is a need to build upon their methods in an attempt to develop a standard high performance and cost-effective calibration target that can be used across a variety of applications.
The calibration coefficients developed by Wang and Myint (2015), and those of many other authors, are limited by the idealized conditions of the case study. It is not always practical to conduct field work under such idealized conditions, especially in rapid response situations. The reality of many UAS applications introduces a variety of environmental conditions that change throughout the day and have substantial seasonal variation. These applications, and especially water operations, are also constrained by the amount of gear that can be transported, so the use of large and numerous calibration targets is impractical. Therefore, the objective of the current study is to investigate whether the atmosphere significantly impacts UAS imagery and determine the applicability of current remote sensing image correction techniques in real-world UAS applications.

**North Farm Experiment**

The first question that needs to be answered is whether the atmosphere has a significant impact on low-altitude UAS imagery. If the atmosphere has a significant impact on the imagery, then there will be a need to correct UAS imagery for atmospheric effects prior to the use of the imagery in certain applications.

An experiment was carried out on 22 April 2016 at North Farm on the campus of Mississippi State University. The UAS included an X8 octocopter carrying a three band Canon DSLR modified to collect color-infrared (CIR) imagery and an additional five band MicaSense RedEdge camera (Figure 2.1 & Table 2.1).
Table 2.1 UAS Specifications

<table>
<thead>
<tr>
<th>Sensor</th>
<th>UAVs</th>
<th>Resolution (pixels)</th>
<th>FOV (degrees WxH)</th>
<th>GSD @ 125m (cm)</th>
<th>Bands</th>
<th>Center Wavelengths (nm)</th>
<th>Imaging Frequency (seconds / img)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canon EOS Rebel SL1</td>
<td>Nova Block, 3</td>
<td>5184 x 3456</td>
<td>58.27 x 40.86</td>
<td>2.62</td>
<td>3</td>
<td>Maxmax modified G, R, NIR</td>
<td>&lt; 0.5</td>
</tr>
<tr>
<td>MicaSense RedEdge</td>
<td>X8</td>
<td>1280 x 960</td>
<td>47.2 x 35.4</td>
<td>8</td>
<td>5</td>
<td>475, 560, 668, 840, 717</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 2.1 Caption: Specifications of the components making up the UASs flown

Figure 2.1 Camera Spectral Response

Figure 2.1 Caption: Canon DSLR (a) and MicaSense RedEdge (b) spectral response functions

Imagery of three gray scale (6%, 22%, 44%) spectrally homogenous calibration panels was collected continuously from altitudes of 15 to 800 feet AGL. The imagery was collected over a one hour period from 1500 – 1600 (LDT). A nearby Natural Resources Conservation Service Soil Climate Analysis Network (SCAN) site recorded
relative humidity values ranging from 36% to 40%. Illumination conditions were fairly constant; however, some intermittent clouds cast a shadow and affected some of the imagery. Attempts were made to filter out these affected images from the analysis. Discrepancies between the recorded GPS altitudes and visual assessment of the Canon DSLR imagery required additional filtering of images (Figure 2.2). The final dataset used in the statistical analysis for the Canon DSLR camera became underwhelming as a result of this required filtering. Emphasis is therefore put on the statistical analysis using the MicaSense RedEdge camera because of its advanced spectral properties and GPS response time. A similar filtering process was used to remove MicaSense RedEdge images that were unsatisfactory due to cloud conditions. There were no obvious discrepancies between the GPS recorded altitude and a visual assessment of the MicaSense RedEdge imagery. A bootstrap resampling technique was used to improve the dataset’s ability to describe the greater population that these samples represent and to produce 95% confidence intervals (CIs) of the flight altitude sample means for statistical comparisons. Plots of 95% CIs for each elevation provide a visual tool for interpreting how the panel-recorded radiance is affected by changes in flight altitude (Figure 2.3 & Figure 2.4). 100 feet was considered ground constant for comparison against imagery collected at altitudes aloft.
Figure 2.2  GPS Response

Figure 2.2 Caption: Comparison of MicaSense and Canon GPS response time. Canon GPS has a smoother profile due to a slower response time compared to MicaSense.

Figure 2.3  Canon Plotted Confidence Intervals

Figure 2.3 Caption: Plots of the 95% confidence intervals created from the bootstrapped means for each altitude. The rows are the spectral bands and are displayed from top to bottom as Green, Red, and NIR. The columns are the tarp panels and are displayed from left to right as 6%, 22%, and 44%. 100 feet AGL was considered “ground” constant. Confidence intervals falling outside of 100 foot AGL confidence interval (red dash) are considered statistically different imagery. These comparisons were confirmed by permutation tests.
Figure 2.4  MicaSense Plotted Confidence Intervals

Figure 2.4 Caption: Plots of the 95% confidence intervals created from the bootstrapped means for each altitude. The rows are the spectral bands and are displayed from top to bottom as Blue, Green, Red, Red Edge, and NIR. The columns are the tarp panels and are displayed from left to right as 6%, 22%, and 44%. 100 feet AGL was considered “ground” constant. Confidence intervals falling outside of 100 foot AGL confidence interval (red dash) are considered statistically different imagery.
Plotted CIs from the Canon images indicate significant differences (p < 0.05) between the imagery collected at the ground and imagery collected at an altitude of 200 feet for all three spectral bands; however, challenges faced with the Canon DSLR camera GPS response time led to an underwhelming dataset. In contrast, the MicaSense RedEdge camera GPS response time was adequate to produce a sufficient dataset for a more robust analysis. CIs from the MicaSense images show agreement with the Canon that there are significant differences (p < 0.05) between images collected at the ground and images collected at 200 feet (Figure 2.4). In fact, most spectral bands suggest that there are significant differences between images collected at the ground and images collected at higher flight altitudes. The aerosol concentrations from 200 to 800 feet likely stay relatively consistent which would suggest the slight variations between altitudes are a signal of BRDF effects. Another reoccurring signature is that some CIs appear to have a range of zero. These ranges of zero coincide with oversaturated images and a recording of maximum radiance by the MicaSense RedEdge Camera. This is most apparent for the Blue and Green spectral bands, those wavelengths that are more readily scattered by the atmospheric column. It is possible that these oversaturation signals are additional proxies for atmospheric effects on the imagery. From this experiment, it was apparent that the atmosphere significantly impacts low-altitude UAS imagery and requires the correction of UAS imagery.

**Image Correction**

**Study Area**

The study spanned a 90 km² area over the Lower Pearl River Watershed just southeast of Slidell, Louisiana (Figure 2.5). The region is a complex coastal estuary with
a mix of low marsh grasses, braided streams, dense forests, manmade structures, and prominent aquatic vegetation. The area is subject to relatively high humidity values and sometimes strong sea breeze events, often initiating sporadic precipitation events during the warm season.

![Study Area](image)

Figure 2.5  Study Area

Figure 2.5 Caption: Study area is highlighted by green polygons.

**Materials and Methods**

*Unmanned Aerial System*

The UAS was composed of the Nova Block 3 unmanned aerial vehicle (UAV) flying with a modified Canon DSLR camera (Figure 2.6; Table 2.2).
Table 2.2 UAS Specifications

<table>
<thead>
<tr>
<th>Sensor</th>
<th>UAVs</th>
<th>Resolution (pixels)</th>
<th>FOV (degrees WxH)</th>
<th>GSD @ 125m (cm)</th>
<th>Bands</th>
<th>Center Wavelengths (nm)</th>
<th>Imaging Frequency (seconds / img)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canon EOS Rebel SL1 Nova Block 3</td>
<td>5184 x 3456</td>
<td>58.27 x 40.86</td>
<td>2.62</td>
<td>3</td>
<td>Maxmax modified B G, R, NIR</td>
<td>&lt; 0.5</td>
<td></td>
</tr>
</tbody>
</table>

Table 2.2 Caption: Specifications of the components making up the UAS flown

Figure 2.6 Canon Camera Spectral Response

Figure 2.6 Caption: Spectral response functions of the original Canon DSLR camera without the yellow filter (a) and the Canon DSLR camera with the yellow filter (b)

The Nova Block 3 was capable of flying for 80 minutes on a single battery charge. The Canon DSLR collected data in three spectral wavebands at a 2.62 cm horizontal spatial resolution when flown 125 m above ground level (AGL). The original camera collected data from 300 – 1000nm (Figure 2.6a); however, the sensor was retrofitted with a Wratten #12 yellow filter that blocks wavelengths less than 450 nm (Figure 2.6b). The purpose of this filter was to convert the blue sensor to a surrogate NIR sensor. By limiting the response to blue visible wavelengths, the camera is converted to a Green-Red-NIR color-infrared (CIR) camera; however, the resulting NIR spectral band has a rather noisy and weak signature.
Calibration Targets

Three gray panels were used as calibration targets in the UAS imagery. Three panels were used to balance the size of the equipment and the need to quantify any non-linearity in the UAS-spectroradiometer relationship. The darkest panel was a 6% reflectivity value. The lightest panel was capped at a 44% reflectivity value to try to avoid overexposure of the sensor. A 22% reflectivity value was selected as the moderate panel. Specific requirements for the calibration targets were as follows:

1. The targets should be large enough to fill the IFOV of multiple sensors, yet small enough to be portable and used on a boat.
2. The targets should lay orthogonal to the sensor both on land and on a boat.
3. The targets should have near-Lambertian reflectance properties.
4. The targets should be able to lay on top of vegetation without interference of reflecting surfaces.
5. The targets should be durable and washable.

