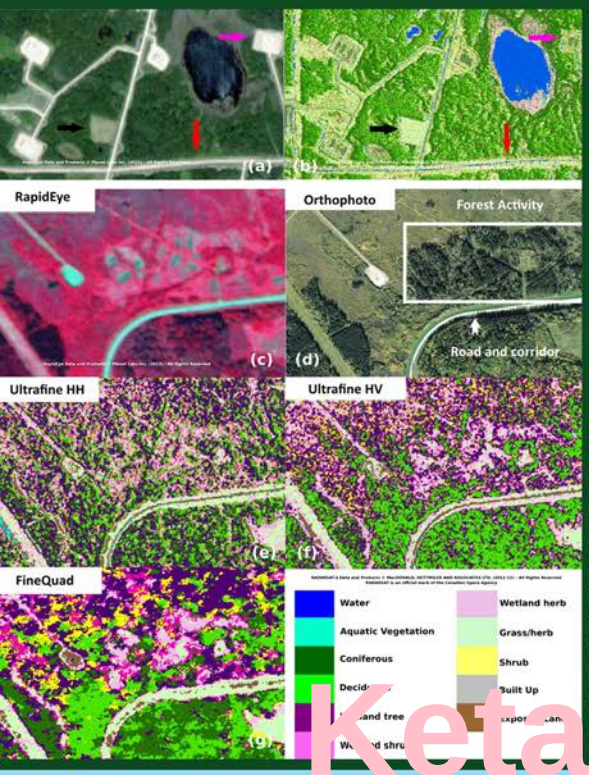
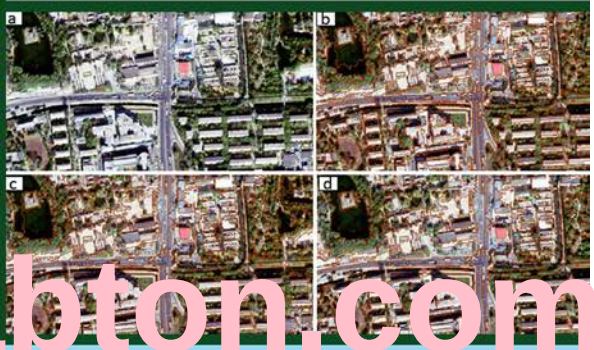
Taylor & Francis Series in Imaging Science

QIHAO WENG, Series Editor

HIGH SPATIAL RESOLUTION REMOTE SENSING

Data, Analysis, and Applications





Edited by

Yuhong He and Qihao Weng

CRC Press
Taylor & Francis Group

High Spatial Resolution Remote Sensing

Data, Analysis, and Applications

Taylor & Francis Series in Imaging Science

Series Editor

Qihao Weng
Indiana State University

Published Titles

Remote Sensing Time Series Image Processing
Qihao Weng

High Spatial Resolution Remote Sensing: Data, Analysis, and Applications
Yuhong He and Qihao Weng

For more information about this series, please visit: www.crcpress.com

High Spatial Resolution Remote Sensing

Data, Analysis, and Applications

Edited by
Yuhong He
Qihao Weng

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Printed on acid-free paper

International Standard Book Number-13: 978-1-4987-6768-2 (Hardback)

Library of Congress Cataloging-in-Publication Data

Names: He, Yuhong, 1978- editor. | Weng, Qihao, editor.
Title: High spatial resolution remote sensing : data, analysis, and applications / edited by Yuhong He and Qihao Weng.
Description: Boca Raton, FL : Taylor & Francis, 2018. | Includes bibliographical references.
Identifiers: LCCN 2018006261 | ISBN 9781498767682 (hardback : alk. paper)
Subjects: LCSH: Remote sensing. | Optical radar. | Environmental sciences--Remote sensing.
Classification: LCC G70.4 .H53 2018 | DDC 621.36/78--dc23
LC record available at https://lccn.loc.gov/2018006261

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

Contents

SECTION II Algorithms and Techniques

SECTION III Case Studies and Applications

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Preface

1 INTRODUCTION

High spatial resolution data provide a novel data source for addressing environmental questions with an unprecedented level of detail. These remote sensing data are a result of significant advances in image acquisition platforms and sensors, including satellite, manned aircraft, and unmanned aerial vehicle (UAV) platforms. Furthermore, the recent development of commercially operated satellite platforms with high spatial resolution sensors allows for the collection of a large amount of images at regular time intervals, with relatively large footprints (i.e., image swathes). For example, the WorldView series, with the WorldView-1 satellite launched on September 18, 2007, and WorldView-4 launched on November 11, 2016, are capable of resolving objects at 31 cm in the panchromatic band and at 1.24 m in 4 (or 8)-band multispectral over a 13.1-km-wide swath. For a specific study site, these image data can be easily searched and acquired at a reasonable cost through service companies, such as DigitalGlobe. In addition, the recent proliferation of UAVs has made it possible to collect images at spatial and temporal scales that would be impossible using traditional platforms. Many recent works have focused on collecting imagery using UAV-equipped multispectral, hyperspectral, and thermal sensors as well as laser scanners at centimeter resolutions.

High (meters) and ultra-high (centimeters) spatial resolution images open a door for fine-scale analysis of objects at the Earth's surface. A number of scientific journal articles have highlighted the usefulness of high spatial resolution remote sensing, including the use of remote sensing in studying the physical environmental system, the human system, and the interactions between them. Examples in physical environmental studies include fine-scale forest inventory (Mora et al., 2013), wetland plant community identification (Zweig et al., 2015), grassland mapping (Lu and He, 2017), and water resources (Debell et al., 2016). In terms of the human system, high spatial resolution remote sensing has been used to study urban impervious surfaces (Yang and He, 2017), public health (Hartfield et al., 2011), and epidemics (Lacaux et al., 2007). As for human-environment interactions, high-resolution remote sensing has been used for land degradation (Wiesmair et al., 2016), precision farming (Zarco-Tejada et al., 2013), water and air pollution (Yao et al., 2015), and natural hazards (e.g., earthquakes, typhoons, floods, landslides) (Joyce et al., 2009).

This increased spatial resolution exasperates the intraclass variability found in an image. For example, in a grassland scene, vegetation leaves, gaps, shadows, and stems are all visible in the pixels of a high spatial resolution image. While this information is potentially useful for mapping purposes, the added details in a high-resolution image pose challenges for image segmentation and feature selection. Furthermore, the number of detectable entities or classes increases with spatial resolution. Traditional information extraction techniques may not operate well at high spatial resolutions due to large data volume and heterogeneous spectral information (Wulder et al., 2004), spurring the need for the development of innovative image processing techniques. To

this end, techniques operated in the spatial domain have the potential for successfully extracting information from high spatial resolution images (Culvenor, 2003).

To effectively utilize information contained in high spatial resolution imagery, some key questions must be addressed, including

1. What are the challenges of using new sensors and new platforms?
2. What are the cutting-edge methods for fine-level information extraction from high spatial resolution images?
3. How can high spatial resolution data improve the quantification and characterization of physical-environmental or human patterns and processes?

This book intends to address the above-mentioned questions in three separate parts: (1) data acquisition and preprocessing, (2) algorithms and techniques, and (3) case studies and applications. This chapter aims to provide an overview of the book and an outline of each chapter. Section 2 discusses the opportunities and challenges of using new sensors and platforms and of using high spatial resolution remote sensing data. It also introduces the chapters that address recent developments in platforms, sensors, and data with a focus on UAVs. Section 3 addresses the issues related to high spatial image processing and introduces cutting-edge methods. Section 4 summarizes state-of-the-art high spatial resolution applications and demonstrates how high spatial resolution remote sensing can support the extraction of detailed information needed in different systems. Readers should bear in mind that platforms, sensors, data, methods, and applications are related to each other. Although we intend to organize the chapters based on their primary focuses, we also acknowledge that the authors of each chapter may tell a complete story, sometimes from introducing an image acquisition system all the way to a specific application.

2 IMAGE ACQUISITION SYSTEMS AND PREPROCESSING

An image acquisition system, which includes a platform and one or more sensor(s), plays an important role in the sense of how efficiently the areas under study can be surveyed. Traditional remote sensing platforms include satellite, aircraft, and space shuttle platforms, while a more recent addition is the UAV remote sensing system. Each of these platforms has its own technological and operational specifications (Table 1). Satellite platforms can survey large geographical areas, but their spatial resolutions are relatively coarse, the sensors are fixed, and the image acquisition plan is out of the end user's control. In other words, most high spatial resolution satellites can only acquire data after they have been programmed or paid to do so (Joyce et al., 2009). One exception is that during a major disaster event, satellite operators usually schedule imagery collection without needing an approved request.

Different from satellite platforms, the use of both manned and unmanned aircraft can be tailored to meet the needs of an end user, and acquisition sensors and parameters can be adjusted to ensure the best possible imagery. In terms of surveying area, manned aircraft can survey in kilometer squares with ideal sensors, while UAVs are only good for studying relatively small areas due to their limited battery life. However, UAVs have attracted increased attention in recent years due to their high

TABLE 1
Typical High Spatial Resolution Platforms and Their Technological and Operational Specifications

Parameters \ Platforms	Satellite	Airborne	UAV
Footprint	>100 km²	10–100 km²	1–10 km²
Sensor	Multispectral Hyperspectral Thermal Microwave	Multispectral Hyperspectral Thermal Microwave LiDAR	Multispectral Hyperspectral Thermal Microwave LiDAR
Spatial resolution	0.3 cm to 10 m	0.5 to 5 m	Centimeters
Temporal resolution	Day	Hours	Minutes
Cost of the platform	Very costly	Costly	Affordable
Cost of a scene	>$2000	~$1000	~$200
Image preprocessing	Procedures/parameters available	Procedures/parameters unavailable	Procedures/parameters unavailable

Source: Adapted from Toth, C., Jóźków, G. 2016. *ISPRS Journal of Photogrammetry and Remote Sensing*. 115: 22–36.

flexibility, low operational costs, and ability to collect very high spatial resolution images. In comparison with manned aircraft, which can mount large sensors, UAVs with limited payloads can only carry compatible sensors. With the advancement in sensing and computing technologies, sensors have become more compatible. It is hoped that multiple sensors, such as multiple cameras and/or LiDAR sensors, can be mounted on UAVs in the near future (Asner et al., 2012). However, aerial image acquisitions involve additional cost and efforts for image preprocessing because they are a function of the camera or sensor optics. The image preprocessing efforts include, but are not limited to, obtaining ground control data, performing specific radiometric correction, and image mosaicking. Clearly, the optimal platform depends not only on the size of the study area and the objective of study, but also on the available budget and resources.

Many studies have reviewed challenges in relation to UAV image collection and preprocessing (e.g., Whitehead and Hugenholtz, 2014). It is almost impossible to process UAV images using the existing image preprocessing tools because they were developed mainly for well-calibrated platforms and sensors. A UAV acquires images with inconstant scales, large overlaps, variable image orientations, and high amounts of relief displacement (Hardin and Jensen, 2011). In spite of these challenges, UAV remote sensing systems have proved useful for many different applications (see Chapters 1, 2, and 3). In Chapter 1**,** UAV remote sensing technology with high spatial and temporal resolutions is considered critical for high-throughput phenotyping and precision agriculture. In the field experiments, the authors used a variety of UAV platforms

and sensors, including PrecisionHawk Lancaster 5 with a Nikon 1 J4 visual camera, Tuffwing UAV Mapper with a Sony A6000 visual camera, and DJI Matrice 100 UAV with a MicaSense RedEdge multispectral camera. Many considerations were applied during image acquisition and preprocessing steps. The authors conclude that UAV has great potential for obtaining imagery that can provide timely and accurate information for decision support, but further research regarding UAV engineering configurations that enable navigation, and communication and data transfer capabilities are needed. The authors point out a variety of issues related to UAV imagery preprocessing, diagnostic analysis of environmental and plant conditions, and artificial intelligence for decision support. Similarly, Chapters 2 and 3 also describe the need to build a UAV-based imaging system, but with a focus on hyperspectral imagers. The author highlights the need for a UAV-based hyperspectral system, factors to be considered when building and operating a UAV-based hyperspectral system, and the methods for preprocessing UAV-based hyperspectral data.

Other high spatial resolution data from a variety of sensors, including optical, LiDAR, and synthetic-aperture radar (SAR), are introduced in Chapter 4, 5, and 6 with a focus on data preparation and preprocessing. In particular, Chapter 4 provides a brief overview of data preparation needed for integrating high-resolution multispectral and LiDAR data through a case study of mapping coastal wetlands. Using these data, the authors quantified important structural metrics for different wetland compositional types, and suggested that the combined use of LiDAR, multispectral remote sensing, and terrain metrics achieved the best model performance. The authors suggest new research directions to address the challenges often experienced in the integration of LiDAR and multispectral data. Chapter 5 discusses multiview image matching steps, and introduces advanced image matching techniques. In addition, multiview image matching accuracy and related challenges, limitations, and opportunities are also discussed. Chapter 6 focuses on high-resolution radar data acquisition, preprocessing, and processing. The chapter starts with a nonmathematical presentation of the fundamentals of SAR that are relevant to the high-resolution imaging of natural environments, followed by basic SAR image preprocessing, image processing, four SAR satellite systems, and case studies of SAR application for identifying manmade and natural features.

3 HIGH SPATIAL RESOLUTION IMAGE PROCESSING

The availability of enormous amounts of high spatial resolution remote sensing data has necessitated the development of appropriate image processing methods. This book presents a few algorithms and workflows, including structure from motion (SfM) techniques, stepwise procedures, spectral unmixing, object-based image classification, and convolutional neural networks to extract information from different high spatial resolution data. SfM reconstructs 3-D geometry and camera position from a sequence of 2-D images captured from multiple viewpoints (Ullman, 1979). This algorithm was developed decades ago, but has recently become popular for processing UAV remote sensing images. Using a consumer-grade DJI Phantom 3 Professional Quadcopter with a RGB camera, the authors in Chapter 7 collected UAV images from multiple positions and analyzed them using various SfM programs to establish a workflow for generating reliable estimates of wood chip volume.

While optical image processing exploits mostly spectral properties, in the case of airborne laser scanning (ALS) data, methods should be developed to take advantage of both range and spectral information from a single source. Chapter 8 proposes a stepwise procedure including land-cover classification, tree crown segmentation, and regression modeling to examine the feasibility of multispectral ALS data for tree carbon stock estimation. The proposed workflow provides a benchmark for processing emerging multispectral ALS data in land-cover mapping.

Linear spectral unmixing has captured growing attention for its ability to extract fine-scale features (Yang et al., 2014). When endmembers are identified in a 2-D spectral mixing space, all pixels in the images can be processed as a linear combination of the identified endmembers. However, the choice of an appropriate triangle structure (from a spectral mixing space) is yet to be automated. To address this gap, Chapter 9 introduces an indicator that is capable of selecting suitable feature space objectively, and thus holds great potential for automated high-resolution remote sensing image unmixing.

High spatial resolution image classification has generally advanced from pixel-based to object-based approaches, with the latter delineating real-world objects composed of many pixels (Wulder et al., 2004). The objects thus add several new possible layers to image analysis ranging from spectral descriptive statistics to textural and geometric information. However, object-based image analysis also introduces a number of analytical issues to consider. In the image segmentation step, an inappropriate scale can negatively affect classification results, and therefore researchers must decide which segmentation scale to apply to the image (Ma et al., 2015). Chapter 10 provides a review on five scale-selection methods, and compares their advantages and disadvantages using WorldView-2 imagery. The authors suggest that scale-selection methods be explored before choosing an optimal scale to apply for segmenting objects.

Computer vision aims to mimic the human visual system for automatic image extraction, analysis, and interpretation. Chapter 11 summarizes the relevant computer vision technologies with a focus on convolutional neural networks. Specifically, the authors discuss how computer vision can be adapted to work in ecological studies, and introduce experiments to assess the effectiveness of convolutional networks in scenarios similar to the camera trap scenario. The chapter ends with several considerations for the use of these techniques in ecological studies.

4 CASE STUDIES AND APPLICATIONS

Using various high spatial resolution data, Part 3 of this book covers a range of unique applications. For grasslands, UAV-based multispectral images were used in Chapter 12 to investigate vegetation biophysical and biochemical properties in a tall grassland, while manned-aircraft-based hyperspectral images were used in Chapter 13 to invert a radiative transfer model for mapping leaf chlorophyll in a mixed grassland. Chapter 12 demonstrates that canopy leaf area index (LAI) and chlorophyll content can be accurately retrieved from UAV-acquired imagery using spectral indices. Chapter 13 concludes that the inverted model is able to estimate leaf chlorophyll content for green canopies with high accuracy, but overestimates both mixed green and brown canopies and mixed green and brown canopies with exposed soil.

Both Chapter 14 and 15 delineate wetlands, with Chapter 14 focusing on wetland mapping using GeoEye-1 images, and Chapter 15 characterizing geomorphic and biophysical properties of wetlands with airborne LiDAR. Chapter 14 indicates that wetland complexity, including varying sizes and shapes and the aquatic vegetation communities found within, can be detected using high spatial resolution optical imagery. Airborne LiDAR has gained popularity in wetland science research in the past two decades and Chapter 15 reviews this expanding field with examples from a wide range of wetland ecosystem types. Using a case study conducted in a northern peatland complex, the authors demonstrate concepts related to the accuracy of LiDAR-derived ground surface elevations, as well as geomorphic and hydrologic analysis.