The final calibration target was composed of three 2’ x 4’ gray strips of coated Type 822 fabric sewn together using a 1” overlapping seam and a 2” folded hem around the perimeter. The Type 822 fabric is a durable, high-strength woven polyester fabric with an Oxford weave that provides near-Lambertian reflectance properties. The coating of the panels was a pigmented acrylic/silicone polymer that was neutral in hue and devoid of spectral content from 420 to 1600 nm, with minor spectral content to 2200 nm. The targets were laboratory-calibrated using a Perkin-Elmer 1050 Spectrophotometer with an integrating sphere so that the band average diffuse hemispherical reflectivity (DHR) of the individual fabric webs were 6%, 22%, and 44% (R +- 0.05R) (Figure 2.7; Table 2.3).
Table 2.3  Laboratory-Measured Band Reflectance

<table>
<thead>
<tr>
<th>Band Average Reflectance (%)</th>
<th>420 - 700 nm</th>
<th>420 - 1050 nm</th>
<th>900 - 1700 nm</th>
</tr>
</thead>
<tbody>
<tr>
<td>6%</td>
<td>7.2</td>
<td>6.7</td>
<td>5.6</td>
</tr>
<tr>
<td>22%</td>
<td>23.2</td>
<td>22.7</td>
<td>20.9</td>
</tr>
<tr>
<td>44%</td>
<td>45.8</td>
<td>45.7</td>
<td>45</td>
</tr>
</tbody>
</table>

Table 2.3 Caption: Laboratory-measured band average reflectance for each calibration panel

Figure 2.7  Laboratory-Measured DHR

Figure 2.7 Caption: Laboratory-measured DHR from 300-2500 nm for the 6%, 22%, and the 44% calibration panels

The panels were then sewn onto a larger 4’ x 6’ piece of uncoated Type 822 fabric. The larger 4’ x 6’ piece of uncoated Type 822 included a pouch with a Velcro opening and six size 4 reinforced clawgrip grommets. The reinforced clawgrip grommets allowed for the tarp to be securely affixed to a boat. The pouch allowed for an insert (e.g. plywood or other rigid material) so that the calibration target would lie flat on the ground (Figure 2.8a) or on the top of a boat canopy (Figure 2.8b). The greatest challenge with the design was preventing creasing when deploying the calibration target.
Figure 2.8  Calibration Tarp Deployment

Figure 2.8 Caption: Constructed calibration targets deployed on land (a) and on boat canopy (b). Tie downs are loosened during flights to allow the board to lie flat on top of the boat canopy. The darkest panel is 6\% reflectance, the medium panel is 22\% reflectance, and the lightest panel is 44\% reflectance.

Field Operations

The relatively high humidity values and morning temperature inversions posed challenges to prevent fogging of the camera lens. Flight range was limited by the Federal Aviation Administration (FAA) line of sight requirements in place at the time of the study. Vegetation across the study region changes from marsh grass in the southern portion to dense forest in the northern portion; therefore, this limited the northern extent of flight coverage because line of sight with the UAS became a challenge as vegetation
converted to a forested landscape. Also, the north-south orientation of the river network limited the east-west extent of flight coverage.

Regular operations involved two boats, one dedicated to flight operations and the other dedicated to water operations. The flight operations boat was in charge of the flight of the UAS and the download and storage of flight data. The water operations boat worked with the flight operations boat so that water quality information was collected every time the UAS flew directly overhead the water operations boat. The water operations boat also carried the calibration target. Ground reference data of the calibration target panels was taken in tandem with the UAS flyover using a GER 1500 spectroradiometer. The workflow from field operations to the final image correction and calibration from these missions is summarized in Figure 2.9.

Figure 2.9  Data Processing Workflow

Figure 2.9 Caption: Workflow of data collection and calibration of UAS imagery
Table 2.4 UAS Flight Missions

<table>
<thead>
<tr>
<th>Date</th>
<th>Sensors Flown</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dec 15-18, 2014</td>
<td>5 cm Canon CIR</td>
</tr>
<tr>
<td>Mar 02-06, 2015</td>
<td>5 cm Canon CIR</td>
</tr>
<tr>
<td>May 17-22, 2015</td>
<td>5 cm Canon CIR</td>
</tr>
<tr>
<td>Aug 09-14, 2015</td>
<td>5 cm Canon CIR &amp; Test 16 cm MicaSense</td>
</tr>
<tr>
<td>Dec 14-18, 2015</td>
<td>5 cm Canon CIR &amp; 16 cm MicaSense</td>
</tr>
</tbody>
</table>

Table 2.4 Caption: Date of missions, sensors flown, and image resolution based on flight altitude

**Flight Data**

Flight data used in this study began in December 2014 with follow up missions at about a two month interval (Table 2.4). While regular operations began in December 2014, the use of the previously described calibration panels did not begin until the August 2015 mission. Flight mosaics were generated for each mission using a camera alignment technique in Agisoft Photoscan Pro software. This technique used the orientation of the camera at the time of the image registration and common points in overlapping images to generate tie points to triangulate the cameras position. A point cloud was generated from this alignment procedure and was examined for any anomalies. The final point cloud was used to generate a mesh on which a GeoTIFF orthomosaic was exported. All post-processed imagery can be retrieved online (http://www.gri.msstate.edu/geoportal/).

**Empirical Line Calibration**

The wavelengths of the spectroradiometer data collected for each calibration panel (6%, 22%, and 44%) during the August 2015 and December 2015 missions were converted to whole numbers, interpolated, and weighted to match the Canon DSLR sensor spectral response function (SRF). This weighting process adjusted the
spectroradiometer data to mimic those values the UAS would have recorded in the absence of the atmosphere. With the spectroradiometer data adjusted to match that of the UAS, the UAS imagery was analyzed to find locations where the water operations boat could be seen in the imagery. If the georeferenced information of the boat in the imagery matched the location information recorded at the time of the ground reference data, then the UAS DN pixel values were extracted and averaged for each of the calibration panels. Hence, the final dataset used for the regression was a series of UAS-weighted ground spectroradiometer radiance and extracted UAS DN data pairs (Figure 2.10).

Figure 2.10  Spectroradiometer-UAS Scatterplot

Figure 2.10 Caption: UAS-weighted ground spectroradiometer-UAS image pairs for Canon DSLR green band (a), red band (b), and NIR band (c). Labels at each point indicate the sample site where the data was obtained.
Using these ground reference-UAS image pairs, regressions between the adjusted spectroradiometer radiance values and the UAS DN values of the calibration panels were developed for each spectral band. Ideally, new calibration equations would be created for each of the missions flown to account for diurnal and seasonal environmental differences; however, this type of analysis is not feasible in most UAS applications due to the numerous flights required to image the full study area. Therefore, ground reference-UAS image pairs were combined from numerous missions and flight conditions to develop robust calibration equations that can minimize errors across a variety of UAS applications.

Once the imagery was calibrated and the UAS brightness values were converted to radiance, the imagery was normalized for illumination by converting from radiance to reflectance using the following equation:

\[
\text{Reflectance} = \frac{\Pi L_\lambda d^2}{E_{sun\lambda} \cdot \sin\theta_e}
\]  

(3.1)

where \(L_\lambda\) is the spectral radiance, \(d\) is the Earth-Sun distance in astronomical units, \(E_{sun\lambda}\) is the exoatmospheric solar spectral irradiance for each sensor spectral band, and \(\sin\theta_e\) is the solar elevation angle.