When dealing with karst, one of the most fragile and heterogeneous landscapes, high spatial resolution imagery can provide detailed information that aids in the exploration of mechanisms of vegetation dynamics. In Chapter 16, vegetation cover in a degraded karst area was extracted using multispectral high spatial resolution ALOS imagery. The authors conclude that high spatial resolution imagery, when processed with the multiple endmember spectral mixture analysis approach, is able to successfully extract fine-scale features in karst areas.

Chapter 17 uses IKONOS multispectral and panchromatic images to estimate cherry orchard acreage. High spatial resolution imagery combined with object-based image analyses was found to be effective in estimating the cherry orchard acreage, within a $\pm 3.1\%$ error margin of the U.S. Department of Agriculture's census data. The author suggests using a UAV multispectral system to collect fine-resolution images because it may provide detailed information on fruit crop yield and/or potential diseases that may adversely affect orchard trees.

5 SUMMARY

High spatial resolution remote sensing data have proven useful for the extraction of ground features at a level not possible with medium- or coarse-resolution images. However, until now, the novelty of UAV platforms and sensor configurations, high cost of commercial-satellite-based data, inconsistent revisit times, and underdeveloped image processing methods have limited the usefulness of high spatial resolution images for large area applications at an operational level, not to mention real-time assessment and monitoring. The best practice in high spatial resolution remote sensing data acquisition and processing has yet to be developed.

Spatial resolution, image swath, spectral features, and temporal revisit are all critically important in determining whether a particular platform and/or sensor or data processing algorithm is capable of providing specific assessment and monitoring capabilities. It often takes at least 24 hours, or even several days, before a satellite operator can respond to an image acquisition request. This is because response time is determined by the position of the satellite within its orbit, as well as by weather conditions in the case of optical sensors (Joyce et al., 2009). Other than satellite platforms, both manned and unmanned airborne platforms are capable of acquiring images for real-time mapping and monitoring, but are limited to local scales. Further, data preprocessing, processing, and delivery systems are currently being developed

for manned and unmanned systems and their usefulness is yet to be validated. It is expected that in the near future, end users will be able to receive near-real-time images and mapping products, from manned or unmmaned systems, for an area or a feature of interest.

No single platform, sensor, data type, or processing technique will work for all cases. Multisource data integration, multiresolution exploitation, and multitemporal analysis will likely be key future lines of research (Zhang, 2010) for real-time assessment and monitoring. The application-driven strategy has been encouraged in past decades, with the hope of finding an effective solution for a specific problem. However, a data-driven analysis is possible in the era of big data, and one may achieve better real-time assessment and monitoring using multisource remote sensing big data.

The chapters in this book provide a snapshot of cutting-edge high spatial resolution remote sensing image collection, preprocessing, processing, and applications. More advances in these areas can be expected in the near future. We hope that the collection of chapters in this book will provide a useful benchmark for the high spatial resolution remote sensing community and inspire more studies that would address important scientific and technical challenges in current and future high spatial remote sensing data acquisition, processing, and applications.

ACKNOWLEDGMENTS

We wish to thank all of the authors for their invaluable contributions. Thanks go to the reviewers who provided constructive and comprehensive reviews to improve the quality of the book. We thank the research assistants Eliza King, Anna Shadrova, and Mitchell Bonney for their contribution to formatting and editing the chapters, and for their help in communicating with the contributors and reviewers. Special thanks go to the senior editor from CRC Press, Taylor & Francis Group, Ms. Irma Shagla Britton, who helped to keep the project on schedule.

Yuhong He
University of Toronto
Qihao Weng
Indiana State University

LIST OF REVIEWERS

Dr. Holly Croft, Department of Geography, University of Toronto, Toronto, Canada.
Dr. Juan Carlos Fernandez Diaz, Department of Civil and Environmental Engineering, University of Houston, Houston, Texas.
Dr. Yuhong He, Department of Geography, University of Toronto Mississauga, Mississauga, Canada.
Dr. Bing Lu, Senior Remote Sensing Scientist, Fiera Biological Consulting Ltd., Edmonton, Canada.
Dr. Cameron Proctor, Department of Geography, University of Toronto Mississauga, Mississauga, Canada.

Dr. Chen Shi, Beijing Advanced Innovation Center for Imaging Technology, Capital Normal University, Beijing, China.
Dr. Jian Yang, Department of Geography, University of Toronto, Toronto, Canada.
Dr. Chuiqing Zeng, Canada Centre for Inland Waters (CCIW), Environment and Climate Change Canada, Burlington, Canada.

REFERENCES

Asner, G.P., Knapp, D.E., Boardman, J., Green, R.O., Kennedy-Bowdoin, T., Eastwood, M., Martin, R., Anderson, C., Field, C.B. 2012. Carnegie Airborne Observatory-2: Increasing science data dimensionality via high-fidelity multi-sensor fusion. *Remote Sensing of Environment*. 124: 454–465.

Culvenor, D. 2003. Extracting individual tree information: A survey of techniques for high spatial resolution imagery. Pages 255–277 in Wulder, M., Franklin, S., eds. *Remote Sensing of Forest Environments: Concepts and Case Studies*. Boston: Kluwer Academic.

DeBell, L., Anderson, K., Brazier, R.E., King, N., Jones L. 2016. Water resource management at catchment scales using lightweight UAVs: Current capabilities and future perspectives. *Journal of Unmanned Vehicle Systems*. 4: 7–30.

Hardin, P.J., Jensen, R.R. 2011. Small-scale unmanned aerial vehicles in environmental remote sensing: Challenges and opportunities. *GIScience & Remote Sensing*. 48 (1): 99–111.

Hartfield, K.A., Landau, K.I., Leeuwen, W.J.D. 2011. Fusion of high resolution aerial multispectral and LiDAR data: Land cover in the context of urban mosquito habitat. *Remote Sensing*. 3: 2364–2383.

Joyce, K., Belliss, S., Samsonov, S., McNeill, S., Glassey, P. 2009. A review of the status of satellite remote sensing and image processing techniques for mapping natural hazards and disasters. *Progress in Physical Geography*. 33 (2): 183–207.

Lacaux, J.-P., Tourre, Y.M., Vignolle, C., Ndione, J.-A., Lafaye, M. 2007. Classification of ponds from high-spatial resolution remote sensing: Application to Rift Valley Fever epidemics in Senegal. *Remote Sensing of Environment*. 106: 66–74.

Lu, B., He, Y. 2017. Optimal spatial resolution of unmanned aerial vehicle (UAV)-acquired imagery for species classification in a heterogeneous ecosystem. *GIScience and Remote Sensing*. 55: 205–220.

Ma, L., Cheng, L., Li, M., Liu, Y., Ma, X. 2015. Training set size, scale, and features in geographic object-based image analysis of very high resolution unmanned aerial vehicle imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*. 102: 14–27.

Mora, B., Wulder, M.A., Hobart, G., White, J.C., Bater, C.W., Gougeon, F.A., Varhola, A., Coops, N.C. 2013. Forest inventory stand height estimates from very high spatial resolution satellite imagery calibrated with lidar plots. *International Journal of Remote Sensing*. 34 (12): 4406–4424.

Toth, C., Jóźków, G. 2016. Remote sensing platforms and sensors: A survey. *ISPRS Journal of Photogrammetry and Remote Sensing*. 115: 22–36.

Ullman, S. 1979. The interpretation of structure from motion. *Proceedings of the Royal Society of London Series B, Biological Sciences*. 203: 405–426.

Whitehead, K., Hugenholtz, C.H. 2014. Remote sensing of the environment with small unmanned aircraft systems (UASs), part 1: A review of progress and challenges. *Journal of Unmanned Vehicle Systems*. 2 (3): 69–85. doi:10.1139/juvs-2014-0007.

Wiesmair, M., Feilhauer, H., Magiera, A., Otte, A., Waldhardt, R. 2016. Estimating vegetation cover from high-resolution satellite data to assess grassland degradation in the Georgian Caucasus. *Mountain Research and Development*. 36 (1): 56–65.

Wulder, M.A., Hall, R.J., Coops, N.C., Franklin, S.E. 2004. High spatial resolution remotely sensed data for ecosystem characterization. *BioScience*. 54: 511–521.

Yang, J., He, Y. 2017. Automated mapping of impervious surfaces in urban and suburban areas: Linear spectral unmixing of high spatial resolution imagery. *International Journal of Applied Earth Observation and Geoinformation*. 54: 53–64.

Yang, J., He, Y., Oguchi, T. 2014. An endmember optimization approach for linear spectral unmixing of fine-scale urban imagery. *International Journal of Applied Earth Observation and Geoinformation*. 27: 137–146.

Yao, F., Wang, C., Dong, D., Luo, J., Shen, Z., Yang, K. 2015. High-resolution mapping of urban surface water using ZY-3 multi-spectral imagery. *Remote Sensing*. 7 (9): 12336–12355.

Zarco-Tejada, P.J., Guillén-Climent, M.L., Hernández-Clemente, R., Catalina, A., González, M.R., Martín, P. 2013. Estimating leaf carotenoid content in vineyards using high resolution hyperspectral imagery acquired from an unmanned aerial vehicle (UAV). *Agricultural and Forest Meteorology*. 171: 281–294.

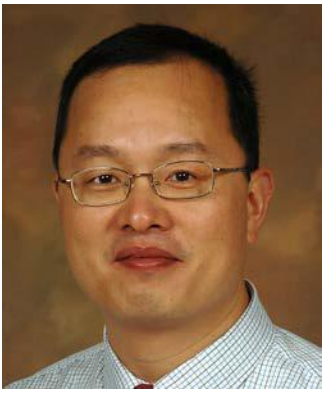
Zhang, J. 2010. Multi-source remote sensing data fusion: Status and trends. *International Journal of Image and Data Fusion*. 1: 5–24. doi:10.1080/19479830903561035.

Zweig, C.L., Burgess, M.A., Percival, H.F., Kitchens, W.M. 2015. Use of unmanned aircraft systems to delineate fine-scale wetland vegetation communities. *Wetlands*. 35: 303–309.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Editors

Yuhong He is an associate professor and geographic information system (GIS) program advisor in the Department of Geography at the University of Toronto Mississauga (UTM). She received her PhD in geography from the University of Saskatchewan in 2008, joined UTM as an assistant professor in 2009, and was promoted to associate professor in 2014. Her research focuses on remote sensing techniques, spatial analysis, climate data, and ecosystem modeling of natural or managed systems. Her approach exploits observation windows at various spatial and temporal scales to investigate the linkages between changes in plant biophysical and biochemical traits in relation to environmental and anthropogenic driving factors. Her work has been supported by major grants from various internal and external sources including the Natural Sciences and Engineering Research Council (NSERC), Canada Foundation for Innovation (CFI), and Early Researcher Award (ERA). Her innovative research at the start of her career has been recognized. She received the Early Researcher Award of Ontario in 2015 and the Early Career Scholar Award from the Remote Sensing Specialty Group at the American Association of Geographers (AAG) 2011 annual meeting. Yuhong He and her students have published more than 50 peer-reviewed journal articles and presented over 80 papers at national and international conferences. She has been serving as an associate editor for the *Journal of Applied Remote Sensing*, and is an editorial advisory board member of the *ISPRS Journal of Photogrammetry and Remote Sensing*. Her past service included being a director of the AAG Remote Sensing Specialty Group (2012–2014).



Qihao Weng, PhD, IEEE Fellow, is the director of the Center for Urban and Environmental Change and a professor at Indiana State University, and worked as a senior fellow at the National Aeronautics and Space Administration (NASA) from December 2008 to December 2009. He received his PhD degree from the University of Georgia in 1999. Weng is currently the lead of the Group on Earth Observation (GEO) Global Urban Observation and Information Initiative, and serves as an editor-in-chief of the *ISPRS Journal of Photogrammetry and Remote Sensing* and the series editor of the Taylor & Francis Series in Remote Sensing Applications.

He has been the organizer and program committee chair of the biennial International Workshop on Earth Observation and Remote Sensing Applications conference series sponsored by IEEE, the International Society for Photogrammetry and Remote Sensing (ISPRS), and GEO since 2008, a national director of the American Society for Photogrammetry and Remote Sensing from 2007 to 2010, and a panelist of the U.S. Department of Energy's (DOE)'s Cool Roofs Roadmap and Strategy in 2010. In 2008, Weng received a prestigious NASA senior fellowship. He is also the recipient of the Outstanding Contributions Award in Remote Sensing in 2011 and the Willard and Ruby S. Miller Award in 2015 for his outstanding contributions to geography, both from the American Association of Geographers. In 2005 at Indiana State University, he was selected as a Lilly Foundation Faculty Fellow, and in the following year he also received the Theodore Dreiser Distinguished Research Award. In addition, he was the recipient of the 2010 Erdas Award for Best Scientific Paper in Remote Sensing (first place) and the 1999 Robert E. Altenhofen Memorial Scholarship Award, which were both awarded by the American Society for Photogrammetry and Remote Sensing. He was also awarded the Best Student-Authored Paper Award by the International Geographic Information Foundation in 1998. Weng has been invited to give 100 talks by organizations and conferences held in the United States, Canada, China, Brazil, Greece, the United Arab Emirates, and Hong Kong. Weng's research focuses on remote sensing applications to urban environmental and ecological systems, land-use and land-cover changes, urbanization impacts, environmental modeling, and human-environment interactions. Weng is the author of 210 articles and 11 books. According to Google Scholar, as of April 2018, his Science Citation Index (SCI) citation reached 13,000 (h-index of 52), and 33 of his publications had more than 100 citations each. Weng's research has been supported by funding agencies that include the National Science Foundation (NSF), NASA, U.S. Geological Survey (USGS), U.S. Agency for International Development (USAID), National Oceanic and Atmospheric Administration (NOAA), National Geographic Society, European Space Agency, and Indiana Department of Natural Resources.

Contributors

Bob Avant
Texas A&M AgriLife Research
Texas A&M University
College Station, Texas

Muthukumar V. Bagavathiannan
GEOSAT Center
and
Department of Soil and Crop Sciences
Texas A&M University
College Station, Texas

Michael P. Bishop
GEOSAT Center
and
Department of Geography
Texas A&M University
College Station, Texas

Joseph R. Buckley
Department of Physics and Space Science
Royal Military College of Canada
Kingston, Ontario, Canada

Gang Chen
Laboratory for Remote Sensing and Environmental Change
Department of Geography and Earth Sciences
University of North Carolina at Charlotte
Charlotte, North Carolina

Xinqu Chen
Department of Geography and Environmental Management
University of Waterloo
Waterloo, Ontario, Canada

Dale A. Cope
Department of Mechanical Engineering
Texas A&M University
College Station, Texas

Shihong Du
Institute of Remote Sensing and GIS
Peking University
Beijing, China

Anthony M. Filippi
GEOSAT Center
and
Department of Geography
Texas A&M University
College Station, Texas

Michael Guerzhoy
Department of Computer Science
University of Toronto
Toronto, Ontario, Canada

Monika Havelka
Department of Geography
University of Toronto Mississauga
Mississauga, Ontario, Canada

Dirk B. Hays
GEOSAT Center
and
Department of Soil and Crop Sciences
Texas A&M University
College Station, Texas

Yuhong He
Department of Geography
University of Toronto Mississauga
Mississauga, Ontario, Canada

Travis L. Howell
Center for Geospatial Analytics
and
Department of Forestry & Environmental Resources
North Carolina State University
Raleigh, North Carolina

Da Huo
GEOSAT Center
and
Department of Geography
Texas A&M University
College Station, Texas

Jonathan Li
Department of Geography and Environmental Management
University of Waterloo
Waterloo, Ontario, Canada

Bing Lu
Department of Geography
University of Toronto Mississauga
Mississauga, Ontario, Canada

Kin M. Ma
Geography & Sustainable Planning
Grand Valley State University
Allendale, Michigan

Lonesome Malambo
Department of Ecosystem Science and Management
Texas A&M University
College Station, Texas

Bill McCutchen
Texas A&M AgriLife Research
Texas A&M University
College Station, Texas

Koreen Millard
Department of Geography and Environmental Studies
Carleton University
Ottawa, Ontario, Canada

Dongping Ming
School of Information Engineering
China University of Geosciences
Beijing, China

Amy B. Mui
Environmental Science
Dalhousie University
Halifax, Nova Scotia, Canada

Seth C. Murray
Department of Soil and Crop Sciences
Texas A&M University
College Station, Texas

Jeffrey A. Olsenholler
GEOSAT Center
and
Department of Geography
Texas A&M University
College Station, Texas

Sorin C. Popescu
GEOSAT Center
and
Department of Ecosystem Science and Management
Texas A&M University
College Station, Texas

Cameron Proctor
Department of Geography
University of Toronto Mississauga
Mississauga, Ontario, Canada

Xiangkun Qi
Institute of Subtropical Agriculture
Chinese Academy of Sciences
Changsha, Hunan, China

Nithya Rajan
Department of Soil and Crop Sciences
Texas A&M University
College Station, Texas

Murray Richardson
Department of Geography and Environmental Studies
Carleton University
Ottawa, Ontario, Canada

William L. Rooney
Department of Soil and Crop Sciences
Texas A&M University
College Station, Texas

Joshua Seltzer
Department of Computer Science
University of Toronto
Toronto, Ontario, Canada

Kunwar K. Singh
Center for Geospatial Analytics
and
Department of Forestry &
Environmental Resources
North Carolina State University
Raleigh, North Carolina

Vijay P. Singh
Department of Soil and Crop Sciences
Texas A&M University
College Station, Texas

Lindsey Smart
Center for Geospatial Analytics
and
Department of Forestry &
Environmental Resources
North Carolina State University
Raleigh, North Carolina

J. Alex Thomasson
Department of Biological and
Agricultural Engineering
Texas A&M University
College Station, Texas

Alexander Tong
Department of Geography
University of Toronto Mississauga
Mississauga, Ontario, Canada

John Valasek
GEOSAT Center
and
Department of Aerospace Engineering
Texas A&M University
College Station, Texas

Misty Vidrine
Texas A&M AgriLife Research
Texas A&M University
College Station, Texas

Jinfei Wang
Department of Geography
University of Western Ontario
Ontario, Canada

Kelin Wang
Institute of Subtropical Agriculture
Chinese Academy of Sciences
Changsha, Hunan, China

Jian Yang
Department of Geography
University of Toronto Mississauga
Mississauga, Ontario, Canada

Brennan W. Young
GEOSAT Center
and
Department of Geography
Texas A&M University
College Station, Texas

Chuiqing Zeng
Department of Geography
University of Western Ontario
Ontario

Chunhua Zhang
Department of Geography and
Geology
Algoma University
Sault Ste. Marie, Ontario, Canada

Xiuyuan Zhang
Institute of Remote Sensing and GIS
Peking University
Beijing, China

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Section I

Data Acquisition and Preprocessing

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

1 High-Resolution UAS Imagery in Agricultural Research *Concepts, Issues, and Research Directions*

Michael P. Bishop, Muthukumar V. Bagavathiannan, Dale A. Cope, Da Huo, Seth C. Murray, Jeffrey A. Olsenholler, William L. Rooney, J. Alex Thomasson, John Valasek, Brennan W. Young, Anthony M. Filippi, Dirk B. Hays, Lonesome Malambo, Sorin C. Popescu, Nithya Rajan, Vijay P. Singh, Bill McCutchen, Bob Avant, and Misty Vidrine

CONTENTS

1.1 INTRODUCTION

Crop yield (production per acre) has increased up to eightfold over more than a century of concerted scientific research into agricultural systems and genetic improvement of crops (Brummer et al. 2011; Hall and Richards 2013). The global population is expected to increase to over 9 billion people by 2050, and increases in the standard of living will require more food, fiber, and fuel (Godfray et al. 2010; FAO 2017). There will also be new constraints from climate change, decreased availability of fertilizer and irrigation water inputs, and external pressure from consumers, companies, and governments to produce agricultural products in a more sustainable way to address food quality and security issues (Godfray et al. 2010; Tilman et al. 2011; Challinor et al. 2014; FAO 2017). This will require new developments and utilization of agricultural and information technologies.

Given the advent of unmanned aerial systems (UASs) and high-resolution multispectral imagery, managers and agricultural specialists have new tools and information for optimizing management decisions and enabling precision agriculture solutions. Enabling technologies include global positioning systems (GPSs), high-resolution multispectral and hyperspectral sensors, geographic information system (GIS) technology, and field sensor and network monitoring capabilities. In addition, these information technologies also support improved information production and decision support capabilities using knowledge representation, artificial intelligence, and visualization techniques. Collectively, such enabling geospatial technologies permit mapping and monitoring of crops (Hunt et al. 2010; Torres-Sánchez et al. 2014, 2015), weed assessment and management (Armstrong et al. 2007; Gray et al. 2008; López-Granados 2010; Eddy et al. 2013; Peña et al. 2013; Torres-Sánchez et al. 2013), plant stress detection (Hunt et al. 2010; Zarco-Tejada et al. 2012), and many other precision agricultural applications throughout every aspect of preplanting through postharvest (Herwitz et al. 2004; Blackburn 2006; Castaldi et al. 2017). Nevertheless, although the potential for geospatial technologies to significantly impact agricultural research and management practices is high, numerous concepts and issues must be addressed to generate diagnostic and actionable information that can be relied upon for scientific inquiry, management, and optimization problem solving. There are numerous issues associated with UAS data acquisition strategies, image preprocessing, geospatial data management, information production, information synthesis, and the development and evaluation of agricultural decision support systems that are specific to agriculture.

Perhaps the greatest agriculture challenge is understanding the complex interactions between plants and their agricultural environment. Understanding these interactions may enable optimal management of plant growth by controlling input factors to maximize crop yield, sustainability, safety, and nutrition. Using objective information on these and other measurable factors, and applying site-specific management, is the basis of *precision agriculture*. Another major challenge is to breed and select the best genetics in a cultivar to meet crop production goals. Automated and semiautomated methods of plant phenotypes are termed *high-throughput phenotyping* (HTP), which seeks to use sensors deployed on various types of platforms to conduct measurements rapidly so that larger populations and/or more replicates can be measured and

assessed. Consequently, developing and evaluating UAS technology and high-resolution imaging systems can significantly assist in precision agriculture and high-throughput phenotyping activities to help managers make better decisions and assist plant breeders and geneticists to screen more varieties more quickly and accurately, and with less cost. Even more exciting is the possibility of allowing plant breeders to observe traits or phenotypes that have never before been possible, such as the utility of UAS technology to temporally screen large fields and estimate plant growth curve trajectories.

Although UAS technology and the collection of high-resolution imagery can help address primary agricultural challenges, numerous concepts and research directions must be further investigated and formalized to address a variety of scientific and engineering issues associated with the development and evaluation of these technologies. Consequently, the primary objective of this chapter is to report on selected aspects of our multidisciplinary experiences involving the Texas A&M AgriLife Unmanned Aerial Systems Project at the field research facility near College Station, Texas, where we have addressed key issues in utilizing geospatial technologies. We accomplish this by providing fundamental background information about phenotyping, precision agriculture, and geospatial technology issues in agriculture. We then provide information about the study area and go into details about challenges involved with flight planning, image acquisition, data preprocessing, and information extraction. We then discuss challenges and issues related to crop mapping, plant stress detection, and weed management.

1.2 BACKGROUND

1.2.1 Phenotyping

Crop yields need to increase at a rate greater than 2% per year to ensure adequete food supply to meet projected human populations under a changing climate (Godfray et al. 2010; Cairns et al. 2013; Gleadow et al. 2013). Current yield increase rates are consistently lower than projected and average annual crop yield goals are not met in much of the land currently used for cereal production (Ray et al. 2013).

Systematic improvements in crop yield and quality are typically due to new technology adoption of crop production methods (agronomic approaches) and the development of cultivars that are genetically more productive (plant breeding approaches). While the exact contribution attributed to each approach varies with each crop, on average these two approaches have contributed equally to increases in productivity (Duvick 2005). While these increases have been impressive, further increases will require more information, creativity, and integration to achieve the same levels.

As a science, crop improvement is a relatively new endeavor with its nascent beginnings occurring in the early twentieth century with the rediscovery and application of Mendelian genetics (Acquaah 2014). From these early beginnings, genetic technologies have become economically feasible and provided logical solutions for many problems. These included molecular markers, gene isolation and transformation, and, more recently, DNA sequencing technology, which has made

genotyping of individual breeding lines feasible in large crop improvement programs (Acquaah 2014). Combining genotyping with incremental improvements in the breeding process, such as off-season nurseries (reduced cycle times), plot equipment (improved phenotypic data collection and accuracy), and statistical analysis techniques (improved interpretation of the data), crop improvement programs are poised to make greater strides (Acquaah 2014).

Application of these molecular genetic technologies was transformational, and new technologies have the potential to be even more transformational, but only if the genetic information can be analyzed in relation to large amounts of informative phenotypic data. Herein lies the current bottleneck in crop genetic improvement programs: Traditional methods of phenotyping simply do not allow the collection of accurate data in quantities to be informative at a transformational level. Plant breeding is often referred to as a "numbers game," meaning that the more material evaluated increases the chances of finding the transformational genetic segregant. Consequently, methods that will allow breeding programs to phenotype larger breeding nurseries faster, more accurately, and repeatedly are essential to making these transformational changes (Furbank and Tester 2011). Furthermore, this phenotyping must be completed in field conditions to accurately identify which genotypes are the most efficient and which will allow crop improvement programs to effectively utilize the genetic technology, and to evaluate crop improvement methods in new and innovative ways.

Several approaches can be used to mitigate this bottleneck. For initial gene and trait discovery, greenhouse and growth chamber studies are useful for dissecting phenotypic traits because they minimize environmental effects. However, for traits of a quantitative nature, field phenotyping is essential to transcending the complex interactions between genetics, environment, and management that routinely occur in production systems. Consequently, field-based, high-throughput techniques, especially those that utilize remote sensing platforms, are the most promising new tool to assist in breaking this bottleneck in the crop improvement pipelines (Tester and Langridge 2010; Furbank and Tester 2011; Araus and Cairns 2014; Shi et al. 2016).

Application of these HTP systems could have several benefits. The first and most obvious is the rapid phenotyping of agronomically important traits. Examples of such traits include, but are not limited to, plant height, maturity, and biotic and abiotic stress tolerance. Currently, these traits are evaluated by individuals who walk through fields and make quantitative or qualitative measurements of individual plant genotypes. If this could be accomplished using HTP and image analysis, greater numbers of genotypes could be evaluated in a breeding cycle. The greater the number of individuals evaluated, the greater the potential to identify the most useful genetic segregants.

The second benefit is less obvious because it was not possible to apply in traditional crop improvement evaluation. Systematic evaluation of a single trait is rarely conducted in a breeding program due to the cost and time to collect the data. With remote sensing, however, this is very conceivable. For example, weekly plant-height measurements on thousands of genotypes was not feasible for most programs, but methods involving photogrammetry or LiDAR remote sensing could be applied. Digital surface model data could be compiled to develop plant-height growth curves and delimit growth curves of different genotypes (Apelt et al. 2015). Plant scientists

have long known these capabilities exist, but have not had a feasible approach to measure them at an appropriate operational scale sufficient for crop improvement (Hoffmann and Rooney 2014).

Several approaches to remote sensing are actively being investigated in crop improvement programs. Ground-based systems are based on a tractor system or other ground vehicle that passes through the field collecting data from sensors attached to the equipment (Pauli et al. 2016). In short crops, ground-based systems (i.e., sensors mounted on farming equipment that passes through the field collecting data) have large sensor capacities and are effective provided that ground conditions do not limit movement through the crop and that repeated trips do not affect crop growth by soil compaction. UASs offer a greater range of coverage but are limited by weather conditions that affect flight and a smaller sensor payload. Ultimately, the capability of these systems to cover vast areas in a short period, as well as their ability to carry various payloads consisting of different sensors, makes them very appealing to crop scientists (Shi et al. 2016).

Once phenotypic data are collected using either system, a significant amount of data preprocessing is required to prepare data for analysis and the production of quantitative information that plant scientists and breeders can use to make decisions regarding the relative value of one genotype relative to another. Geospatial data preprocessing is not trivial because scientific and methodological issues must be taken into consideration (Shi et al. 2016). Currently, the use and application of both the hardware (platform and sensors) and software (data preprocessing, analysis, and visualization) are in their infancy. New approaches and technologies for data collection and processing are continually being developed and need to be rigorously evaluated.

Regardless of the potential of using geospatial technologies, crop improvement programs must first objectively evaluate these numerous aspects of UASs and remote sensing technologies before implementing them into the complexities of agricultural research. From a crop improvement perspective, the value of a measured trait can be assessed in two different ways. First, the measurements can be assessed for accuracy (i.e., the remotely measured trait is consistent with values as they have been traditionally measured). Second, the measurements can be assessed for relative accuracy (i.e., the actual numbers may not be consistent with traditional measurements but they are consistent relative to the other genotypes in the evaluation). In the second approach, the relative performance of any given genotype is consistent in ranking or relative utility. Most programs have baseline performance supplemented with relative performance to make decisions relative to advancement within the program. Consequently, both approaches are important, but, arguably, the latter method is most important in a crop breeding program because decisions must be made on whether individual genotypes are kept or eliminated from further evaluation.

Clearly, the application of remote sensing technologies to plant improvement programs is essential if transformational changes in productivity are to be accomplished. There remains a large amount of research to identify the best approaches for collecting, processing, and analyzing the data that can be generated. The expectation is that these technologies will aid the programs in screening existing traits faster and more accurately than currently possible, and they will elucidate

new traits or means of evaluating existing traits to be more informative to the crop improvement process.

1.2.2 Precision Agriculture

In today's large-scale agricultural operations, farmers use tractors and specific implements (even aircraft in some cases) to apply inputs such as seed, water, fertilizer, and tillage across large fields that have tremendous spatial and temporal variability in soil conditions, crop health, weed and insect pressure, etc. This variability is often significant at a very fine spatial scale, meaning that two adjacent plants could need significantly different amounts of a particular input. Precision agriculture seeks to optimize farm management at the smallest scale possible. To be successful, precision agriculture requires three things: (1) sensors that collect data on position and field characteristics, (2) analytical capabilities that evaluate sensor data and provide decisions or decision support enabling optimization of inputs, and (3) equipment that enables varying the rate of inputs as field location changes. For example, multispectral sensors can be used to map the magnitude of reflectance in certain wavelengths indicative of nitrogen deficiency in a crop, and a spray application of nitrogen fertilizer could potentially apply the appropriate amount of fertilizer where needed. The constraints of precision agriculture optimization are to maximize economic output and minimize environmental risk. Reports of the benefits of precision agriculture have included, for example, major reductions in the amount of pesticide used on a field without significant reduction in crop yield. A recent example related to remote sensing (Yang et al. 2014) involves the mapping of a soil-borne fungal disease in cotton. Because of the tendency of this particular fungus to inhabit a consistent area of a given field, and because the effects of the disease are clearly evident in remote sensing images, mapping the incidence of the disease enables fungicide treatments in subsequent growing seasons to be limited to only the infected areas of the field. If the infected portion of a field is significant, precision agriculture in this case significantly lowers the cost for the farmer and the amount of fungicide applied (along with environmental risks).

Agriculture has benefited greatly from advances in computers and electronics, and many key enabling technologies in precision agriculture are based on these advances. The concept of precision agriculture took hold in the early days of civil access to the U.S. Department of Defense's GPS, which, including the satellite assets of several other countries, is now commonly referred to as the Global Navigation Satellite System (GNSS). Access to the GNSS provides data on field position and time for any recorded field attribute. Coupled with computers and data-recording technologies, GNSS has enabled the mapping of field variability. The development of GISs—computer information systems adept at handling geographic data—has enabled mapping and analysis of field variability. Developments in sensors, both proximal and remote, have enabled mapping of more and more field characteristics. Developments in variable-rate technologies (VRTs) have enabled mechanized input applications to be varied in real time according to field position. Examples of currently available VRTs include variable-rate planters and fertilizer applicators. Advances in analytics like machine learning are enabling sophisticated characterization of

relationships between numerous field properties and key output metrics like crop yield. The "Internet-of-things in agriculture"—including stationary wireless sensor networks and ubiquitous sensors on field equipment coupled with wireless networks and cloud-based data storage and analysis—has provided vast amounts of data that provide the raw material for model development. Advances in automation are enabling real-time control of vehicles and implements in the field. UASs are now being used as remote sensing platforms to provide image data indicative of field variability. Fully autonomous robots are also being developed for the agriculture of tomorrow, and early versions are already finding use in high-value crops for harvesting and thinning fruits and vegetables.

In spite of the onslaught of enabling technologies, current agricultural systems and practices limit precision agriculture variations to management zones, which are typically contiguous multihectare areas within a field that are noticeably different from other areas of the field in some important, manageable characteristic like soil fertility. As technologies continue to improve, site-specific management will focus on increasingly smaller field zones such as 1 ha or even 1 m^2. In the early days of GPS/GNSS, real-time position accuracy was on the order of 20 m, but advances in the technology such as real-time kinematic (RTK) differential correction have made it possible for equipment to determine field position to within 1 cm or so. This level of precision enables automated guidance systems for tractors and harvesters to maintain an extremely precise path. One result is that sensor data can be associated with precise field locations, even to the level of individual plants. The slowest aspect of precision agriculture to take hold is development of analytical capabilities. What is currently used is a combination of algorithms and intuition, but artificial intelligence has the potential to vastly improve precision agriculture analytics.

The numerous sensors now commonly aboard agricultural machinery, in combination with strictly sensing platforms like UASs, have plunged precision agriculture into the new realm of big data. With this change have come several questions that must be answered as data-centric agricultural technologies progress:

- Data ownership: If data are collected by numerous sensors on a farm, who owns the data?
- Data validity: How do we know the data being collected with all these sensors are accurate, have the right contextual information, have been maintained carefully, etc.?
- Data standardization: So many people are collecting and using data with so many types of devices and equipment, how can we make the data interchangeable?
- Data bandwidth: If we are collecting large volumes of data on-farm, how do we transmit the data to a location where they can be analyzed and converted to actionable information, particularly when considering that rural bandwidth is typically very limited?
- Model practicality: Are the analytical models being developed practical in terms of data availability and equipment capability to implement the decisions? For example, if we have a model that identifies on what day, at what location on a 1-m^2 basis, how much irrigation water to apply to a field,

we cannot implement the prescription if we have an irrigation scheme that involves flood irrigation or a center-pivot system.
- Data availability: The volume of data and technological complexity in handling and analyzing it are such that users commonly struggle with major bottlenecks in the workflow of big data for farm-management decision-making.