This method of converting to radiance first, then later normalizing for illumination was selected because the solar illuminations conditions in future UAS applications will be different than the solar illuminations conditions experienced during the development of the calibration equations. Rather than develop new calibration equations for each time of the year, this allows for the application of the calibration equations to the same imaging sensor throughout the year.
Results

Image Correction

All Canon DSLR spectral bands exhibited an exponential relationship between the raw UAS DNs and the ground spectroradiometer-measured values of the three calibrations panels (Figure 2.10). These exponential relationships are in agreement with past research (Smith and Milton 1999; Wang and Myint 2015; Kelcey and Lucieer 2012) and are an indication that the data follows initial expectations. Two-thirds of the available site data from August 2015 and December 2015 were randomly selected and used to develop the calibration equations. The remaining one-third of data were used to verify the performance of the calibration equations.

Table 2.5 provides the three resulting calibration equations, one for each spectral band. Performance of the calibration equations was assessed using the root mean square error (RMSE), the mean absolute error (MAE), goodness of fit ($R^2$), and a Mann-Whitney U test. While RMSE and MAE would be applicable to the current dataset, any future applications of this methodology may see very different RMSE and MAE due to the nature of the data. To account for this, normalized RMSE (NRMSE) and normalized MAE (NMAE) were calculated and presented. This normalization method simply divides the RMSE and the MAE by the range of the observed data. The Mann-Whitney U test is a nonparametric test that assesses whether the distributions of two datasets are equal. Therefore, the Mann-Whitney U test was used to test the null hypothesis ($H_0$) that the distribution of the predicted and actual datasets are equal. Table 2.5 provides a summary of the verification results for each calibration equation.
Table 2.5: Regression Results

<table>
<thead>
<tr>
<th>Band</th>
<th>Calibration Equation</th>
<th>R-squared</th>
<th>NRMSE (%)</th>
<th>NMAE (%)</th>
<th>Mann-Whitney</th>
</tr>
</thead>
<tbody>
<tr>
<td>Green</td>
<td>$y = 2773.7\exp(0.0168x)$</td>
<td>0.77</td>
<td>20.2</td>
<td>3.05</td>
<td>U = 317, p = .405</td>
</tr>
<tr>
<td>Red</td>
<td>$y = 2247.1\exp(0.0171x)$</td>
<td>0.79</td>
<td>18.5</td>
<td>1.61</td>
<td>U = 317, p = .361</td>
</tr>
<tr>
<td>NIR</td>
<td>$y = 2501.7\exp(0.0182x)$</td>
<td>0.77</td>
<td>19.3</td>
<td>1.92</td>
<td>U = 317, p = .412</td>
</tr>
</tbody>
</table>

Table 2.5 Caption: Resulting calibration equations and verification metrics

The goodness of fit for each band ($R^2_{\text{green}} = 0.77$, $R^2_{\text{red}} = 0.79$, $R^2_{\text{nir}} = 0.77$) is relatively high considering the field nature of this work and the relatively large sample size compared to the idealized conditions and small sample sizes common in previous studies. Normalized RMSE ($\text{NRMSE}_{\text{green}} = 20.2\%$, $\text{NRMSE}_{\text{red}} = 18.5\%$, $\text{NRMSE}_{\text{nir}} = 19.3\%$) and normalized MAE ($\text{NMAE}_{\text{green}} = 3.05\%$, $\text{NMAE}_{\text{red}} = 1.61\%$, $\text{NMAE}_{\text{nir}} = 1.92\%$) are also within acceptable ranges for each spectral band. More important was that the Mann-Whitney U test results were not significant ($U_{\text{green}} = 317$, $p = .405$; $U_{\text{red}} = 317$, $p = .361$; $U_{\text{nir}} = 317$, $p = .412$) meaning that we cannot reject the $H_0$, so we can conclude that the distribution of the predicted and measured dataset are equal. This confirms that the methods developed produce an acceptable camera calibration using data collected during a complex field experiment. Therefore, the calibration equations will adequately minimize the atmospheric effects on the data when applied to the full mission imagery.

The current work separated the calibration equations from the solar illumination normalization so that the calibration equations could be applied to any imagery collected with the same Canon DSLR camera. Once the calibration equations were applied to the mission imagery, the time of each flight was extracted and used to normalize the atmosphere-corrected imagery for solar illumination using equation 2.1. Vegetation should return a strong NIR signal; therefore, 2000 vegetation pixels were selected and
their median spectral response plotted for the August 2015 mission imagery to illustrate the results of the ELC (Figure 2.11). The uncorrected NIR signal from vegetation appears much lower than expected although it may not be that the NIR signal is low; it is possible that atmospheric scattering has added erroneous Green and Red signal to the image. After applying the calibration equations, the Green and Red signals are decreased and a more typical vegetation spectral response results. The last step normalizing for solar illumination helps to minimize additional erroneous signals due to changing solar conditions throughout the image collection period.

Figure 2.11  August Vegetation Spectral Response

Figure 2.11 Caption: Spectral response of water samples from 2000 vegetation pixels in the August 2015 imagery for the (a) uncorrected image, (b) atmospherically corrected image, and (c) atmospherically corrected and solar illumination normalized image
**Discussion**

The discovery that environmental factors, such as atmospheric scattering, have significant impacts on relatively low altitude UAS imagery has important implications for UAS applications that require spectrally accurate information, like harmful algae bloom detection and multitemporal analyses. In the North Farm experiment, it was discovered that the median signal recorded between 200 feet and 800 feet was sometimes over 50% greater than the median signal recorded at 100 feet AGL. While a nearby SCAN site indicated ground relative humidity values around 40%, it is also important to record water vapor concentrations throughout the boundary to better quantify the atmospheric effects on UAS imagery. Follow-up experiments will add on-board meteorological sensors and will take place in a wide variety of conditions. These follow-up experiments will provide a robust assessment of the atmospheric effects on UAS imagery in various conditions.

A unique aspect of the current work is that the image correction procedure split the correction for atmospheric effects and the solar illumination normalization. Separating these steps allows for the correction of any imagery collected with the same Canon DSLR camera no matter the time of collection. In addition, the development of these calibration equations using numerous sampling sizes across multiple missions created more robust and operational calibration equations that can be applied in various UAS applications to minimize atmospheric and solar illumination effects on imagery. While this method also produced a lower goodness-of-fit value compared to idealized studies, the calibration equations better represent the varied conditions typical in UAS applications. 

*Error! Reference source not found.*
As can be seen in **Error! Reference source not found.**, a single site could have been used to develop the calibration equations to obtain the high goodness-of-fit results found in idealized scenarios; however, to employ these site-specific equations on new data would lead to substantial errors because of the varied environmental conditions and field challenges experienced in real-world UAS applications. Thus, developing calibration equations from numerous sites across multiple missions generates more representative calibration equations and leads to less overall error in the calibration of new imagery.

Consistent illumination conditions across the flight mosaics were one of the greatest challenges because it was rare that the sampling coincided with an entirely clear-sky day. Cloud shadows were common in the mission mosaics which affected local illumination conditions. This will be a common challenge in UAS applications because one of the greatest advantages of these systems is their ability to fly in operational
settings when cloud conditions prohibit other remote sensing systems. These local illumination variations led to some situations where the UAS flyover of the boat and the time of the ground reference spectroradiometer sampling occurred under a cloud shadow, so those local illumination effects were quantified to some degree by the calibration equation. However, it was more often the case that the UAS flyover of the boat and the time of the ground reference spectroradiometer sampling occurred under sunny conditions while other portions of the image were influenced by cloud shadows. In addition to increasing the sample data included in the calibration equation development, the development of automated techniques to normalize for cloud shadow contamination would help improve the performance of the calibration equations.