Big data will have limited value in increasing farm productivity, profitability, and environmental risk reduction until there is a practical way to readily and reliably produce timely, actionable information from the data in common farm-level applications. New precision agriculture technologies like remote sensing, particularly with UASs, have created heretofore unseen capabilities to answer agricultural questions, but a remaining major difficulty, requiring a great deal of domain-specific expertise, is in knowing what questions to ask of the data. As all these capabilities improve, the potential to increase farm profit and environmental stewardship continues to grow. Precision agriculture has already significantly improved productivity and environmental risk mitigation, but the majority of the improvements remain to be made as the industry strives to transition from the scale of large management zones to individual plants.

1.2.3 Geospatial Technologies

An extensive amount of remote sensing literature highlights ongoing geospatial technology research to address challenges associated with agricultural research. Recent advancements in new configurations and specifications of UASs have permitted investigators to test the efficacy of utilizing UAS platforms to test data collection strategies for a multitude of applications (e.g., Hunt et al. 2010; Zhang and Kovacs 2012; Peña et al. 2013; Torres-Sánchez et al. 2013, 2014, 2015; Shi et al. 2016; Yi 2016). A key theme in this research involves data collection to characterize extreme spatial and temporal variability of agricultural field conditions, required for phenotyping and quantitative characterization of plant growth and stress conditions.

Key research themes also involve the engineering aspects of demonstrating the proof of concept that GPS receivers, irradiance sensors, and new commercially available multispectral, hyperspectral, and LiDAR sensors can be mounted effectively to acquire accurate quantitative and image information. Other research addresses the need to configure UASs with Wi-Fi and other communication equipment to address the need of downloading large volumes of high-resolution imagery (e.g., Herwitz et al. 2004) and connecting UASs to high-speed data networks to address the needs for rapid information generation. The importance of this engineering-related research needs to be recognized because flight strategies, sensor payloads, and choice of sensors govern image sampling, data quality, preprocessing requirements, and the nature of image analysis. Classic examples include image sampling and fieldwork control to ensure quality geometric correction and radiometric calibration of imagery. Another example is sensor choice because the spectral response functions of sensors may not be appropriate for extracting accurate vegetation information caused by sampling in regions of the electromagnetic spectrum that record various matter/energy interactions, thereby significantly reducing the scientific validation of using

vegetation indices (i.e., sampling red and near-infrared [NIR] reflectance in one image). Research has clearly revealed that it is not just a matter of acquiring data on a UAS, and that numerous issues must be accounted for related to preprocessing, analysis, and information synthesis (Hardin and Jensen 2011; Kelcey and Lucieer 2012; Shi et al. 2016).

Image preprocessing also represents a very complex topic that needs to be understood. High spatial resolution images must be captured with a high degree of side-lap and overlap so that subscenes can be geometrically corrected and mosaicked. Ground control points (GCPs) are required to ensure a high degree of geometric fidelity. The number of GCPs becomes difficult to manage economically with an increase in the study area size and the lack of control points in subscenes. Our experience and those of others demonstrate that increasing the number of GCPs does not always lead to an increase in geometric accuracy and that field validation is often required (Hruska et al. 2012; Gómez-Candón et al. 2014).

Similarly, radiometric calibration is required to facilitate the use of vegetation indices and biophysical modeling, and enable temporal analysis. Unfortunately, many users are not aware of the need to account for sensor band calibration differences; irradiance variations caused by space-time image sampling, atmospheric factors, and topographic factors; and anisotropic reflectance variations caused by the architectural complexity of vegetation canopies. Rasmussen et al. (2016) addressed this last point by indicating that anisotropic reflectance conditions need to be accounted for in drawing conclusions about using vegetation indices for assessing crop responses to various kinds of treatments. The situation is exacerbated by commercially available sensors that have dynamic exposure control, and depending upon the flight data acquisition strategy, can result in imagery with banding and reflectance anomalies caused by surrounding highly reflective features (e.g., roads and buildings) that trigger changes in the exposure control settings. Dynamic exposure control complicates the radiometric calibration process and requires the imagery to be radiometrically calibrated before geometric correction and mosaicking. In addition, it requires strategic sampling of radiometric calibration panels in the field to facilitate empirical calibration approaches because irradiance varies spatially and temporally and governs radiance measurements obtained in the field by spectral radiometers. Calibration tarps can also be used, although there are limitations related to accounting for solar geometry and viewing angles (Moran et al. 2001).

Given the tremendous volume of multispectral and hyperspectral data that can be acquired over the growing season, geospatial research efforts have also focused on evaluating various information extraction approaches to assess, monitor, and map weeds, crops, and plant stress conditions. A strong research emphasis has been on evaluating the use of spectral bands and vegetation indices for the detection and discrimination of weeds and crops (e.g., Armstrong et al. 2007; Gray et al. 2008; López-Granados 2010; Eddy et al. 2013; Peña et al. 2013). Similarly, research involving hyperspectral remote sensing of plant pigments and stress (e.g., Haboudane et al. 2004; Blackburn 2006, 2007; Feret et al. 2008; Jacquemoud et al. 2009; Sanches et al. 2014) is critical for precision agriculture applications. Investigators are attempting to evaluate new data sets for spectral discrimination capabilities via more advanced processing techniques including absorption feature modeling (Jacquemoud and Baret

1990; Jacquemoud et al. 2009), wavelet analysis (Koger et al. 2003), object-oriented analysis (Ruiz et al. 2011; Torres-Sánchez et al. 2015), and various classification approaches (Armstrong et al. 2007; Gray et al. 2008; Siachalou et al. 2015; Haug and Ostermann 2016; Hernández-Hernández et al. 2016). Standardized and highly accurate quantitative and thematic information about plants and crop cover conditions remains elusive due to plant spectra similarity resulting in low classification accuracies and the complexities of interacting plant and environmental factors that limit causal diagnostic assessment of plant stress conditions.

Finally, there is a paucity of research on geospatial information synthesis to support agricultural optimization efforts and provide acceptable and reliable decision support. Although it has been recognized that various forms of geospatial data and information can be integrated into crop growth models that provide decision support (e.g., Thorp et al. 2008), advances in semantic modeling, spatial analysis, knowledge representation, artificial intelligence, and scientific visualization have not been effectively incorporated into prototype decision support systems. Such systems will inevitably incorporate UAS remote sensing technologies, geospatial information processing systems, field sensor and network systems, crop modeling systems, and visualization and communication functionality for stakeholders. In the meantime, UAS and remote sensing technologies need to be evaluated with respect to a multitude of issues that promote and/or limit diagnostic assessment of agricultural environments.

1.3 STUDY AREA

We discuss various issues resulting from our research at the Texas A&M University System Research Farm, located in Burleson County, Texas (Figure 1.1). The eastern border of the farm is the Brazos River, which separates the farm from Brazos County, Texas. Intermittent flooding of the Brazos over thousands of years has led to the high soil fertility that enabled row crop production at this site, despite large portions of the rest of both surrounding counties only being suitable for pasture land. The majority of the Brazos Bottom Research Farm exhibits a Ships clay loam soil, fertile with adequate water-holding capacity for typical row crop growth, when sufficient rainfall and some supplemental irrigation water is available. To the east of the river is located a small regional airport (Easterwood). At their closest, Easterwood's runways are less than 5.90 km from the farm. All research plots, however, are greater than 7 km away from the airport. Texas A&M received authorization to fly UASs for research on agriculture and was one of the first research universities to do so.

The primary research and production crops include corn (*Zea mays* L.), cotton (*Gossypium hirsutum* L.), sorghum (*Sorghum bicolor* (L.) Moench), wheat (*Triticum aestivum* L.), cowpeas (*Vigna unguiculata* (L.) Walp.), and perennial grasses (various species), although other crops are also grown. Plot research that occurs on the farm consists of plant breeding and genetics experiments (consisting of testing many varieties and very small plots), agronomy, entomology, plant pathology, weed science (consisting of trials evaluating pest control options), and soil science experiments (typically measuring soil characteristics and their spatial effect on crop varieties).

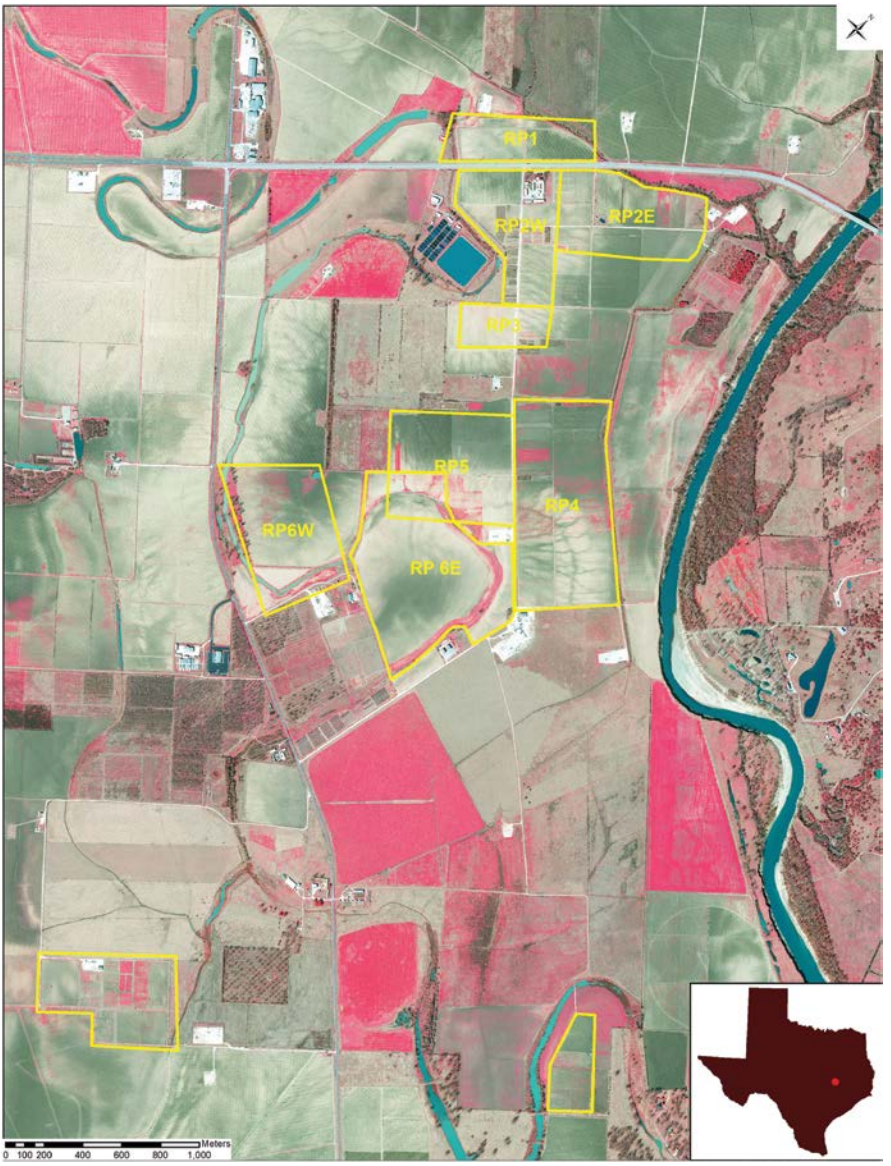


FIGURE 1.1 False-color composite (green, red, NIR) orthoimagery over the study area in Burlseon County, Texas. The Brazos River is on the right of the image, and the FM 60 highway runs left to right near the top of the image and crosses the Brazos River. Data collection was attempted in one UAS flight in route packs (RPs), outlined in yellow. The 6-in and 12-in 4-band (blue, green, red, NIR) orthoimageries were acquired between January and February 2013 during leaf-off conditions. The data set covers approximately 1963 km^2 (758 mi^2). Five local entities contributed funds to collect the new imagery using the High Priority Imagery and Data Sets (HPIDS) State contract administered by the Texas Natural Resource Information System (TNRIS) Strategic Mapping Program (StratMap).

1.4 IMAGE ACQUISITION

The ability to conduct useful science under the framework of agricultural applications is not only dependent upon the collection of high-quality multi- and hyperspectral imagery, but also dependent upon platform and sensor issues in complex interdependent ways. This requires the proper matching and integration of air vehicle, sensors, mission design, and image preprocessing. Although commercial UASs are starting to be equipped with autopilots, sensors, and simple data processing software, they are often limited to only one sensor and frequently lack cross-platform integration expandability. Our work has focused on addressing issues using highly integrated fixed-wing and a multirotor unmanned aerial system that is customized for precision agriculture science. Numerous issues associated with flight operations and data capture need to be addressed.

1.4.1 Flight Operations and Sensors*

It is essential that interdisciplinary knowledge and trained pilots be used for image acquisition. This requires the coordination of research scientists to account for all the phases of research including data acquisition, image preprocessing, information extraction, fieldwork, image visualization, and decision support. Our multidisciplinary team includes the Colleges of Agriculture, Geosciences, and Engineering, with expertise in agricultural crop research, remote sensing and geospatial technology, and UAS flight operations. Collaboration among the experts in these fields ensures collection of high-quality imagery to meet research objectives.

Operation of UASs on Texas A&M University System property requires compliance with Texas A&M System Regulation 24.01.07 and the small UAS regulation (Part 107) issued by the Federal Aviation Administration (FAA). To operate the controls of a small UAS under Part 107, the operator needs a remote pilot airman certificate with a small UAS rating, or be under the direct supervision of a person who holds such a certificate. All UAS operations were conducted by researchers who have their FAA remote pilot certification, and all UAS flight operations were conducted under Part 107 requirements. These requirements include registration of the unmanned aerial vehicles (UAVs), operating in Class G airspace under daylight-only visual flight rules (VFR) conditions, maintaining visual line of sight of the UAV, and operating at a maximum altitude of 400 ft (∼122 m) above ground level (AGL).

Operationally, we have used two types of UAVs, fixed-wing and multirotor, to collect imagery of crops depending on the type of imagery needed to assess the condition of the crop field. A fixed-wing UAV uses a conventional wing design to generate lift through its forward motion. It generally flies at faster speeds (12–18 m/s) and higher altitudes (80–100 m), which provides more efficient coverage of larger areas. A fixed-wing UAV also has a longer endurance time, which is an essential capability for production-scale farms. Its design does require an open area for launching and landing

* Listing of commercial aerial platforms and sensors does not represent an endorsement from individual researchers or Texas A&M University, but are provided as examples of commercial equipment being evaluated for research purposes.

the UAV. The fixed-wing UAVs have robust design and durable construction, allowing for long hours of repeated use throughout a growing season.

A multirotor UAV uses individual rotors, or propellers, to generate lift for the vehicle. The number of rotors typically varies from three to eight, depending on the size of the UAV. It typically flies at slower speeds (5–7 m/s) and lower altitudes (20–80 m), which provides high-resolution images of small areas. The multirotor UAV usually has short flight times and limited payload capacity. With its vertical takeoff and landing capability, it can be operated in confined areas near obstructions, such as power lines, towers, or buildings. The autonomous flight software used in multirotor UAVs provides ease of use for the operator, while also providing fail-safe features to ensure that the operator does not violate regulations for flying in the National Airspace System.

We have used a variety of platforms and sensors including:

- PrecisionHawk Lancaster 5 with a Nikon 1 J4 visual camera: The PrecisionHawk Lancaster 5 (Figure 1.2a) is a fully autonomous UAS that automatically optimizes its flight plan to efficiently collect data. Equipped with plug-and-play sensors that can be exchanged in the field without configuration changes, the UAS can capture a variety of imagery.
- Tuffwing UAV Mapper with a Sony A6000 visual camera: The Tuffwing UAV Mapper (Figure 1.2b) is suitable for aerial mapping. The airframe was designed for the Pixhawk autopilot and can carry cameras including Sony mirrorless cameras, Canon S110, Sequoia, RedEdge Flir Vue, and SLANTRANGE. The UAS can stay in the air for about 45 minutes on a single charge, which will easily photograph 1 km^2 while flying at an altitude of 100 m. The flying wing design flies better in strong winds than a traditional airplane.
- DJI Matrice 100 UAV with MicaSense RedEdge multispectral camera: The DJI Matrice 100 (Figure 1.2c) is an easy-to-fly multirotor UAV. It is equipped with a visual camera and can be configured with additional sensors, such as a multispectral or thermal camera. We have evaluated the MicaSense RedEdge sensor.

FIGURE 1.2 Select UASs evaluated in this study: (a) PrecisionHawk Lancaster 5 with Nikon 1 J4 visual camera; (b) Tuffwing UAV Mapper with Sony A6000 visual camera; and (c) DJI Matrice 100 with MicaSense RedEdge multispectral camera.

The traditional paradigm for most science applications of UASs has been that good science is a function of using good UASs and image sensor technologies. This paradigm is not completely correct because there can be significant gaps between commercial off-the-shelf (COTS) UASs and research-purposed unmanned vehicles that can carry and properly fly high-performance sensors. For customizable sensor integration, commercial UASs are starting to be equipped with autopilots, sensors, and simple data processing software, but can be limited to carrying only one sensor. Additionally, there is often a lack of cross-platform integration expandability. This gap negatively impacts the science objectives. Many factors related to UAS and sensor technologies must be considered to obtain quality imagery, and not appropriately addressing them has shown to be problematic.

1.4.2 Data Acquisition Flight Strategies

Flight planning is an inherent process required for obtaining imagery (Valasek et al. 2016, 2017). Seasonal prevailing wind direction and wind intensity, sensor-specific distance between passes, image overlap, image side-lap, GCP locations, and auto-triggering cycle time are all crucial for quality data collection. The ultimate goal of flight path design is to ensure geographic coverage and improve mosaic quality under a restricted time frame. Flight time is restricted due to battery restrictions and the inherit nature of large fluctuations in outgoing thermal radiance. Image quality can be increased by reducing flight speed to reduce motion blur, reducing cross-wind paths, and increasing image overlap and side-lap.

1.4.2.1 Prevailing Wind

Prevailing wind intensity is crucial for both the reduction of cross-wind flight paths and the necessary speeds for takeoff and landing. Cross wind during flight reduces the accuracy of the waypoint tracking and increases the vehicle bank angle, and, as a result, increases the complexity of image mosaicking (Hardin and Jensen 2011). Due to the nature of UAS image acquisition flight environments, the vehicle usually carries payloads close to the maximum takeoff capacity and can have limited takeoff and landing space. Therefore, by carefully utilizing the seasonal prevailing wind, both the percentage of cross-wind flight paths and takeoff and landing distance can be minimized.

1.4.2.2 Waypoint Scheduling

Common issues with the standard and cross-stitch waypoint scheduling methods (Figure 1.2a and b) are that the high banking angles required for agile turns are not easily achieved with fixed-wing UASs. Additionally, agile maneuvers for fixed-wing UASs often result in missed GCPs as well as images with high banking angles, which leads to image distortion and poor mosaic results. A moving-box waypoint design method was developed from experience with collecting imagery from multiple flight tests, and provides the advantages of generally reducing the commanded banking angle and does not require large control surface maneuvers. With the moving-box method, waypoints are scheduled in a rectangular pattern that migrates along one direction, such that data are collected along the perimeter of the moving rectangle until coverage of the target area is achieved (Figure 1.2c). The pattern can be

scheduled either clockwise or counterclockwise according to the wind direction and is dependent on specific mission conditions and parameters. Issues with the moving-box method lay in the imagery coverage of the center region.

Flight planning research activities indicate that, in order to create high-quality mosaics with minimal distortion, it is recommended to have at least 75% overlap between successive images and 60% side-lap between flying tracks. This is sufficient for flights with low wind measurements. However, in flight days with wind gusts above 6.7 m/s (15 mph), when the vehicle turns the cross winds induce the vehicle to bank above the banking threshold required for good imaging.

1.4.2.3 Waypoint Auto-Triggering

Several COTS autopilot software programs provide auto-triggering capability. Mission Planner provides triggering according to distance or triggering according to a fixed time interval. For fixed-wing imaging flights, the first approach often results in a skewed distribution of images throughout a flight path. This is the result of either the delay time in the autopilot system adjusting to wind gusts and therefore miscalculating the distance traveled, or reduction of speed when the vehicle is banking. Consequently, the fixed time interval auto-triggering (Figure 1.3c) is preferred.

1.5 GEOMETRIC CORRECTION: UAS PHOTOGRAMMETRY

UAS photogrammetric practices begin with direct georeferencing, which can be defined as "the process of independently reconstructing the geometry of remotely sensed data

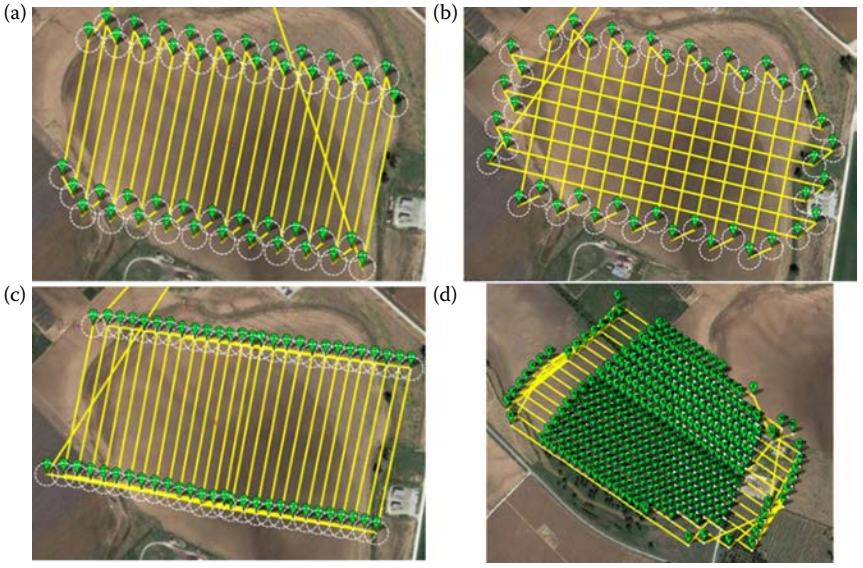


FIGURE 1.3 Flight path patterns and planning. (From Shi et al. 2016. *PLoS ONE*, *11*(7), e0159781. doi:10.1371/journal.pone.0159781.): (a) conventional flight path planning; (b) cross-stitch flight path planning; (c) moving-box flight path planning; and (d) distribution of data collection locations with fixed time interval auto-triggering.

by integration of remote sensing and navigation systems" (Perry 2009). This is generally accomplished with a GPS and an inertial measurement unit (IMU) onboard the UAS that records the position (latitude, longitude, and elevation from the GPS) and attitude (roll, pitch, and yaw from the IMU) of the UAS at the time of each image acquisition. The photogrammetric solution is usually supplemented with surveyed GCPs.

Few UAS GPS units reach the RTK positional error of $\leq$2 cm. Therefore, a network of GCPs is required to photogrammetrically generate orthoimages and digital elevation models (DEMs) with which reasonable measurements can be made. How many GCPs and where to locate them are questions currently answered with more art than science. For example, Benassi et al. (2017) used 12 GCPs for a field of 0.2 km^2, while Gillan et al. (2017) used three for plots of 0.0025 km^2, and Glendell et al. (2017) used 8–36 GCPs for fields of 0.02–0.18 km^2. The fields in this work ranged from 0.1 to 0.5 km^2 and usually included three to five GCPs. Environmental factors can sometimes determine the usability of GCPs because they can be obscured due to floods, soil deposition, weeds, and ant mounds. Farm machinery, other off-road machinery, and wild animals can damage or destroy GCPs. As a result, a desired horizontal root-mean-squared error (RMSE) of less than 1 pixel can usually be achieved only with considerable manual intervention.

Environmental conditions, especially wind and cloud cover, affect accuracy. Benassi et al. (2017) made four flights under three regimes (12 GCPs, RTK only, and RTK with one GCP) over the same area and found an inverse correlation between wind speed and accuracy. Interestingly, RTK with one GCP was the most accurate, followed by RTK only, except for the flight with the highest winds (in that flight, the 12 GCP solution was best). This may be due to the effects of excessive roll and pitch, which can cause considerable errors in the generation of orthomosaics (Figure 1.4). There is likely some threshold value for roll and pitch below which errors are unobservable and beyond which errors are considerable and ruinous.

Wind-induced pitch and roll can be reduced through the use of gimballed mounts. However, these add considerable weight to the UAS and engineering or operational considerations may preclude their use. When such is the case, increased side- and overlap increases the likelihood that sufficient near-nadir images are acquired. Grenzdörffer et al. (2008) found that a nominal side- and overlap of 60% could be insufficient, and found success with 80% for both. With higher overlap, images with excessive roll can be eliminated from the photogrammetric solution, and better orthomosaics and DEMs can be produced (Figure 1.5).

1.6 CROP ASSESSMENT AND MAPPING

An extensive amount of literature has focused on the use of vegetation indices for mapping and biophysical assessment of crops (e.g., Thenkabail et al. 2000; Haboudane et al. 2002; Zarco-Tejada et al. 2003, 2004; Eddy et al. 2013; Peña et al. 2013; Candiago et al. 2015; Kross et al. 2015). Mapping studies have focused on evaluating spectral and spatial features for discriminating between various crops and land-cover classes. A focus of such thematic mapping has been on the evaluation of new imagery and vegetation indices to increase the statistical separability of classes

FIGURE 1.4 Orthoimage with overall subpixel RMS error, but with duplication and mislocation errors of ~25 pixels due to inclusion of images with excessive aircraft roll. Note ghostly white and dark gray concrete tiles (GCPs). True locations given by black and white squares.

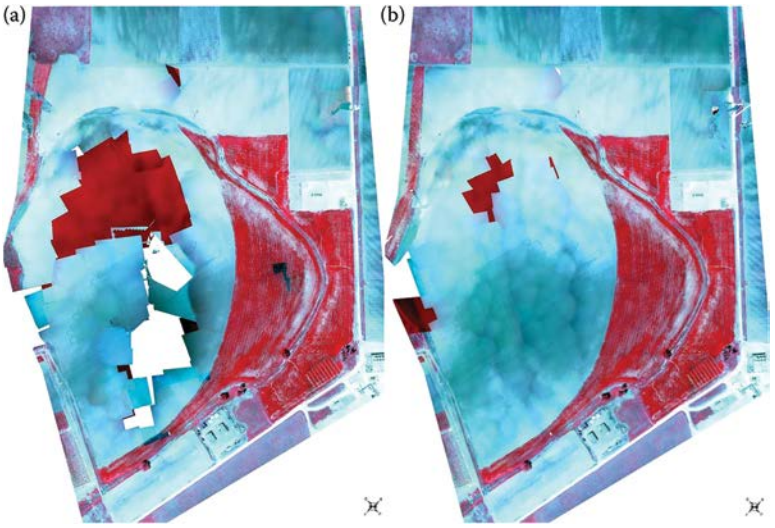


FIGURE 1.5 Orthoimages of RP 6E (see Figure 1.1) acquired on a windy day. Both have a reported RMS of around 1 pixel, but (b) had images with a pitch or roll greater than half the camera's field of view removed from the solution.

in feature space to permit accurate classifications. A significant amount of research has also focused on evaluating various types of classification algorithms, including statistical and artificial neural computing classification approaches (Moreno et al. 2014; Peña et al. 2014; Kumar et al. 2015; Haug and Ostermann 2016). Given the advent of new multi- and hyperspectral sensors, such empirical thematic mapping research will continue.

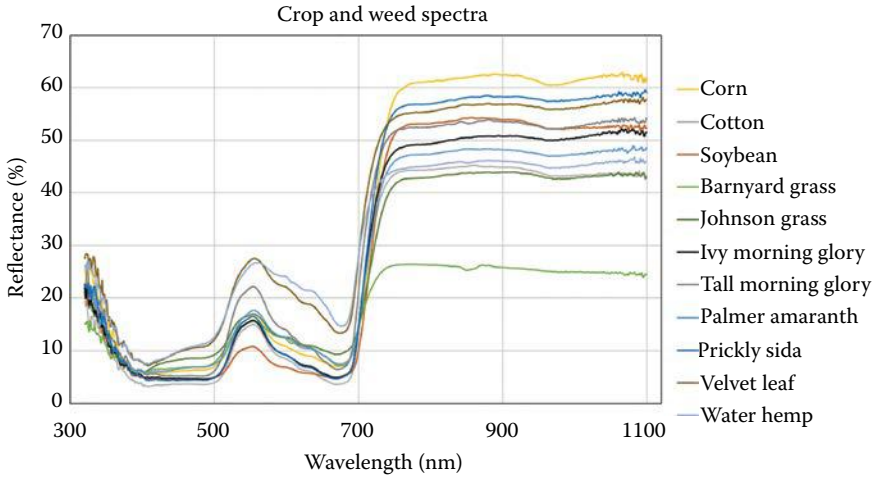


FIGURE 1.6 Select crop and weed spectra sampled in the study area with a Spectral Evolution PSR-1100 field portable spectroradiometer with a leaf clip attachment. Data were collected in the field area (Figure 1.1) July through September 2016 except the corn spectra, which were collected March through April 2017.

It is important to note, however, that crop and weed spectral reflectance patterns can be very similar (Figure 1.6), and a variety of environmental factors including topography, soil macro- and micronutrient status, and other abiotic and biotic factors can influence plant pigment concentrations, mesophyll cellular structure, leaf moisture, and other biochemical properties. Although it is widely recognized that hyperspectral data can be used to assess some of these conditions (Haboudane et al. 2004; Blackburn 2006, 2007; Feret et al. 2008; Sanches et al. 2014), empirical differentiation and mapping of crops and weeds can be extremely difficult to accomplish in a diagnostic fashion (Rasmussen et al. 2016). Consequently, we have been focusing on assessment of other analytical approaches that have diagnostic potential such as hyperspectral and spatial wavelet analysis.

Wavelet analysis characterizes the scale-dependent periodicity of a signal, which is superior to Fourier analysis because it both reveals the frequency components of the signal and identifies where a certain frequency exists in the spectrum. Wavelet analysis can be a very useful tool to study UAS hyperspectral imagery because it is capable of revealing trends, breakdown points, higher derivative discontinuities, and self-similarity in the reflectance pattern (Misiti et al. 1996).

Researchers have reported that unique spectral features can be extracted from wavelet coefficients, and that these features are useful in crop discrimination (Hsu and Tseng 2000), weed detection (Koger et al. 2003; Bossu et al. 2009), and assessment of crop nutrient stress levels (Liu et al. 2011). Wavelet coefficients can be calculated using the continuous wavelet transform (CWT) method or the discrete wavelet transform (DWT) method. The CWT is superior because it operates at every

scale and the shifting of the wavelet function is continuous and smooth. The CWT is defined as:

$$WT(a,\tau) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(x)\psi^*\left(\frac{x-\tau}{a}\right)dx$$

where a is scale, τ is position, and ψ^* is the mother wavelet function. A wavelet coefficient is a function of both scale and position. Scale controls the compression or stretching of the wavelet and position controls the shifting of the wavelet function.

For a one-dimensional analysis of a hyperspectral curve, the CWT generates wavelet coefficients in wavelength-scale space and the magnitude of the coefficients indicates how closely the data correlate with the mother wavelet for a specific spectral band location. There is a correspondence between wavelet scales and frequency, such that a smaller scale corresponds to a compressed wavelet, which is high in frequency, while larger scales correspond to a stretched wavelet, representing lower frequency. Scale can be converted to frequency if needed.

Figure 1.7 shows the continuous wavelet transform on soybean (a crop) and water hemp (a weed) spectra. The coefficient plot provides a wavelength-scale view of the spectrum. After the transform, modulus maxima lines are extracted from the CWT coefficients at each scale, which highlight spectral features such as the green peak, red edge, and near-infrared absorptions. These features and potentially other features that we are developing have the potential to be used for diagnostic species classification. Nevertheless, more research is required to determine the degree to which numerous genetic, environmental, and managerial factors influence wavelet coefficient spectral features.

Another important approach for mapping and assessment of plant communities involves spatial analysis of plant architecture and canopy structure. UAS-based LiDAR systems can be used to assess plant architectural characteristics (e.g., Chisholm et al. 2013; Anthony et al. 2014; Grenzdörffer 2014), while image-based spatial features can provide valuable information regarding canopy texture and anisotropy.

We demonstrate this by characterizing the one-dimensional variation in reflectance for crop canopies for four different species. Specifically, we selected four transects, each over a soybean, cotton, sorghum, or corn field. Data from each transect were analyzed via wavelet analysis to characterize spatial patterns in reflectance (Figure 1.8). Different spatial scales (i.e., spatial frequency) relate to variations in canopy structure. Wavelet analysis reveals a unique spatial frequency pattern for soybeans and cotton compared to sorghum and corn, because these two crop groups have vastly different canopy structures (Figure 1.8). Because canopy structure changes over time in response to plant maturity, plant health, presence of weeds, and other environmental factors, a multitemporal wavelet analysis approach may have significant application related to biomass or crop yield estimation and

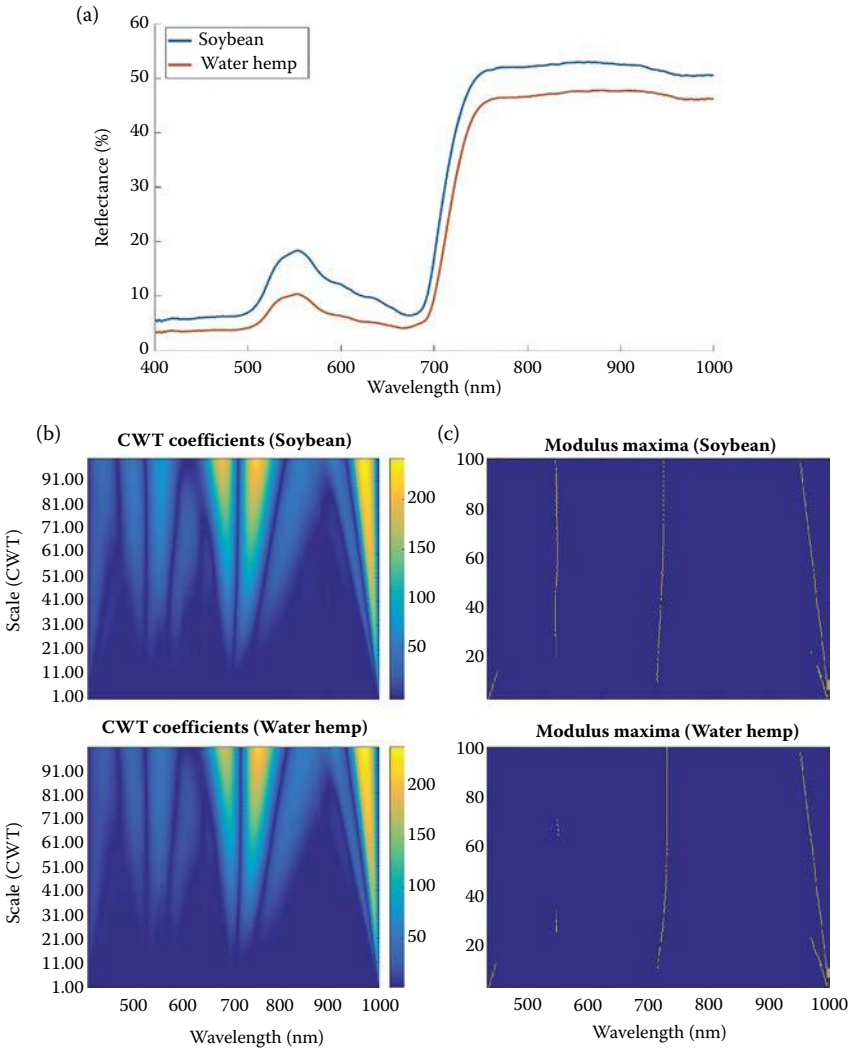


FIGURE 1.7 Continuous wavelet analysis on soybean and water hemp. (a) Reflectance spectra of soybean and water hemp. (b) CWT coefficients, using Daubechies 3 as the mother wavelet. Color represents the magnitude of wavelet coefficients, where yellow is high magnitude. (c) Extracted modulus maxima lines.

weed control. High-resolution imagery is required to better characterize plant and canopy structural patterns.