A great advantage of an ELC is that it can help reduce external sources of error beyond atmospheric effects. The nature of field work, especially on the water, introduces challenges to the data collection process. For example, the calibration panels were made to have near-Lambertian reflectance properties, so they should have shown very little to no bidirectional reflectance when they were lying orthogonal to the sensor. The reality of field work on the water is that this is rarely the case. Wave action often caused the boat to rock which would complicate the assumption that the calibration targets remained orthogonal to the UAS. It is possible that the wave action also induced bidirectional reflectance effects of the calibration target. These factors appeared to have caused the pixel values of the 44% panel to ‘bleed’ into the 22% panel (Figure 2.13). By capturing these variations in field conditions within the calibration equations, the errors contributed by these external factors were also reduced.
Conclusions

The objective of the current study was to investigate whether the atmosphere significantly impacts UAS imagery and determine the applicability of current remote sensing image correction techniques in real-world UAS applications. Experiments indicated that the atmosphere has a significant image on relatively low altitude UAS imagery. An ELC framework was developed to correct UAS imagery in a real-world UAS field application. The water-based nature of this study required the modification of current ELC methods and led to the development of a cost-effective and feasible ELC framework for correcting UAS imagery in various applications. Results agreed with previous studies that there is an exponential relationship between the UAS recorded DNs and spectroradiometer radiance values. Calibration equations developed from these exponential relationships performed adequately, indicating that ELC remote sensing image correction techniques are effective in real-world UAS applications. A great benefit of the ELC framework presented in the current study is that the calibration equations
incorporate a variety of conditions from multiple missions. In addition to correcting atmospheric effects on the imagery, this helped minimize other field sampling errors and extended the applicability of the calibration equations to various UAS applications. Furthermore, the calibration equations were developed independent from the solar illumination normalization so that the calibration equations can be applied to any future imagery collected with the same imaging sensor.
CHAPTER III
LAND SURFACE CLASSIFICATION OF UNMANNED
AERIAL SYSTEM IMAGERY

Literature Review

Unmanned Aerial Systems (UASs) offer the opportunity to characterize land
surface features in unprecedented detail. The high spatial resolution, low costs, and rapid
deployment capabilities of UAS have gained substantial attention in land use land cover
(LULC) classification applications. Recent findings that UAS imagery can produce
classification accuracies comparable to manned aerial systems (MAS) and satellites
(Casado et al. 2016; Casado et al. 2015; Laliberte et al. 2011; Laliberte and Rango 2009;
Husson et al. 2016) have given researchers all the more reason to pursue the employment
of UASs in their own research. However, with the benefits of UASs come new challenges
and questions of whether current remote sensing image classification configurations can
be applied to UAS applications without modification.

There are many factors that play into the precision and accuracy of LULC
classifications including: remotely sensed data selection; selection of training samples;
image preprocessing (e.g. geometric rectification, radiometric calibration, atmospheric
correction); feature representation; classification technique; and post-classification
processing (Lu and Weng 2007). While regional LULC classifications may only require
moderate resolution space-borne data, local scale LULC classifications require much
higher resolution imagery collected by airborne aircrafts or low orbit satellites. This is especially true for wetlands where high amounts of spectral mixing occurs (Ozesmi and Bauer 2002). There are advantages and disadvantages to using high resolution data. The ability to identify small-scale features that would otherwise be spectrally mixed in a moderate resolution data pixel (Cingolani et al. 2004) is a great advantage of high resolution data. Husson et al. (2016) successfully used high spatial resolution UAS imagery to identify aquatic vegetation growth-form and dominant-taxon in a complex coastal estuary. Casado et al. (2016) found that high resolution (2.5 cm) UAS imagery produced higher accuracy classifications compared to 5 cm and 10 cm resolution imagery. However, Casado et al. (2016) also found that even in their small study area (1.4 km river reach), processing time was substantially increased with increased image resolutions. This, along with the added challenges of shadowing, spectrally similar features, and data volume (Myint et al. 2011), suggests the possible need for UAS specific classification configurations.

Some degree of image preprocessing is always required prior to an image classification. This can include mosaicking images to create a seamless version of the area of interest. This mosaicking requires that overlapping cells in the images are combined by either taking the value of a single cell or blending the two cell values together. Another common preprocessing step can include radiometric calibrations where the image digital number (DN) values are either converted to match the expected surface radiance, are converted to radiance without regard to measured surface radiance, or are normalized for illumination where the radiance values are converted to reflectance to negate the influence of solar illumination (Song et al. 2001; Hu 2002). However, these
image corrections are not always required prior to classification (Song et al. 2001). In fact, Song et al. (2001) determined that many image classifications using MAS and satellite data do not require image correction; image correction was only required in cases where multitemporal data were being used with a single classification algorithm, or certain indices (e.g. normalized difference vegetation index). The extent to which this guidance can be applied to UAS imagery is yet to be determined. This inquiry is important given that UAS imagery is often a mosaic of multiple flights rather than the current single snapshot in time.

Overall accuracy of any classification algorithm relies heavily on the proper selection of training pixels (Hubert-Moy et al. 2001; Chen and Stow 2002; Lu and Weng 2007; Pal and Mather 2004, 2003). This is especially important when using high resolution data because two physically different features may have very similar spectral responses (Zhang and Wang 2003; Myint et al. 2011). However, logistics, time, and costs make it challenging to acquire large field-based sample datasets (Buchheim & Lillesand, 1989; Jackson & Landgrebe, 2001). The limitations to creating large field-based sample datasets can be overcome with the help of image analyst experts using high resolution imagery to build sample datasets (Lu and Weng 2007). As with any sample dataset, it is possible that the selection of too many training samples could reduce the accuracy of the classification (Hughes 1968; Price et al. 2003; Pal and Mather 2003). Therefore, some rules of thumb for MAS and satellite imagery recommend a range from 51-100 training samples per class (Fitzpatrick-Lins 1981; Congalton 1991).

The way in which features are characterized when ingested by a classification technique will affect the overall performance and appearance of a LULC classification.
The most common method for characterizing features has been on a per-pixel basis, though recent work has demonstrated the benefits of first segmenting (i.e. pixel grouping) the image to characterize features on an object basis (Myint et al. 2011; Robertson and King 2011; Im et al. 2008; Park 2015). Myint et al. (2011) compared a per-pixel based classifier versus an object-based classifier, finding that the object-based classification resulted in a 23% improvement in classification accuracy compared to a per-pixel.

Object-based classifications have the distinct advantage in high spatial and spectral resolution imagery where two different features may have very similar spectral color, but such classifications come with the downside of increasing image preprocessing time. Another drawback to an image segmentation is the opportunity to introduce error to the classification by inappropriately clustering features.

While past classification techniques such as the maximum likelihood classifier (MLC) and the logistic regression (LogR) classifier have proven to be valuable classification techniques (Michelson et al. 2000; Robertson and King 2011; Pal and Mather 2003), increased computing power and access to high resolution remotely sensed data has led to the use of more sophisticated classification techniques including artificial neural networks (ANN), random forest (RF) and support vector machines (SVM). Many studies have compared the applications and performance of these more advanced classification techniques with mixed results determining the most effective technique (Lu and Weng 2007; Shao and Lunetta 2012; Srivastava et al. 2012). Srivastava et al. (2012) found that ANN produced superior results compared to SVM, whereas Shao and Lunetta (2012) found that SVM produced superior results compared to ANN. It has been well documented that RF algorithms can be a viable alternative to the more commonly used
SVM and ANN classifier algorithms (DeFries et al. 1998; Friedl et al. 1999; Pal and Mather 2003; Lawrence et al. 2004; Im et al. 2008; Friedl and Brodley 1997). A RF classifier is different than most classifiers because it is a non-parametric classifier that works just as the name suggests, by recursively testing a set of training samples to build a classification tree. A great advantage of RF classifiers is that they do not make statistical assumptions about the data distribution nor do they require the magnitude of computing power involved with an ANN classification (Friedl and Brodley 1997; Im et al. 2008; Pal and Mather 2003). As such, RF classifiers have been the primary classification technique employed in UAS image classifications (Husson et al. 2016; Feng et al. 2015). Feng et al. (2015) demonstrated the effectiveness of using an RF classifier versus a MLC to distinguish features in an urban area using UAS true color imagery. Similarly, Husson et al. (2016) identified aquatic vegetation growth-form and dominant-taxon from UAS imagery with greater accuracy using an RF classifier compared to a threshold classifier. It appears that the RF classifier can outperform standard threshold and MLC classification techniques in UAS applications; however, these comparisons have been limited to a single classification instance. There is a need for a robust assessment of how MLC and LogR classifiers compare to the RF classifier in UAS applications.

Recent UAS research continues to employ current remote sensing image classification methods with little awareness as to how variations to image correction techniques, training sample configurations and classification techniques affect UAS image classification performance. There is a need to investigate whether current remote sensing image classification configurations can continue to be applied to UAS applications without modification. Therefore, the current study provides an assessment of
how current remote sensing image preprocessing techniques, training sample configurations, and classification techniques perform in UAS LULC image classifications.

**Methods**

**UAS Imagery**

The data used in this study was collected over a 90 km² portion of the Lower Pearl River Watershed just southeast of Slidell, Louisiana (Figure 3.1).

![Study Area](image)

**Figure 3.1** Study Area

Figure 3.1 Caption: Study area is outlined by green polygons
This region was selected in cooperation with the Lower Mississippi River Forecast Center with the goal of generating high resolution maps of the complex stream network. The region is a complex coastal estuary where the southern portion of the study area is dominated by a mix of low marsh grasses, braided streams, and prominent aquatic vegetation. In the northern portion, low marsh grasses are replaced with dense forests where there is less salt water intrusion.

The data was collected by a UAS flying with a modified Canon DSLR camera (Table 3.2). A detailed account of the data collection field campaign is provided in Chapter 2. The UAS was composed of a Nova Block 3 unmanned aerial vehicle (UAV) and a Canon DSLR modified to collect in three wavebands (Green, Red, and NIR) at a 2.62 cm horizontal spatial resolution when flown 125 m above ground level (AGL). Data was collected at approximately two-month intervals over a one year period resulting in five field campaigns (Table 3.3). Each mission imagery was collected throughout the day over a one week period. This means that each mosaic is essentially multitemporal imagery because the time of day and the day on which data was collected varies.

<table>
<thead>
<tr>
<th>Sensor</th>
<th>UAVs</th>
<th>Resolution (pixels)</th>
<th>FOV (degrees WxH)</th>
<th>GSD @ 125m (cm)</th>
<th>Bands</th>
<th>Center Wavelengths (nm)</th>
<th>Imaging Frequency (seconds / img)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canon EOS Rebel SL1</td>
<td>Nova Block 3</td>
<td>5184 x 3456</td>
<td>58.27 x 40.86</td>
<td>2.62</td>
<td>3</td>
<td>Maxmax modified G, R, NIR</td>
<td>&lt; 0.5</td>
</tr>
</tbody>
</table>

Table 3.2 Caption: Specifications of the components making up the UASs flown
Table 3.3 UAS Flight Missions

<table>
<thead>
<tr>
<th>Date</th>
<th>Sensors Flown</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dec 15-18, 2014</td>
<td>5 cm Canon CIR</td>
</tr>
<tr>
<td>Mar 02-06, 2015</td>
<td>5 cm Canon CIR</td>
</tr>
<tr>
<td>May 17-22, 2015</td>
<td>5 cm Canon CIR</td>
</tr>
<tr>
<td>Aug 09-14, 2015</td>
<td>5 cm Canon CIR &amp; Test 8 cm MicaSense</td>
</tr>
<tr>
<td>Dec 14-18, 2015</td>
<td>5 cm Canon CIR &amp; 8 cm MicaSense</td>
</tr>
</tbody>
</table>

Table 3.3 Caption: Date of missions, sensors flown, and image resolution based on flight altitude

**Image Preprocessing**

Three versions of the UAS imagery are used in the classifications to assess how image corrections affect UAS image classification performance. The first is the raw DN value imagery, the second corrects the UAS imagery for noise contributed by the atmosphere using an empirical line calibration (ELC) technique (Chapter 2), and the third takes the corrected UAS imagery and normalizes for changes in illumination caused by changing solar illumination conditions (Chapter 2).

After the images were corrected, the multiple flight mosaics that make up each mission were mosaiced to create a large single GeoTIFF for each mission. Due to the substantial data volume and processing time that would be required for the classification of each mosaic at the native 5 cm resolution, the mission mosaics were resampled to 0.25 meter pixels. This resolution was selected because it was half the resolution required to identify the smallest features of interest (0.5 m) in the imagery.

**Image Classification**

August 2015 mission imagery was used in the development of the classification framework because it provided the most complete coverage of the study area. The
remaining missions were used to test the practical implementation of the top performing
classification framework. Three multispectral wavebands — Green, Red, and NIR—
were used as input to three commonly used classification techniques, LogR, MLC, and
RF, to classify open water (OW), terrestrial vegetation (TV), and aquatic vegetation
(AV).

Samples were composed of 6000 total pixels, 2000 OW, 2000 TV, and 2000 AV
pixels. It was important to separate aquatic vegetation from terrestrial vegetation because
aquatic vegetation typically grows and floats on top of water. The aquatic vegetation
appears similar to terrestrial vegetation in remotely sensed imagery but acts to
camouflage the underlying water column. Therefore, the aquatic vegetation needs to be
characterized separately from terrestrial vegetation to allow for the accurate mapping of
water bodies. It was impractical to collect a sufficient sample dataset from fieldwork;
thus, the samples were selected by an expert image analyst using the high resolution
August UAS imagery.

Twenty different training/testing ratios were analyzed to investigate the effect
training/testing ratio had on classifications (Table 3.4). In theory, a typical 70:30
training/testing sample classification ratio should be a sufficient ratio in an image
classification.
<table>
<thead>
<tr>
<th>Split Fraction</th>
<th># Training Samples</th>
<th># Verification Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>300</td>
<td>5700</td>
</tr>
<tr>
<td>0.1</td>
<td>600</td>
<td>5600</td>
</tr>
<tr>
<td>0.15</td>
<td>900</td>
<td>5100</td>
</tr>
<tr>
<td>0.2</td>
<td>1200</td>
<td>4800</td>
</tr>
<tr>
<td>0.25</td>
<td>1500</td>
<td>4500</td>
</tr>
<tr>
<td>0.3</td>
<td>1800</td>
<td>4200</td>
</tr>
<tr>
<td>0.35</td>
<td>2100</td>
<td>3900</td>
</tr>
<tr>
<td>0.4</td>
<td>2400</td>
<td>3600</td>
</tr>
<tr>
<td>0.45</td>
<td>2700</td>
<td>3300</td>
</tr>
<tr>
<td>0.5</td>
<td>3000</td>
<td>3000</td>
</tr>
<tr>
<td>0.55</td>
<td>3300</td>
<td>2400</td>
</tr>
<tr>
<td>0.6</td>
<td>3600</td>
<td>2400</td>
</tr>
<tr>
<td>0.65</td>
<td>3900</td>
<td>2100</td>
</tr>
<tr>
<td>0.7</td>
<td>4200</td>
<td>1800</td>
</tr>
<tr>
<td>0.75</td>
<td>4500</td>
<td>1500</td>
</tr>
<tr>
<td>0.8</td>
<td>4800</td>
<td>1200</td>
</tr>
<tr>
<td>0.85</td>
<td>5100</td>
<td>900</td>
</tr>
<tr>
<td>0.9</td>
<td>5600</td>
<td>600</td>
</tr>
<tr>
<td>0.95</td>
<td>5700</td>
<td>300</td>
</tr>
<tr>
<td>1</td>
<td>6000</td>
<td>6000</td>
</tr>
</tbody>
</table>

Table 3.4 Caption: A breakdown of the number of training samples used and number of verification samples used based on each split factor

Next, the effect that training sample size per class had on the classification was investigated by increasing the number of randomly selected training samples per class from 50-500 at 30 sample intervals. So for a training/testing ratio of 70:30 and a per class sample size of 200, 420 total samples would be randomly selected for classification and the remaining 180 samples would be used for verification. These two analyses were iterated 1000 times using a resampling with replacement technique to create a distribution of the classification overall accuracy (OA) values. 95% Confidence intervals (CIs) were calculated and plotted for each classification OA distribution to compare the impacts of image corrections, training/testing sample ratios, training sample size, and classification
techniques on the final classification performance. In total 540,000 classifications were produced: 20,000 classifications to determine the effect of training/testing ratio for each of the three classification techniques and three image versions, and 16,000 to determine the effect of training sample size for each of the classification techniques and three image versions.

The top performing training/testing sample ratio and sample size configuration was then applied to the three classification techniques and iterated 1000 times to calculate median Kappa, OA, User’s accuracies (UA) and Producer’s accuracies (PA). This allowed for a more robust assessment of classification performance for each of the three image versions. After determining the desired training/testing sample ratio, training sample size, image correction method, and classification technique, this classification configuration was applied to the remaining mission imagery (i.e. December 2014, March 2015, May 2015, and December 2015) to test this framework in a practical application.