1.7 WEED ASSESSMENT

Weed competition is one of the major limiting factors to crop production (Zimdahl 1999). Weeds compete with crops for critical resources such as water, nutrients, light,

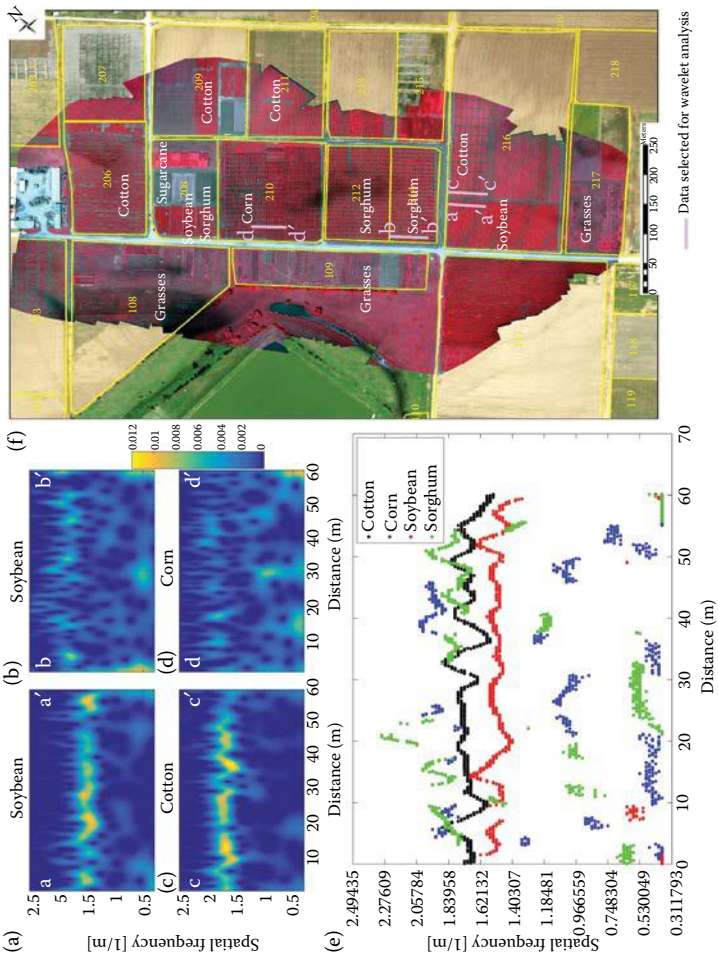


FIGURE 1.8 Wavelet analysis for characterizing canopy structure with near-infrared reflectance spatial variations for soybean, cotton, sorghum, and corn. (a–d) Wavelet coefficient plots of spatial frequency (*Y*-axis, from CWT analysis, converted from scale) of the reflectance value at a given location at a distance along the transect (*X*-axis). Color represents the magnitude of the wavelet coefficients in frequency-location space, where yellow is high magnitude. (e) Maxima along each transect, extracted from (a–d) for comparison. (f) Annotated map indicating transect locations (pink lines) for each crop species.

and space, resulting in crop yield loss. Herbicides are relied upon as an important tool for weed management in broad-acre cropping systems. However, improper use of herbicides typically leads to reduced weed management efficacy and lost profitability. To be effective, weed management decisions must be taken based on a thorough assessment (i.e., field scouting) of the dominant weed species present, area of infestation, and growth stage and size of crops and weeds (Wiles et al. 1992). Field scouting for weeds serves as an important component of integrated weed management and is conducted at routine intervals to identify problem weed issues and make suitable management decisions.

Currently, growers usually employ field scouts and crop consultants to perform weed infestation assessments. However, manual field scouting is expensive to growers, estimated to cost approximately $10/acre (Prostko 2015), and is often inaccurate due to various difficulties and challenges associated with ground-based field scouting. Because weed distributions are often patchy, the weed infestation is better viewed above the crop canopy. When crop stands are dense, field scouts mainly walk around the field edges, thus bypassing weed issues in field centers. Deep standing water makes walking inside a field extremely challenging in crops such as rice.

Weed identification and management is highly time sensitive because effective weed control is achieved only when weeds are treated prior to a certain growth stage (Carey and Kells 1995; Johnson and Hoverstad 2002). However, unfavorable weather conditions can severely delay field scouting for several days or weeks. Furthermore, field scouting adds a tremendous workload in an already busy season (Czapar et al. 1997), meaning that not all fields receive timely scouting. An inability to take timely action can lead to irreversible crop damage and increased weed pressure in subsequent years. Visual weed infestation estimates are also subject to variability and human error. There is no standardized visual assessment protocol that applies across weed species and is consistent among individuals. In this regard, UAS-based tools can be invaluable in diagnosing and assessing weed infestations and facilitating precision weed management. Other related applications of UAS technology include the assessment of weed and crop growth stages and estimation of weed density and spatial distribution, as well as the assessment of herbicide injury and off-target herbicide drift damage on crop fields.

Historical use of remote sensing technology for weed assessment and other applications uses sensors installed on satellite or other aerial devices. These methods have been useful, but not robust due to inadequate spatial resolution. Sensors have also been mounted on ground vehicles for simplified weed detection and herbicide delivery. However, ground vehicles typically lack the ability to utilize field-level spatial information in management decision-making and have limitations under wet soil conditions. Compared to ground-vehicle-based platforms, UASs that fly at relatively low altitudes provide a convenient means to acquire high-resolution imagery of field-level weed infestations.

Earlier research on sensors that could collect data for weed assessment were primarily based on surface reflectance (Haggar et al. 1983). For example, the optical sensors such as GreenSeeker and WeedSeeker technologies have utilized surface reflectance of leaves to differentiate weeds and bare soil (Andújar et al. 2011). Herbicide applications carried out using the WeedSeeker technology have resulted

in significant savings in herbicides (Hummel and Stoller 2002). Haggar et al. (1983) developed a plant detection system based on the radiance ratio of the red and near-infrared wavelengths, which is typically greater for green vegetation compared to soil. Researchers have also experimented with simple machine vision systems. Fennimore et al. (2016) developed a machine-vision-based system that can remove intrarow weeds as well as thin crop stands to desired population densities, though this system had a limited ability to distinguish weeds at large growth stages and at high densities.

Robust weed detection and differentiation will require the utilization of advanced sensors that use reflectance, transmittance, or fluorescence images under ultraviolet, visible, or near-infrared illumination. Other sensors, such as LiDAR and thermal sensors, may also provide valuable information. Utilization of recent advances in sensor technology can be instrumental in this regard (Shi et al. 2016). Meyer et al. (1998) have shown that leaf shape and textural analysis can be helpful for distinguishing weed species for spot spraying. Additionally, spatial structural pattern analysis of production fields and knowledge integration can be incorporated into the existing approaches. Given the complex characteristics exhibited by weeds, we argue that a combination of tools and approaches that include surface reflectance, plant height, arrangement of individuals and distribution patterns, canopy temperature, plant or leaf shape and structural features, spatial field patterns, and knowledge integration will be invaluable in achieving more reliable outcomes.

Research is, however, limited in utilizing UASs for weed assessments and other related applications in large-scale agricultural production fields. Some of the existing limitations include the lack of advancements in the application of spatial and textural analysis of large-scale field images to extract useful diagnostic information and develop decision support tools to guide management decision-making.

1.8 CONCLUSIONS

Given the advent of new sensor imaging technologies and advances in UAS platforms, agricultural researchers have new opportunities to collect high spatial and temporal resolution imagery that can revolutionize the way in which farmers, planners, and managers optimize crop production and management strategies. Similarly, plant breeders can benefit from rapid assessment and characterization of plant biophysical properties, ranging from above-ground and below-ground biomass assessment using LiDAR and ground-penetrating radar sensors, plant physiological stress conditions, and plant and canopy architectural and structural conditions. Nevertheless, effective use and evaluation of new sensors, data acquisition strategies, imagery preprocessing requirements, information extraction algorithms and approaches, and the synthesis of information and knowledge for use in decision support remain important research topics. Remote sensing and information technologies have yet to enable diagnostic plant characterization solutions and reliable decision support for optimal agricultural management strategies. This will require additional engineering, remote sensing, algorithm, and information system development to address a multitude of issues related to soil and crop sciences.

New research directions that need to be pursued and require new innovative technology solutions include:

- UAS payload optimization, including multiple sensors, such as GPS, pyranometer, LiDAR, and hyperspectral sensors, and the inclusion of onboard processing and CPUs/GPUs to permit near-real-time image preprocessing and environmental awareness.
- Intelligent UAS navigation capabilities that ensure optimized image acquisition based on environmentally aware platforms that can rapidly adjust flight planning strategies based on changing conditions. Navigation controllers also need to be developed to ensure platform and sensor communication with other UAS platforms that enable swarm technologies to be utilized effectively.
- Data communication and transfer capabilities to address data storage, environmental awareness, and efficient turnaround time for stakeholders. New antenna technologies and high transfer rates will be needed to support navigation, storage, computation, and decision support functionality, given all the technical issues associated with geospatial big data. Effective data management solutions are critical because data, intermediate processing steps, information products, and modeling results must all be organized, accessible, and readily available.
- Effectively addressing radiometric calibration and geometric correction and mosaicking of imagery over large areas. This is required because traditional approaches do not effectively account for the issues of space (large geographic areas), time intervals, topography, and atmospheric conditions related to surface irradiance variations and optimized placement of calibration panels and GCPs. In the future, all UASs should include irradiance and GPS sensors to address the complexity of radiometric and geometric calibration and correction.
- Investigation of diagnostic information extraction approaches. Active remote sensing research is predominately based on the evaluation of empirical approaches to agricultural thematic mapping and biophysical assessment, focused on the evaluation of new sensors and image information content. Diagnostic information extraction algorithms and approaches are sorely needed to produce accurate and reliable information.
- Development and evaluation of artificial intelligence technology to enable decision support capabilities. This includes addressing semantic modeling, knowledge representation, and information synthesis, and enabling conceptual knowledge of concepts and causation to be used to address mathematically intractable agricultural problems. Similarly, geospatial decision support systems will need to be integrated with plant and crop production models that account for landscape intakes (i.e., energy, water, nutrients) and outputs (i.e., energy, water, biomass).
- Finally, the nature of new technological development and geospatial big data issues dictates that we require new and improved scientific visualization capabilities that support data exploration, computation, information synthesis, knowledge discovery, and decision support. People need to more effectively interact with software programs and systems based on a collective view and understanding of environmental

conditions. Modern commercial-based software systems do not provide for these capabilities. This is not a trivial task because basic image display is totally inadequate for information synthesis and many other issues that we have addressed.

ACKNOWLEDGMENTS

The authors would like to thank Texas A&M AgriLife Research for funding to initiate UAS research activities and an AgriLife Cropping Systems research grant to Bagavathiannan to study the application of UASs for weed and crop stress assessment. We thank everyone in AgriLife Corporate Relations for their contributions to this large multidisciplinary research project at Texas A&M University.

REFERENCES

Acquaah, G. 2014. *Principles of Plant Genetics and Breeding*, 2nd Edition. Wiley and Sons, New York.

Andújar, D., Ribeiro, Á., Fernández-Quintanilla, C., and Dorado, J. 2011. Accuracy and feasibility of optoelectronic sensors for weed mapping in wide row crops. *Sensors*, 11(3), 2304–2318. doi:10.3390/s110302304.

Anthony, D., Elbaum, S., Lorenz, A., and Detweiler, C. 2014. On crop height estimation with UAVs. *Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference*, 4805–4812. IEEE.

Apelt, F., Breuer, D., Nikoloski, Z., Stitt, M., and Kragler, F. 2015. Phytotyping^{4D}: A light-field imaging system for non-invasive and accurate monitoring of spatio-temporal plant growth. *The Plant Journal*, 82(4), 693–706.

Araus, J. L., and Cairns, J. E. 2014. Field high-throughput phenotyping: The new crop breeding frontier. *Trends in Plant Science*, 19(1), 52–61.

Armstrong, J.-J. Q., Dirks, R. D., and Gibson, K. D. 2007. The use of early season multispectral images for weed detection in corn. *Weed Technology*, 21, 857–862. doi:10.1614/WT-06-074.1.

Benassi, F., Dall'Asta, E., Diotri, F., Forlani, G., Morra di Cella, U., Roncella, R., and Santise, M. 2017. Testing accuracy and repeatability of UAV blocks oriented with GNSS-supported aerial triangulation. *Remote Sensing*, 9(2), 172. doi:10.3390/rs9020172.

Blackburn, G. A. 2006. Hyperspectral remote sensing of plant pigments. *Journal of Experimental Botany*, 58(4), 855–867. doi:10.1093/jxb/erl123.

Blackburn, G. A. 2007. Wavelet decomposition of hyperspectral data: A novel approach quantifying pigment concentrations in vegetation. *International Journal of Remote Sensing*, 12(28), 2831–2855. doi:10.1080/01431160600928625.

Bossu, J., Gée, C., Jones, G., and Truchetet, F. 2009. Wavelet transform to discriminate between crop and weed in perspective agronomic images. *Computers and Electronics in Agriculture*, 65(1), 133–143. doi:10.1016/j.compag.2008.08.004.

Brummer, E. C., Barber, W. T., Collier, S. M., Cox, T. S., Johnson, R., Murray, S. C., Olsen, R. T., Pratt, R. C., and Thro, A. M. 2011. Plant breeding for harmony between agriculture and the environment. *Frontiers in Ecology and the Environment*, 9(10), 561–568. doi:10.1890/100225.

Cairns, J. E., Hellin, J., Sonder, K., Araus, J. L., MacRobert, J. F., Thierfelder, C., and Prasanna, B. M. 2013. Adapting maize production to climate change in sub-Saharan Africa. *Food Security*, 5(3), 345–360.

Candiago, S., Remondino, F., De Giglio, M., Dubbini, M., and Gattelli, M. 2015. Evaluating multispectral images and vegetation indices for precision farming applications from UAV images. *Remote Sensing*, 7(4), 4026–4047. doi:10.3390/rs70404026.

Carey, J. B., and Kells, J. J. 1995. Timing of total postemergence herbicide applications to maximize weed control and corn (*Zea mays*) yield. *Weed Technology*, 9(2), 356–361.

Castaldi, F., Pelosi, F., Pascucci, S., and Casa, R. 2017. Assessing the potential of images from unmanned aerial vehicles (UAV) to support herbicide patch spraying in maize. *Precision Agriculture*, 18, 76–94. doi:10.1007/s11119-016-9468-3.

Challinor, A. J., Watson, J., Lobell, D. B., Howden, S. M., Smith, D. R., and Chhetri, N. 2014. A meta-analysis of crop yield under climate change and adaptation. *Nature Climate Change*, 4, 287–291. doi:10.1038/nclimate2153.

Chisholm, R. A., Cui, J., Lum, S. K., and Chen, B. M. 2013. UAV LiDAR for below-canopy forest surveys. *Journal of Unmanned Vehicle Systems*, 1(1), 61–68. doi:10.1139/juvs-2013-0017.

Czapar, G. F., Curry, M. P., and Wax, L. M. 1997. Grower acceptance of economic thresholds for weed management in Illinois. *Weed Technology*, 11(4), 828–831.

Duvick, D. N. 2005. Genetic progress in yield of United States maize (*Zea mays* L.). *Maydica*, 50, 193–202.

Eddy, P. R., Smith, A. M., Hill, B. D., Peddle, D. R., Coburn, C. A., and Blackshaw, R. E. 2013. Weed and crop discrimination using hyperspectral image data and reduced bandsets. *Canadian Journal of Remote Sensing*, 39(6), 481–490. doi:10.5589/m14-001.

Fennimore, S. A., Slaughter, D. C., Siemens, M. C., Leon, R. G., and Saber, M. N. 2016. Technology for automation of weed control in specialty crops. *Weed Technology*, 30(4), 823–837.

Feret, J.-B., François, C., Asner, G. P., Gitelson, A. A., Martin, R. E., Bidel, L. P. R., Ustin, S. L., le Maire, G., and Jacquemoud, S. 2008. PROSPECT-4 and 5: Advances in leaf optical properties model separating photosynthetic pigments. *Remote Sensing of Environment*, 112, 3030–3043. doi:10.1016/j.rse.2008.02.012.

Food and Agriculture Organization of the United Nations (FAO). 2017. *The Future of Food and Agriculture: Trends and Challenges. Rome*.