**Classification Tuning**

Tuning the RF classifier was required to ensure that it is optimized for the data in the current study. Two of the most influential parameters are the number of variables randomly sampled at each node (mtry) and the number of total trees to grow (ntree). A sensitivity analysis of these parameters concluded that the optimal configuration was to allow one variable to be randomly sampled at each node and grow a total of 500 trees (Figure 3.2).
Figure 3.2 Random Forest Classifier Tuning

Figure 3.2 Figure 3.3 Caption: Tuning of the random forest classifier. A series of combinations adjusting the number of variables randomly sampled as candidates at each split (mtry) and the number of trees to grow (ntree) were used to identify the best tuning

Results

Classification performance is illustrated by CI plots of OA values (Figure 3.3 & Figure 3.4). The training/testing sample ratio progresses from fewer training samples to more training samples used in the development of the classification formulas. The plotted blue line provides the 85% OA target recommended by Foody (2002). Most apparent is a drop in the MLC performance when image corrections were applied. The increasing trend in CI spread is due to decreasing the number of verification points. This demonstrates why it is important to have a sufficient number of verification points. Figure 3.3 indicates that training/testing sample ratios between 55:45 and 70:30 are all reasonable options. Accuracy tends to decrease for ratios 50:50 and below, while precision decreases rapidly from 75:25 and above. The LogR and MLC were less sensitive to training/testing sample ratio than the RF classifier.
A 70:30 training/testing sample ratio was selected for further inquiry into the impacts of sample size on UAS image classifications. Figure 3.4 illustrates the evolution of OA CIs relative to the per class sample size from 50-500 samples per class. Additional OA comparison lines from a survey of numerous satellite (82%: Stow et al. 2007; Otukei and Blassche 2010; Li et al. 2014; Myint et al. 2011; Topaloğlu et al. 2016; Myint and Stow 2011; Varga et al. 2015; Pal and Mather 2003; Truax and Cartwright 2012;
Srivastava et al. 2012; Robertson and King 2011), airborne (81%: Platt and Goetz 2004; Lawrence et al. 2006; Camps-Valls and Bruzzone 2005; Pal and Foody 2010; Bazi and Melgani 2006; Lucas et al. 2008; Bagan et al. 2008; Baker et al. 2013; Samat et al. 2016), and UAS (76%: Laliberte et al. 2010; Casado et al. 2015, 2016; Husson et al. 2016; Feng et al. 2015) classification studies were added to the plots in addition to the 85% target OA recommended by Foody (2002).

Figure 3.4 Classification Confidence Interval Plots

Figure 3.4 Caption: Plots classification overall accuracy confidence intervals with changes to the training data sample size. Training data sample sizes increase from n=50 to n=500. Solid line is the 85% target accuracy threshold proposed by Foody 2002. Dashed line is 82% overall average accuracy calculated by averaging past multi-class classification research using satellite data. Dotted line is 81% overall average accuracy calculated by averaging past multi-class classification research using aircraft data. Dash-dotted line is 76% overall average accuracy calculated by averaging past multi-class classification research using UAV data. (Logistic Regression Classifier (LogR); Maximum Likelihood Classifier (MLC); Random Forest classifier (RF))
All classification techniques observed an increasing trend in precision with increasing sample size. In all cases, the OAs between the uncorrected imagery and the two corrected images were not significantly different (p<0.05). While accuracies between classification techniques for any one sample size category were not significantly different (p<0.05), a visual assessment of the plots indicates that median OAs for the LogR and MLC are slightly greater than the median OAs for the RF classifier. OAs begin to plateau around a per-class sample size of 200, and classification precision becomes relatively consistent at a per-class sample size of 260 for all classifiers. This is larger than sample size recommendations offered by Fitzpatrick-Lins (1981), Congalton (1991), and Congalton and Green (2008), suggesting that the high spatial resolution but low spectral precision of the UAS requires more samples per class to achieve the desired degree of confidence in classification accuracy. Having determined that the 70:30 training/testing ratio is valid and that 260 training samples per class is sufficient, an assessment of median user accuracy (UA), producer accuracy (PA), OA, and Kappa was used to further compare each classification technique (Table 3.5 & Figure 3.5).
Table 3.5 August Classification Median Accuracies

<table>
<thead>
<tr>
<th>Brightness Values</th>
<th>Radiance</th>
<th>Reflectance</th>
<th>Runtime Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>LogR</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UA (%)</td>
<td>93</td>
<td>78</td>
<td>74</td>
</tr>
<tr>
<td>PA (%)</td>
<td>94</td>
<td>81</td>
<td>70</td>
</tr>
<tr>
<td>OA (%)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Kappa</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

| MLC               |          |             |               |
| UA (%)            | 90       | 76          | 76            | -             | 85 | 80 | 66 | - |
| PA (%)            | 95       | 83          | 66            | -             | 95 | 66 | 70 | - |
| OA (%)            | -        | -           | -             | 81            | - | - | 77 | - |
| Kappa             | -        | -           | -             | 0.72          | - | - | 0.65 | - |

| RF                |          |             |               |
| UA (%)            | 90       | 75          | 69            | -             | 90 | 75 | 69 | - |
| PA (%)            | 92       | 76          | 66            | -             | 92 | 76 | 66 | - |
| OA (%)            | -        | -           | -             | 78            | - | - | 78 | - |
| Kappa             | -        | -           | -             | 0.67          | - | - | 0.67 | - |

Table 3.5 Caption: Comparison of classification accuracies between classification technique and image correction method. Median accuracies calculated from 1000 iterations of each classification technique. (Logistic Regression Classifier (LogR); Maximum Likelihood Classifier (MLC); Random Forest classifier (RF); User Accuracy (UA); Producer Accuracy (PA); Overall Accuracy (OA); Open Water (OW); Terrestrial Vegetation (TV); Aquatic Vegetation (AV))
Figure 3.5 August Classification Accuracy Metrics

Figure 3.5 Table 3.5 Caption: Confidence intervals of classification accuracy metrics from 1000 iterations of each classifier for the August imagery configured with 260 samples per class and a 70:30 training/testing sample ratio. (Logistic Regression Classifier (LogR); Maximum Likelihood Classifier (MLC); Random Forest classifier (RF); User Accuracy (UA); Producer Accuracy (PA); Overall Accuracy (OA); Open Water (OW); Terrestrial Vegetation (TV); Aquatic Vegetation (AV))

OW was consistently classified with high accuracy (> 90%). AV was the most challenging class to distinguish with OAs often near or less than 70%. Classification accuracies decreased for the LogR and MLC classifiers when image correction procedures were applied. While not significant, this could be because the image correction used a regression technique to reduce atmospheric errors, but this also brings
the pixels closer to a common solution which decreases variance between classes. In contrast, only one variable was randomly sampled at each node of the RF classifier. This acted to protect it from the decreased variance between classes in the Green and Red bands because the NIR band maintained sufficient variance between classes for all image types (Figure 3.6).

Figure 3.6  August Class Spectral Response

Figure 3.6 Caption: Spectral response distributions for the 2000 samples in each class based on image type. (Overall Accuracy (OA); Open Water (OW); Terrestrial Vegetation (TV); Aquatic Vegetation (AV))
The end goal for separating AV and TV was to appropriately quantify water bodies being camouflaged by aquatic vegetation. Therefore, extra emphasis was put on the classifier’s UA and PA to determine which classifier was best identifying AV and OW pixels. The LogR classifier consistently produced the highest UAs and PAs for AV and OW. LogR performed especially well for the uncorrected image. Applying the LogR classifier to the entire August 2015 uncorrected image mosaic shows how well the classification configuration does identifying water bodies (Figure 3.7). It also does well to concentrate aquatic vegetation near or on top of water bodies. This implies areas of water being camouflaged by aquatic vegetation that could be reclassified to correctly identify those water bodies that would otherwise be incorrectly categorized as land. It appears that one of the most influential artifacts in the imagery are local changes in illumination caused by clouds that were not specifically corrected for. These cloud shadows throughout the imagery appear to have the greatest impact on the vegetation classes where TV is sometimes misclassified as AV.
The LogR classification configuration determined from the August analysis was applied to the remaining mission imagery to validate the classification framework (Figure 3.8). Mission mosaics were preprocessed following the same steps used for the August 2015 imagery. Rather than selecting another 2000 samples per class, only 300 samples per class were selected for each image. This saved substantial preprocessing time but still allowed for the random selection of 260 samples per class from the 300 total reference samples per class.
Figure 3.8   Mission Mosaic Classifications

While performance for each class varies slightly based on image type, the changes in accuracy metric between image type are inconsequential (Table 3.6 & Figure 3.9). This confirms that UAS imagery collected throughout the day, or even throughout a relatively short timeframe (e.g. one week), does not need to be treated as multitemporal data. Most
apparent in Table 3.6 is the relatively poor performance of the May 2015 classification compared to other missions. The inferior performance in May 2015 can be attributed to poor image quality for some flights in the mosaic and variations to local illumination conditions, likely due to cloud shadows (Figure 3.10). In fact, these variations in illumination conditions between flights are apparent in multiple mission mosaic imagery, but March and May were most heavily impacted which corresponded to their lower classification performance.

Table 3.6  Mission Classification Median Accuracies

<table>
<thead>
<tr>
<th></th>
<th>Brightness Values</th>
<th>Radiance</th>
<th>Reflectance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dec-14</td>
<td>OW</td>
<td>TV</td>
<td>AV</td>
</tr>
<tr>
<td>UA (%)</td>
<td>91</td>
<td>79</td>
<td>79</td>
</tr>
<tr>
<td>PA (%)</td>
<td>92</td>
<td>84</td>
<td>73</td>
</tr>
<tr>
<td>OA (%)</td>
<td>-</td>
<td>-</td>
<td>83</td>
</tr>
<tr>
<td>Kappa</td>
<td>-</td>
<td>-</td>
<td>0.74</td>
</tr>
<tr>
<td>Mar-15</td>
<td>UA (%)</td>
<td>89</td>
<td>69</td>
</tr>
<tr>
<td>PA (%)</td>
<td>91</td>
<td>73</td>
<td>64</td>
</tr>
<tr>
<td>OA (%)</td>
<td>-</td>
<td>-</td>
<td>76</td>
</tr>
<tr>
<td>Kappa</td>
<td>-</td>
<td>-</td>
<td>0.64</td>
</tr>
<tr>
<td>May-15</td>
<td>UA (%)</td>
<td>84</td>
<td>68</td>
</tr>
<tr>
<td>PA (%)</td>
<td>88</td>
<td>74</td>
<td>54</td>
</tr>
<tr>
<td>OA (%)</td>
<td>-</td>
<td>-</td>
<td>72</td>
</tr>
<tr>
<td>Kappa</td>
<td>-</td>
<td>-</td>
<td>0.58</td>
</tr>
<tr>
<td>Aug-15</td>
<td>UA (%)</td>
<td>93</td>
<td>78</td>
</tr>
<tr>
<td>PA (%)</td>
<td>94</td>
<td>81</td>
<td>70</td>
</tr>
<tr>
<td>OA (%)</td>
<td>-</td>
<td>-</td>
<td>81</td>
</tr>
<tr>
<td>Kappa</td>
<td>-</td>
<td>-</td>
<td>0.72</td>
</tr>
<tr>
<td>Dec-15</td>
<td>UA (%)</td>
<td>96</td>
<td>76</td>
</tr>
<tr>
<td>PA (%)</td>
<td>97</td>
<td>76</td>
<td>74</td>
</tr>
<tr>
<td>OA (%)</td>
<td>-</td>
<td>-</td>
<td>82</td>
</tr>
<tr>
<td>Kappa</td>
<td>-</td>
<td>-</td>
<td>0.73</td>
</tr>
</tbody>
</table>

Table 3.6 Caption: Comparison of classification accuracies between missions and image correction method. Median accuracies calculated from 1000 iterations of each classification technique. (Logistic Regression Classifier (LogR); Maximum Likelihood Classifier (MLC); Random Forest classifier (RF); User Accuracy (UA); Producer Accuracy (PA); Overall Accuracy (OA); Open Water (OW); Terrestrial Vegetation (TV); Aquatic Vegetation (AV))

55
Figure 3.9 Caption: Confidence intervals of classification accuracy metrics from 1000 iterations of the logistic regression classifier for the all mission imagery configured with 260 samples per class and a 70:30 training/testing sample ratio. (Logistic Regression Classifier (LogR); Maximum Likelihood Classifier (MLC); Random Forest classifier (RF); User Accuracy (UA); Producer Accuracy (PA); Overall Accuracy (OA); Open Water (OW); Terrestrial Vegetation (TV); Aquatic Vegetation (AV))
Figure 3.10  Mission CIR Imagery

Figure 3.10 Caption: Mission mosaic imagery for (a) December 2014 (b) March 2015, (c) May 2015, (d) August 2015, and (e) December 2015

Discussion

There are two main outcomes from this work: (1) an assessment on the application of current remote sensing classification methods for UAS image classification, and (2) the identification of a practical, quick response UAS image classification configuration.
A common feature of UAS mission imagery is the collection of data throughout a day, and sometimes across multiple days for a single mission. This means that each UAS mission mosaic is essentially multitemporal data. It was therefore expected that the classification accuracy would increase by normalizing for illumination prior to the classification; however, the changes in accuracies are not statistically significant for any mission. This suggests that the temporal differences within each mission were not great enough to require normalization for changes in solar illumination in a simple classification application. Therefore, the assumption can be made that data collected over a relatively short timeframe (e.g. ~ one week) does not need to be treated as multitemporal data. While correcting for these larger scale solar illumination variations did not significantly affect the classification accuracies, a visual inspection of the full mosaic classification indicated areas where cloud shadows were causing misclassification of features. Similar cloud shadow features are going to be a common theme in operational UAS applications. Some new UAS imaging sensors incorporate upwards viewing irradiance sensors to help adjust for variable cloud conditions; however, there will still be scenarios where the UAS is in full sun but some portions in the sensor FOV are shadowed by clouds. Therefore, it will be important for future work to develop an automated method for correcting for local changes in illumination by clouds.

It was shown that classification accuracy is relatively consistent for training/testing ratios between 55:45 to 70:30. This is in agreement with the current remote sensing training/testing ratio recommendation, so future UAS image classification can confidently select any ratio within this range. While a sample size of 260 per class was deemed reasonable in the current work, results show that other sample sizes could be
used depending on the degree of confidence in classification accuracy required. Therefore, studies unable to collect 260 ground samples per class could use less and reference the tables produced to identify the effect this will have on the uncertainty in classification accuracy. While there were not significant differences in OAs due to changing sample size, it can be visually determined that there is greater uncertainty in classification OAs when fewer samples per class are used. This motivates the need for future work to further investigate whether there is significantly less uncertainty in classification OAs when more samples per class are used. This could be done by resampling the 95% confidence intervals in the current study to calculate and compare inter-quartile range (IQR). A comparison of the IQRs for 50-500 samples per class would provide additional insight about the uncertainty in classification OAs with changes to sample size.

It was determined that the changes in classification accuracies between classifiers were not significant; however, the LogR was most practical for operational UAS application because of its relatively fast runtime and its superior performance classifying OW and AV. While aquatic vegetation proved to be challenging for all classification techniques, many features classified as AV were successfully located near and on top of water (Figure 3.11).
Figure 3.11 Caption: Zoomed-in view of classification show where aquatics are constricting and completely covering streams. Zooming in farther shows an image from a boat taken near the same time of this UAS image in August 2015. This illustrates how the aquatic vegetation is floating on top of the water and acts to camouflage the water. (Open Water (OW); Terrestrial Vegetation (TV); Aquatic Vegetation (AV))

These results are beneficial from a water mapping perspective because this AV class can help identify areas where the river may be wider than it appears in the imagery. It also identifies some connecting streams that would have gone unnoticed, especially in the warm season when aquatic vegetation is prolific. Beyond the ability to successfully quantify both open water and camouflaged water bodies, those areas of dense aquatic vegetation near water bodies can help locate problematic invasive species like water hyacinth and hydrilla.
The classification consistently performed best classifying water features. As expected, the AV class was challenging to classify. This is because aquatic vegetation and terrestrial vegetation have similar spectral response signatures; it is the texture and spatial patterns that best distinguish these two types of vegetation. Adding additional texture or normalized difference vegetation index (NDVI) layers during the preprocessing procedure could help distinguish between them. Adjustments to the classification algorithm parameters would also affect the OA of the classification. Having determined effective training/testing sample ratio and training sample size for classifying features in high spatial resolution UAS imagery, future analyses can incorporate additional techniques to improve classification accuracies. This could include additional classification techniques, sensitivity analyses to identify the most optimal classification algorithm parameters, segmenting the image and conducting an object-based classification, or applying post-processing filters. Another useful analysis could combine methods from Casado et al. (2016) and the current work to provide a robust assessment on the impact of image resolution on classification accuracies.

**Conclusion**

This work set out to conduct an analysis of how current remote sensing image correction techniques, sample training/testing ratios, sample size, and classification technique apply to the classification of UAS imagery. This paper presented a series of classification configurations for the classification of high spatial resolution UAS data by varying training/testing sample ratio, training sample size, and classification technique. It was determined that any training/testing sample ratio between 55:45 to 70:30 is a reasonable option. It was also discovered that 260 samples per class was sufficient for
producing a good balance between classification accuracy and precision, although accuracies with a per class sample size of 260 were not significantly different from other samples sizes, suggesting that future research can use less samples per class at the cost of increased uncertainty in classification accuracy. It was also found that image correction techniques did not significantly improve classification accuracy; therefore, UAS imagery collected throughout a day or across multiple concurrent days does not need to be treated as multitemporal data. While it did not significantly affect overall classification performance, artifacts due to local changes in illumination conditions by cloud shadows were apparent. Additional work is needed to correct for local illumination variations in the absence of an onboard pyranometer.