Furbank R. T., and Tester M. 2011. Phenomics—Technologies to relieve the phenotyping bottleneck. *Trends in Plant Science*, 16(12), 635–644. doi:10.1016/j.tplants.2011.09.005.

Gillan, J. K., Karl, J. W., Elaksher, A., and Duniway, M. C. 2017. Fine-resolution repeat topographic surveying of dryland landscapes using UAS-based structure-from-motion photogrammetry: Assessing accuracy and precision against traditional ground-based erosion measurements. *Remote Sensing*, 9(5), 437. doi:10.3390/rs9050437.

Gleadow, R., Johnson, A., and Tausz, M. 2013. Crops for a future climate. *Functional Plant Biology*, 40(2), 1–4. doi:10.1071/FPv40n2_FO.

Glendell, M., McShane, G., Farrow, L., James, M. R., Quinton, J., Anderson, K., Evans, M., Benaud, P., Rawlins, B., and Morgan, D. 2017. Testing the utility of structure-from-motion photogrammetry reconstructions using small unmanned aerial vehicles and ground photography to estimate the extent of upland soil erosion. *Earth Surface Processes and Landforms*. 42(12), 1860–1871. doi:10.1002/esp.4142.

Godfray, H. C. J., Beddington, J. R., Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F., Pretty, J., Robinson, S., Thomas, S. M., and Toulmin, C. 2010. Food security: The challenge of feeding 9 billion people. *Science*, 327(5967), 812–818. doi:10.1126/science.1185383.

Gómez-Candón, D., De Castro, A. I., and López-Granados, F. 2014. Assessing the accuracy of mosaics from unmanned aerial vehicle (UAV) imagery for precision agriculture purposes in wheat. *Precision Agriculture*, 15, 44–56. doi:10.1007/s11119-013-9335-4.

Gray, C. J., Shaw, D. R., Gerard, P. D., and Bruce, L. M. 2008. Utility of multispectral imagery for soybean and weed species differentiation. *Weed Technology*, 22, 713–718. doi:10.1614/WT-07-116.1.

Grenzdörffer, G. 2014. Crop height determination with UAS point clouds. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-1, 135–140. doi:10.5194/isprsarchives-XL-1-135-2014.

Grenzdörffer, G., Engel, A., and Teichert, B. 2008. The photogrammetric potential of low-cost UAVs in forestry and agriculture. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 21, 1207–1214.

Haboudane, D., Miller, J. R., Pattey, E., Zarco-Tejada, P. J., and Strachan, I. B. 2004. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sensing of Environment*, 90, 337–352. doi:10.1016/j.rse.2003.12.013.

Haboudane, D., Miller, J. R., Tremblay, N., Zarco-Tejada, P. J., and Dextraze, L. 2002. Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sensing of Environment*, 81(2–3), 416–426. doi:10.1016/S0034-4257(02)00018-4.

Haggar, R. J., Stent, C. J., and Isaac, S. 1983. A prototype hand-held patch sprayer for killing weeds activated by spectral differences in crop/weed canopies. *Journal of Agricultural Engineering Research*, 28(4), 349–358.

Hall, A. J., and Richards, R. A. 2013. Prognosis for genetic improvement of yield potential and water-limited yield of major grain crops. *Field Crops Research*, 143, 18–33. doi:10.1016/j.fcr.2012.05.014.

Hardin, P. J., and Jensen, R. R. 2011. Small-scale unmanned aerial vehicles in environmental remote sensing: Challenges and opportunities. *GISciences and Remote Sensing*, 48(1), 99–111. doi:10.2747/1548-1603.48.1.99.

Haug, S., and Ostermann, J. 2016. Plant classification for field robots: A machine vision approach. In *Computer Vision and Pattern Recognition in Environmental Informatics*, ed. J. Zhuo, B. Xioa, and C. Terry, 248–271. IGI Publishing, Hershey, PA.

Hernández-Hernández, J. L., García-Mateos, G., González-Esquiva, J. M., Escarabajal-Henarejos, D., Ruiz-Canales, A., and Molina-Martínez, J. M. 2016. Optimal color space selection method for plant/soil segmentation in agriculture. *Computers and Electronics in Agriculture*, 122, 124–132. doi:10.1016/j.compag.2016.01.020.

Herwitz, S. R., Johnson, L. F., Dunagan, S. E., Higgins, R. G., Sullivan, D. V., Zheng, J., Lobitz, B. M., et al. 2004. Imaging from an unmanned aerial vehicle: Agricultural surveillance and decision support. *Computers and Electronics in Agriculture*, 44, 49–61. doi:10.1016/j.compag.2004.02.006.

Hoffmann Jr., L., and Rooney, W.. 2014. Accumulation of biomass and compositional change over the growth season for six photoperiod sorghum lines. *BioEnergy Research*, 7, 811–815. doi:10.1007/s12155-013-9405-5.

Hruska, R., Mitchell, J., Anderson, M., and Glenn, N. F. 2012. Radiometric and geometric analysis of hyperspectral imagery acquired from an unmanned aerial vehicle. *Remote Sensing*, 4, 2736–2752. doi:10.3390/rs4092736.

Hsu, P.-H., and Tseng, Y.-H. I. 2000. Wavelet based analysis of hyperspectral data for detecting spectral features. *International Archives of Photogrammetry and Remote Sensing*, 33, 61–68.

Hummel, J. W., and Stoller, E. W. 2002. On-the-go weed sensing and herbicide application for the Northern Cornbelt. *2002 ASAE Annual Meeting*, Paper No. 021021. *American Society of Agricultural and Biological Engineers*. doi:10.13031/2013.9300.

Hunt Jr., E. R., Hively, W. D., Fujikawa, S. J., Linden, D. S., Daughtry, C. S. T., and McCarty, G.W. 2010. Acquisition of NIR-green-blue digital photographs from unmanned aerial aircraft for crop monitoring. *Remote Sensing*, 2(1), 290–305. doi:10.3390/rs2010290.

Jacquemoud, S., and Baret, F. 1990. PROSPECT: A model of leaf optical properties spectra. *Remote Sensing of Environment*, 34(2), 75–91. doi:10.1016/0034-4257(90)90100-Z.

Jacquemoud, S., Verhoef, W., Baret, F., Bacour, C., Zarco-Tejada, P. J., Asner, G. P., François, C., and Ustin, S. L. 2009. PROSPECT + SAIL models: A review of use for vegetation characterization. *Remote Sensing of Environment*, 113, S56–S66. doi:10.1016/j.rse.2008.01.026.

Johnson, G. A., and Hoverstad, T. R. 2002. Effect of row spacing and herbicide application timing on weed control and grain yield in corn (*Zea mays*). *Weed Technology*, 16(3), 548–553.

Kelcey, J., and Lucieer, A. 2012. Sensor correction of a 6-band multispectral imaging sensor for UAV remote sensing. *Remote Sensing*, 4, 1462–1493. doi:10.3390/rs4051462.

Koger, C. H., Bruce, L. M., Shaw, D. R., and Reddy, K. N. 2003. Wavelet analysis of hyperspectral reflectance data for detecting pitted morningglory (*Ipomoea lacunosa*) in soybean (*Glycine max*). *Remote Sensing of Environment*, 86(1), 108–119. doi:10.1016/S0034-4257(03)00071-3.

Kross, A., McNairn, H., Lapen, D., Sunohara, M., and Champagne, C. 2015. Assessment of RapidEye vegetation indices for estimation of leaf area index and biomass in corn and soybean crops. *International Journal of Applied Earth Observation and Geoinformation*, 34, 235–248. doi:10.1016/j.jag.2014.08.002.

Kumar, P., Gupta, D. K., Mishra, V. N., and Prasad, R. 2015. Comparison of support vector machine, artificial neural network, and spectral angle mapper algorithms for crop classification using LISS IV data. *International Journal of Remote Sensing*, 36(6), 1604–1617. doi:10.1080/2150704X.2015.1019015.

Liu, M., Liu, X., Wu, L., Duan, L., and Zhong, B. 2011. Wavelet-based detection of crop zinc stress assessment using hyperspectral reflectance. *Computers & Geosciences*, 37(9), 1254–1263. doi:10.1016/j.cageo.2010.11.019.

López-Granados, F. 2010. Weed detection for site-specific weed management: Mapping and real-time approaches. *Weed Research*, 51, 1–11. doi:10.1111/j.1365-3180.2010.00829.x.

Meyer, G. E., Mehta, T., Kocher, M. F., Mortensen, D. A., and Samal, A. 1998. Textural imaging and discriminate analysis for distinguishing weeds for spot spraying. *Transactions of the ASAE*, 41, 1189–1197.

Misiti, M., Misiti, Y., Oppenheim, G., and Poggi, J. M. 1996. *Wavelet Toolbox 4 User's Guide*. The MathWorks Inc., Natick, MA.

Moran, M. S., Bryant, R. B., Clarke, T. R., and Qi, J. 2001. Deployment and calibration of reference reflectance tarps for use with airborne imaging sensors. *Photogrammetric Engineering and Remote Sensing*, 67(3), 273–286.

Moreno, R., Corona, F., Lendasse, A., Graña, M., and Galvão, L. S. 2014. Extreme learning machines for soybean classification in remote sensing hyperspectral images. *Neurocomputing*, 128, 207–216. doi:10.1016/j.neucom.2013.03.057.

Pauli, D., Andrade-Sanchez, P., Carmo-Silva, A. E., Gazave, E., French, A. N., Heun, J., Hunsaker, D. J. et al. 2016. Field-based high-throughput plant phenotyping reveals the temporal patterns of quantitative trait loci associated with stress-responsive traits in cotton. *G3: Genes Genomes Genetics*, 6(4), 865–879. doi:10.1534/g3.115.023515.

Peña, J. M., Gutiérrez, P. A., Hervás-Martínez, C., Six, J., Plant, R. E., and López-Granados, F. 2014. Object-based image classification of summer crops with machine learning methods. *Remote Sensing*, 6(6), 5019–5041. doi:10.3390/rs6065019.

Peña, J. M., Torres-Sánchez, J., De Castro, A. I., Kelly, M., and López-Granados, F. 2013. Weed mapping in early-season maize fields using object-based analysis of unmanned aerial vehicle (UAV) images. *PLoS ONE*, 8(10), e77151. doi:10.1371/journal.pone.0077151.

Perry, J. H. 2009. A synthesized directly georeferenced remote sensing technique for small unmanned aerial vehicles. *MS Thesis*, University of Florida.

Prostko, E. P. 2015. *2015 Peanut Update*. Georgia Peanut Commission. www.gapeanuts.com/growerinfo/2015_ugapeanutupdate.pdf (accessed Aug. 15, 2017).

Rasmussen, J., Ntakos, G., Nielsen, J., Svensgaard, J., Poulsen, R. N., and Christensen, S. 2016. Are vegetation indices derived from consumer-grade cameras mounted on UAVs sufficiently reliable for assessing experimental plots? *European Journal of Agronomy*, 74, 75–92. doi:10.1016/j.eja.2015.11.026.

Ray, D. K., Mueller, N. D., West, P. C., and Foley, J. A. 2013. Yield trends are insufficient to double global crop production by 2050. *PLoS ONE*, 8(6), e66428. doi:10.1371/journal.pone.0066428.

Ruiz, L. A., Recio, J. A., Fernández-Sarriá, A., and Hermosillia, T. 2011. A feature extraction software tool for agricultural object-based image analysis. *Computers and Electronics in Agriculture*, 76(2), 284–296. doi:10.1016/j.compag.2011.02.007.

Sanches, I. D.'A., Filho, C. R. S., and Kokaly, R. F. 2014. Spectroscopic remote sensing of plant stress at leaf and canopy levels using the chlorophyll 680 nm absorption feature with continuum removal. *ISPRS Journal of Photogrammetry and Remote Sensing*, 97, 111–122. doi:10.1016/j.isprsjprs.2014.08.015.

Shi, Y., Thomasson, J. A., Murray, S. C., Pugh, N. A., Rooney, W. L., Shafian, S., Rajan, N. et al. 2016. Unmanned aerial vehicles for high-throughput phenotyping and agronomic research. *PLoS ONE*, 11(7), e0159781. doi:10.1371/journal.pone.0159781.

Siachalou, S., Mallinis, G., and Tsakiri-Strati, M. 2015. A hidden Markov models approach for crop classification: Linking crop phenology to time series of multi-sensor remote sensing data. *Remote Sensing*, 7(4), 3633–3650. doi:10.3390/rs70403633.

Tester, M., and Langridge, P. 2010. Breeding technologies to increase crop production in a changing world. *Science*, 327(5967), 818–822. doi:10.1126/science.1183700.

Thenkabail, P. S., Smith, R. B., and Pauw, E. D. 2000. Hyperspectral vegetation indices and their relationships with agricultural crop characteristics. *Remote Sensing of Environment*, 71(2), 158–182. doi:10.1016/S0034-4257(99)00067-X.

Thorp, K. R., DeJonge, K. C., Kaleita, A. L., Batchelor, W. D., and Paz, J. O. 2008. Methodology for the use of DSSAT models for precision agriculture decision support. *Computers and Electronics in Agriculture*, 64(2), 276–285. doi:10.1016/j.compag.2008.05.022.

Tilman, D., Balzer, C., Hill, J., and Befort, B. L. 2011. Global food demand and the sustainable intensification of agriculture. *Proceedings of the National Academy of Sciences*, 108(50), 20260–20264. doi:10.1073/pnas.1116437108.

Torres-Sánchez, J., López-Granados, F., De Castro A. I., and Peña-Barragán, J. M. 2013. Configuration and specifications of an unmanned aerial vehicle (UAV) for early site specific weed management. *PLoS ONE*, 8(3), e58210. doi:10.1371/journal.pone.0058210.

Torres-Sánchez, J., López-Granados, F., and Peña, J. M. 2015. An automatic object-based method for optimal thresholding in UAV images: Application for vegetation detection in herbaceous crops. *Computers and Electronics in Agriculture*, 114, 43–52. doi:10.1016/j.compag.2015.03.019.

Torres-Sánchez, J., Peña, J. M., De Castro A. I., and López-Granados, F. 2014. Multi-temporal mapping of the vegetation fraction in early-season wheat fields using images from UAV. *Computers and Electronics in Agriculture*, 103, 104–113. doi:10.1016/j.compag.2014.02.009.

Valasek, J., Henrickson, J., Bowden, E., Shi, Y., Morgan, C., and Neely, H. 2016. Multispectral and DSLR sensors for assessing crop stress in corn and cotton using fixed-wing unmanned air systems. *SPIE Autonomous Air and Ground Sensing Systems for Agricultural Optimization and Phenotyping*, 9866-19. SPIE. doi:10.1117/12.2228894.

Valasek, J., Lu, H.-H., and Shi, Y. 2017. Development and testing of a customized low-cost unmanned aircraft system based on multispectral and thermal sensing for precision agriculture applications. *Proceedings of the 2017 International Conference of Unmanned Aircraft Systems (ICUAS'17)*. IEEE. doi:10.1109/ICUAS.2017.7991494.

Wiles, L. J., Wilkerson, J. J., and Gold, H. J. 1992. Value of information about weed distribution for improving postemergence control decisions. *Crop Protection*, 11, 547–554.

Yang, C., Odvody, G. N., Fernandez, C. J., Landivar, J. A., Minzenmayer, R. R., Nichols, R. L. and Thomasson, J. A. 2014. Monitoring cotton root rot progression within a growing season using airborne multispectral imagery. *Journal of Cotton Science*, 18(1), 85–93.

Yi, S. 2016. FragMAP: A tool for long-term and cooperative monitoring and analysis of small-scale habitat fragmentation using and unmanned aerial vehicle. *International Journal of Remote Sensing*, 38(8–10), 2686–2697. doi:10.1080/01431161.2016.1253898.

Zarco-Tejada, P. J., González-Dugo, V., and Berni, J. A. J. 2012. Fluorescence, temperature and narrow-band indices acquired from a UAV platform for water stress detection using a micro-hyperspectral imager and a thermal camera. *Remote Sensing of Environment*, 117, 322–337. doi:10.1016/j.rse.2011.10.007.

Zarco-Tejada, P. J., Miller, J. R., Morales, A., Berjón, A., and Agüera, J. 2004. Hyperspectral indices and model simulation for chlorophyll estimation in open-canopy tree crops. *Remote Sensing of Environment*, 90, 463–476. doi:10.1016/j.rse.2004.01.017.

Zarco-Tejada, P. J., Rueda, C. A., and Ustin, S. L. 2003. Water content estimation in vegetation with MODIS reflectance data and model inversion methods. *Remote Sensing of Environment*, 85, 109–124. doi:10.1016/S0034-4257(02)00197-9.

Zhang, C., and Kovacs, J. M. 2012. The application of small unmanned aerial systems for precision agriculture: A review. *Precision Agriculture*, 13(6), 693–712. doi:10.1007/s11119-012-9274-5.

Zimdahl, R. C. 1999. *Fundamentals of Weed Science*, 2nd Edition. Academic Press, San Diego.

2 Building a UAV-Hyperspectral System I
UAV and Sensor Considerations

Cameron Proctor

CONTENTS

2.1 HYPERSPECTRAL IMAGERY AVAILABILITY

The emergence of commercial Earth imagery products and services companies, such as DigitalGlobe, has exploded the market availability of high-resolution (1–2 m) multispectral imagery. For many urban centers and the surrounding lands, imagery that meets the needs of the majority of the market is available from various resellers, or acquisitions can be ordered at reasonable costs. The hyperspectral market is considerably less developed and lacks centralized repositories of imagery. There is a considerable paucity of hyperspectral data availability in comparison to data from multispectral sensors. The largest provider of free hyperspectral data is NASA (distributed by the U.S> Geological Survey via http://aviris.jpl.nasa.gov/alt_locator/, http://earthexplorer.usgs.gov/, and http://glovis.usgs.gov/), who operates the airborne AVIRIS and satellite EO-1 Hyperion sensors.