The LogR classification configuration produced the best classification accuracies in the least amount of time for the high spatial resolution UAS imagery. OW features were classified with the greatest accuracy followed by TV and AV (respectively). This suggests that it is possible to use the LogR classification configuration to quickly produce a simple classification for UAS imagery. This is important for applications such as the mapping of flood inundation extent, monitoring restoration efforts, and observing changes to stream networks. Future work should investigate how additional preprocessing and post-processing steps can improve classification accuracy, while also keeping the goal of developing a rapid post-flight image analyses framework. This work could include the segmentation of the imagery to conduct an object-based image classification, incorporating additional texture and vegetation index layers, test additional classification techniques, and additional post-processing image filter procedures. There are still many questions to be answered in UAS image classifications, but the assessment
presented in this paper provides valuable classification configuration reference material for future UAS image classifications.
CHAPTER IV
CONCLUDING THOUGHTS

Overview

There have been UAS studies that successfully used current remote sensing image processing techniques to correct UAS image (Kelcey and Lucieer 2012; Wang and Myint 2015). There have also been numerous UAS studies that have adapted current remote sensing image classification configurations to generate reasonable LULC classifications (Casado et al. 2015; Feng et al. 2015; Husson et al. 2016; Laliberte and Rango 2009; Laliberte et al. 2010; Ouellette and Getinet 2016; Zaman et al. 2011). However, the common theme of these studies is that they are forcing UAS into current remote sensing techniques without questioning whether the underlying assumptions of these techniques are valid in UAS application. Therefore, there is a need for assessments to determine whether current remote sensing image processing techniques can be used without modification in UAS applications.

This dissertation made the first attempt to tackle this question by investigating the atmospheric impacts on UAS imagery, assessing the practicality of applying common atmospheric correction techniques to UAS imagery, and assessing the application of current LULC classification configurations to UAS imagery. It was discovered that the atmosphere has a statistically significant impact on the relatively low-altitude UAS
imagery. This indicates the need for the atmospheric correction of UAS imagery for certain applications that require accurate spectral information.

Under idealized conditions, Kelcey and Lucieer (2012) determined that traditional ELC techniques adequately corrected UAS imagery collected at an altitude of 100 m. Under similar conditions but in a different location, Wang and Myint (2015) also found that ELC techniques adequately corrected UAS imagery collected at an altitude of 20 m. The current study again affirmed the application of ELC techniques in UAS imagery correction by successfully conducting an ELC image correction on UAS imagery collected over a coastal estuary at an altitude of 250 m. The water-based nature of the current study required the development of a cost-effective and portable ELC framework that can be adapted in a variety of UAS application. In addition, this study contributed a more robust ELC image correction framework by including numerous sites under various conditions in the development of calibration equations. This means that the calibration equations developed better represent a variety of UAS applications and can minimize errors from both atmospheric effects and various field variables in these applications.

The final inquiry of this dissertation was the assessment of current remote sensing LULC classification techniques in a UAS application. It was determined that the temporal differences across a single mission mosaic were small enough that image corrections did not significantly impact the classification performance. Local changes in illumination conditions due to clouds did not significantly impact the overall accuracy of the UAS image classifications, but the artifacts from these local illumination conditions were visually apparent in the output classification. This outcome, in addition to the
challenges clouds posed for the ELC, suggests that there is a need for a new automated technique that can correct UAS image mosaics for local illumination condition variations.

While training/testing sample ratio requirements are in agreement with current remote sensing techniques, the most effective training sample size requirements are greater in UAS applications. Compared to past research suggesting 51-100 training samples per class (Fitzpatrick-Lins 1981; Congalton 1991), the current study found that 260 samples per class is preferred. This conclusion was based on the balance between the time required for training sample collection and confidence in the expected classification OA. While it was found that there were not significant differences between classification OA associated with the number of samples per class from 50-500, it can be visually determined that there is greater uncertainty in classification OA when fewer samples per class are used.

The three-class hierarchy for this study was rather simple; however, the distinction between the aquatic and terrestrial vegetation classes was expected to be challenging for the relatively low spectral but high spatial resolution Canon DSLR camera. This provided the opportunity to assess whether the commonly used RF significantly outperforms simpler classification techniques like LogR and MLC. Past research has alluded to the possibility that the RF classifier is needed for high spatial resolution UAS imagery. In contrast, this study found that the LogR classification technique performed best in terms of accuracies and runtime. Therefore, high spatial resolution UAS imagery does not necessarily require the implementation of the RF classifier as suggested in previous research.
Final Assessment

The current research, as well as planned future research, will make important contributions to the growing UAS applications in earth and atmospheric sciences. There is a need to further investigate the magnitude of atmospheric effects on low-altitude UAS imagery. Similar experiments, but more rigorous than the North Farm experiments, will be conducted throughout the warm season when solar conditions are appropriate. On-board meteorological sensors will be incorporated into future experiments to more rigorously quantify the external factors affecting the imagery. These experiments will begin in late April and continue through July. A mix of idealized conditions (i.e. solar zenith and clear skies) and inclement conditions (i.e. high relative humidity and varied illumination) will provide a more robust assessment of atmospheric effects on low-altitude UAS imagery in a variety of conditions. This information will be invaluable to ongoing research attempting to use UAS imagery in applications such as calibrating operational satellite sensors.

Efforts stemming from the assessment of classification configurations in Chapter III will improve accuracy and precision of the proposed LULC classification frameworks. There first needs to be the addition of ANN and SVM classification techniques in the analysis. Further investigations as to whether the confidence in classification OAs are significantly increased with increased samples per class are needed. This will be done by resampling the 95% confidence intervals in the current study and calculating inter-quartile range (IQR). A comparison of the IQRs for 50-500 samples per class will conclude whether the confidence in classification OAs are significantly greater when more samples per class are used. Continued efforts will involve the addition of different
classes, comparing per-pixel and object-based classifications, comparing effects of image resolution, incorporating additional feature layers (e.g. texture), and incorporating various post-processing techniques. Streamlining these classification frameworks will provide a semi-automated classification product that can be implemented for rapid LULC classification of UAS imagery. This product will support ongoing research modeling hydrometeorological hazards. In the event of a flood inundation event, UAS imagery will be collected and processed using the streamlined LULC classification product to support the verification of hydrological models. Beyond this, such a product will be a valuable resource to change detection, environmental monitoring, and disaster response efforts. It is hoped that the LULC classifications product can eventually be integrated into a web-based mapping application to aid decision support efforts. Such a product would be an invaluable decision support tool in hazard preparation, response, and mitigation efforts.
REFERENCES

2010 IEEE 39th Applied Imagery Pattern Recognition Workshop (AIPR), IEEE, 

Usage for Imagery Collection in Disaster Research and Management. 9th Int. 

Anderson, K., and K. J. Gaston, 2013: Lightweight unmanned aerial vehicles will 
revolutionize spatial ecology. Front. Ecol. Environ., 11, 138–146, 
doi:10.1890/120150.

Bagan, H., Y. Yasuoka, T. Endo, X. Wang, and Z. Feng, 2008: Classification of airborne 
hyperspectral data based on the average learning subspace method. IEEE Geosci. 

multi-scale comparison of object-based and pixel-based methods for detecting 
change associated with gas well drilling operations. Int. J. Remote Sens., 34, 

Di Baldassarre, G., G. Shumann, and P. Bates, 2009: Near real time satellite imagery to 
support and verify timely flood modelling. Hydrol. Process., 23, 799–803, 

using UAV photogrammetry and SPH modelling of fluids . 1–8.

Bazi, Y., and F. Melgani, 2006: Toward an Optimal SVM Classification System for 


Chou, T., M. Yeh, Y. Chen, and Y. Chen, 2010: Disaster monitoring and management by the unmanned aerial vehicle technology. ISPRS TC VII Symp., XXXVIII, 137–142.


Clark, R. N., G. A. Swayze, K. Heidebrecht, R. O. Green, and F. H. Goetz, 1995: Calibration to surface reflectance of terrestrial imaging spectrometry data: Comparison of methods.


70


Park, J., 2015: Crop Classification Using Imagery of Drone. 91–94.