Although EO-1 orbits in formation with Landsat, its swath width is a paltry 7.7 km versus 185 km and it is primarily a tasking satellite, that is, imagery acquisition is scheduled. Hence, its data availability is limited, although its geographic coverage contains the entire planet. In contrast, AVIRIS is an airborne system based out of California and flies primarily in the United States, although data requests have taken it to individual sites in Canada and the Caribbean. While nearly 2/3 of the California coast and 1/3 of the interior of the state have been subject to at least one imagery acquisition, coverage in other states is severely lacking (AVIRIS data

portal). It is difficult to isolate exact figures, but the majority of hyperspectral imagery acquisitions may be through private contractors. Airborne surveying companies, such as ITRES, Merrick, and SpecTIR, service a number of clients from the mining, oil and gas, agriculture, forestry, defense, and vegetation sectors. Because the imagery was acquired for these clients upon request, the data agreement will dictate whether the imagery can be resold to tertiary markets. Certainly, survey companies do not tend to advertise the availability of previously acquired imagery, suggesting that the majority of clients have no-resale clauses.

2.2 ADOPTION OF HYPERSPECTRAL UNMANNED AERIAL VEHICLES

Although the market for large- and small-scale hyperspectral surveying is small, there are a number of providers that do have the capabilities (e.g., ITRES, Specim). Why, then, construct a hyperspectral unmanned aerial vehicle (UAV) system? Certainly, the technology has not evolved sufficiently that a compact sensor package weighing less than 2 kg offers the same spectral fidelity and performance as systems many times its weight and cost. The hyperspectral UAV market is still in its infancy, and commercial off-the-shelf systems are rare, requiring users to build custom systems wherever technological barriers have not been completely resolved, guidance is difficult to find, and best practices have not been established. This leaves the adopter to take on the unusual and complex role of being responsible for logistics, insurance issues, aircraft maintenance, and other tasks typically reserved for service providers and built into the cost of commercial aerial surveys.

The obvious answer to the question of why construct a hyperspectral UAV is that they offer attractive advantages over contracting commercial aerial survey providers. Foremost is cost. Hyperspectral airborne surveys typically run upwards of $50–$200 per square kilometer, and minimum scanning area requirements invariably place the final cost in the order of tens of thousands of dollars. Mobilization costs consume a large portion of budgets because the plane and sensor package often need to be flown cross-country to the site of investigation, incurring costs from staff time, airport hangar fees, fuel, etc. These costs can place a substantial burden on users in remote locations. Users pay a premium for hyperspectral technology because the client base is smaller, there are few large anchor clients, and repeat business is uncertain, concentrating the costs of doing business such as maintenance and overhead.

The paradigm shift in owning a hyperspectral UAV is that start-up costs are very high, but each successive flight reduces the cost per square kilometer. Since the capital costs of a hyperspectral UAV are in the order of three to four airborne surveys, savings can begin to accrue after a few years of operations. However, the resources spent on getting a hyperspectral UAV program up and running are not trivial. Adopters must envision utilizing hyperspectral UAVs in the long term as a central element of their programs. Ongoing costs further strain research budgets. Annually, operators are required to purchase insurance, acquire flight permits, and oftentimes hire industry experts. A busy season of hyperspectral UAV imagery acquisition may bring the individual flight costs down to ~$1000 per flight (not including capital amortization). In addition, flight schedules or isolated areas that

would be cost-prohibitive with airborne surveys become more feasible with UAVs. For example, acquisition of glacier or wildlife imagery (Bhardwaj et al. 2016, Chabot and Bird 2012) simply requires shipping the UAV through regular channels and a large trunk. Repeat surveys at monthly time intervals are notoriously cheaper when the investigator owns the equipment, as opposed to making use of airborne service providers who are also obligated to service other clients. For wildlife researchers, UAVs have the added benefit of minimizing disturbances. Birds and animals tend to ignore UAVs, whereas in the same circumstances a typical airplane would elicit a reaction of panic.

The availability of the UAV means that there is no need to compromise on flight scheduling. The flexibility offered by a UAV is warranted for investigations of temporally limited or variable phenomena, such as phenological timing (Herwitz et al. 2004). Wine growers in particular have an economic need for repeated sampling within a narrow temporal interval to support decisions of optimal harvest times (Comba et al. 2015) or to detect water stress for irrigation purposes (Baluja et al. 2012) because wine sensorial attributes are negatively affected by drought (Chaves et al. 2007). Follow-up investigations of interesting phenomena or repeat flights to improve sensor performance or correct bad data are simpler with a UAV because possession of the data is near real time and quality control can be performed on site. Lastly, UAVs are a unique opportunity for experimentation and innovation. Traditional hyperspectral remote sensing is top-down, penetrating fewer than 1–2 mm into the vegetation canopy; few studies have investigated multiangle, three-dimensional scanning, or bidirectional reflectance distribution function of vegetation canopies. Unlike airborne surveys, UAVs offer precise flight control, allowing investigations from unique angles and perspectives (Hakala et al. 2010, Suomalainen et al. 2015).

The future prospects of hyperspectral UAVs are difficult to quantify because they lie on the precipice of technological innovation and user experience. The power offered is so new that its applications have yet to be fully realized. Extrapolating from the characteristics of hyperspectral UAVs, we can speculate that applications involving interaction with the environment are going to increase in prevalence. The precision agriculture community has already developed a range of UAVs that can sense and interact with their environment to deliver nutrients and pesticides in a highly spatially resolved manner (Giles and Billing 2015). Sense-and-follow technologies are also in the works for surveillance, wildfire detection, and tracking (Zhang et al. 2016). One day, hyperspectral UAVs might work in coordination with fleets of simple UAVs, with camera systems capable of detecting anomalies that could be further investigated by UAVs with more powerful sensory equipment. Sensor payload combinations will become more common as payloads get lighter, permitting coincident multiangle imagery capture and perhaps even digital surface modeling with hyperspectral overlay (Salamí et al. 2014).

2.3 HYPERSPECTRAL IMAGING FOR UAVs

Unsurprisingly, the military has the longest history of hyperspectral imaging, starting with the airborne imaging spectrometer in 1985 (Richter 2005), barely predating the Jet Propulsion Laboratory's AVIRIS system that began collecting data in 1987.

FIGURE 2.1 Range of UAV-suitable sensors offered by Headwall Photonics. (Left) Nano-Hyperspec; (Middle) Micro-Hyperspec VNIR; (Right) Micro-Hyperspec NIR. Images provided by Headwall Photonics.

Since then, hyperspectral imaging has flourished in the defense and commercial sectors. Early adopters of hyperspectral UAVs noted that the products available were simply downscaled versions of earlier manned-airplane-sized systems, and that many commercial solutions were not optimized for UAVs (Suomalainen et al. 2015). In recent times, the availability of UAV-suitable sensors has grown. Entry into the UAV market by hyperspectral imaging manufacturers with core business interests in the food processing and food quality sectors, among others, has increased the options available (e.g., Figure 2.1). Many companies now offer complete sensor packages, all requiring integrating costs, because there is no standardization for UAV payload size.

At their core, hyperspectral imagers are enabled by two key technologies: a wavelength dispersion element that divides polychromatic light into narrow distinct bands based on wavelength, and a two-dimensional detection array that converts radiance into quantifiable digital information (e.g., Figure 2.2). Examples of splitting techniques include dispersive gratings, prisms, spectrometers, Fabry-Perot interferometer, Michelson interferometers, and acoustic-optical tunable filters. An ideal spectral splitting technology would have explicit barriers between bands and be capable of projecting each band onto an array of photodetectors of 10–20 μm pixel size with no loss in radiance. Practically, no technique can achieve all these criteria, and must elicit trade-offs.

Typical photographic cameras capture imagery in two spatial dimensions, while videographic systems capture three dimensions by collecting a series of images in quick temporal succession. Hyperspectral imaging is a four-dimensional data collection system. To acquire a hyperspectral data cube (or hypercube), data must be collected in four dimensions: two spatial, one spectral, and one temporal. Unfortunately, all detectors consist of *XY* arrays of light-sensitive pixels, requiring either scanning or a snapshot approach where three dimensions are captured on one array.

Spatial scanning is the most common approach, and is also commonly referred to as line scanning, as the second spatial dimension is acquired through sensor movement (Figure 2.3). In line scanning, the spatial dimension perpendicular to the UAV's movement is scanned by blocking all incoming light, save for a narrow slit (∼10–40 μm). Hence, the spatial dimension is projected onto the *X*-dimension of the detector array, and a dispersion element is used to split the incoming light onto the detector's *Y*-axis. Note that a whisk broom sensor is a special case of spatial scanning,

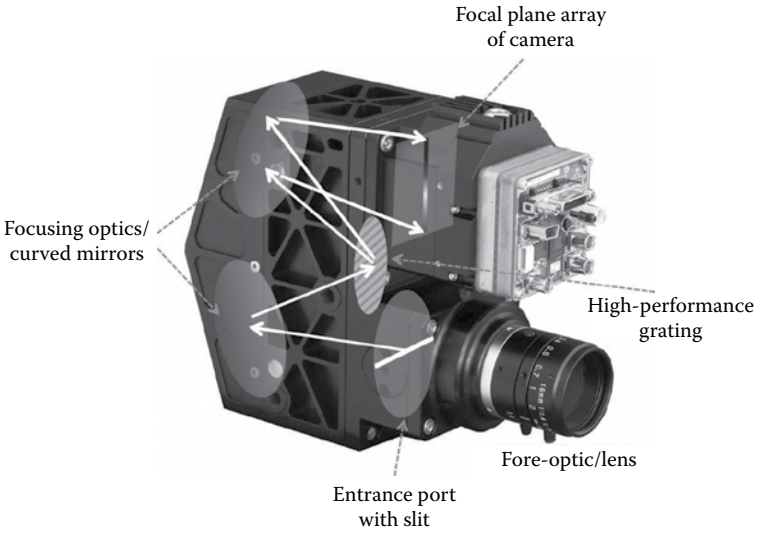


FIGURE 2.2 The optical layout for the Headwall Photonics hyperspectral image sensor comprises a diffraction grating along with focusing optics and curved mirrors. Light enters the unit from a narrow slit behind the lens, where it is dispersed by wavelength and projected onto a two-dimensional detector array. Image provided by Headwall Photonics.

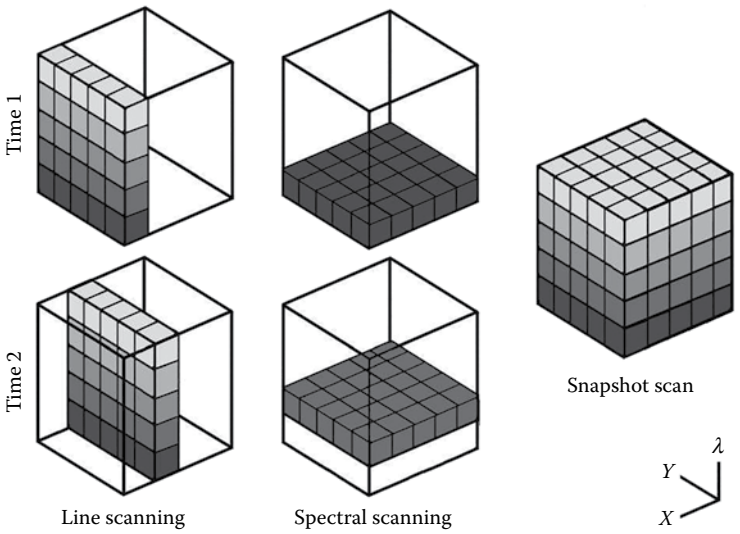


FIGURE 2.3 Various approaches to hyperspectral scanning. (Left) spatial scanning; (Middle) spectral scanning; (Right) snapshot scanner.

but has recently fallen out of favor because the rotating mirror adds mechanical components and weight.

Spectral scanning opts to record then project the entire image (i.e., both spatial dimensions) onto the detector array for a monochromatic (single wavelength interval) map of the scene, utilizing either a fixed or tunable band-pass filter. Hyperspectral imagery can be constructed by exchanging successive filters, if platform movement is minimal during the interval necessary to swap filters. There are certain advantages to a "photographic" representation of a scene. Tunable filters are particularly advantageous because they are not locked into manufacturer bands and thus can be optimized to target the user's wavelength region(s) of interest and acquire data on solely those regions, considerably restricting data output by eliminating the compulsory data collection from nonuseable bands.

Nonscanning or imaging spectrometers collect all spatial and spectral data on a two-dimensional array, yielding the complete hypercube in one snapshot. By nature of the spectral dimension being mixed with the spatial dimension into two dimensions, the detector arrays are quite complex. Certain approaches consist of tiled or mosaicked detector arrays and optical filters. The main advantages of these systems are that they offer higher light throughput and shorter acquisition times. Snapshot hyperspectral imagery is akin to traditional photography, and, as such, can take advantage of the more developed software market, which offers considerably more pushbutton tools for mosaicking imagery and digital surface model generation.

2.3.1 Pushbroom Hyperspectral

As of this writing, there are less than 10 pushbroom hyperspectral manufacturers serving the civil UAV market (Table 2.1). Note that no single sensor operates over the entire 400–2500 nm wavelength region. It is common practice in the hyperspectral remote sensing industry to utilize multiple detectors because no single detector material provides acceptable performance across the entire wavelength region. Field hyperspectral sensors, such as those provided by ASD, utilize three sensors: (1) a visible near-infrared (VNIR) detector (350–1000 nm) comprised of silicon; (2) a short-wave infrared (SWIR) detector (1001–1800 nm) comprised of an index indium gallium arsenide (InGaAs) photodiode; and (3) a two-stage thermoelectric (TE) cooled SWIR detector (1801–2500 nm) comprised of an index InGaAs photodiode. Accordingly, UAV hyperspectral sensor manufacturers offer separate units for the VNIR and near-infrared (NIR) regions.

Regardless of wavelength region, there are two main categories of wavelength dispersive technology commonly utilized: transmission (prism-based) and reflective (grating-based) (Figure 2.4). In transmissive dispersion, a prism splits the incoming light due to the change in refractive index, which deviates the path of light using Snell's law. The resolving power of a prism is proportional to its size and the higher the refractive index is from air (air = ~1.0 vs. optical glass ~1.7 and beyond, i.e., Schott N-LASF31A: $n = \sim 1.9$). Higher refractive index materials tend to have lower transmissivity. These design considerations are nonoptimal for small form factors, hence the development of curved transmission prisms to increase the dispersive power of a prism–grating–prism construction. Incoming light from the entrance

TABLE 2.1
Partial List of Hyperspectral Imager Manufacturers

Manufacturer	Products	Weight (kg)	VNIR	NIR	References
Resonon	PIKA	1.4–7.5	X	X	Hruska et al. (2012) Patterson and Brescia (2010) Swanson et al. (2007)
Headwall	Micro-Hyperspec	0.5–2.0	X	X	Li et al. (2015) Zarco-Tejada et al. (2012)
NovaSol	visNIR microHSI Alpha-vis microHSI	0.45–3.5[a]	X	X	Dayton et al. (2011)
OptoKnowledge	HyperScan VNIR micro	6.2	X		Gat et al. (2015)
Specim	AisaKESTREL	4.75–5	X	X	Rossini et al. (2013) Lin et al. (2013)
BaySpec	OCI-UAV	0.63	X		
Surface Optics	SOC710-GX	1.25	X		
VTT Technical Research	Snapshot camera	0.72	X		Honkavaara et al. (2012)

[a] Camera Only.

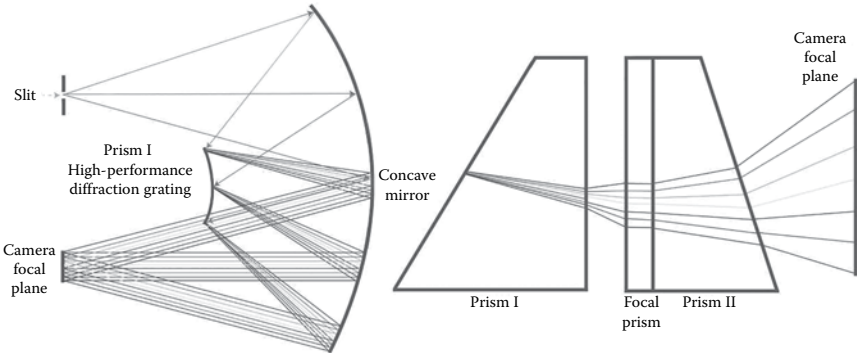


FIGURE 2.4 Comparison of prism and diffraction grating wavelength dispersion devices. Not to scale.

split is collimated by the front lens onto a volume transmission grating between two almost identical prisms. The transmissive optics are arranged such that the dispersed light is parallel and centered onto subsequent collimating lenses that project onto the detector. The advantages of transmissive wavelength dispersion are that the electro-optics are arranged in a straight optical path and offer uniformly high efficiency and low scatter. Transmissivity of visible light can be greater than 90% and vary negligibly by wavelength (Figure 2.5), allowing cheaper detector arrays to be utilized because of the high throughput. Lastly, wavelength dispersion is nonlinear

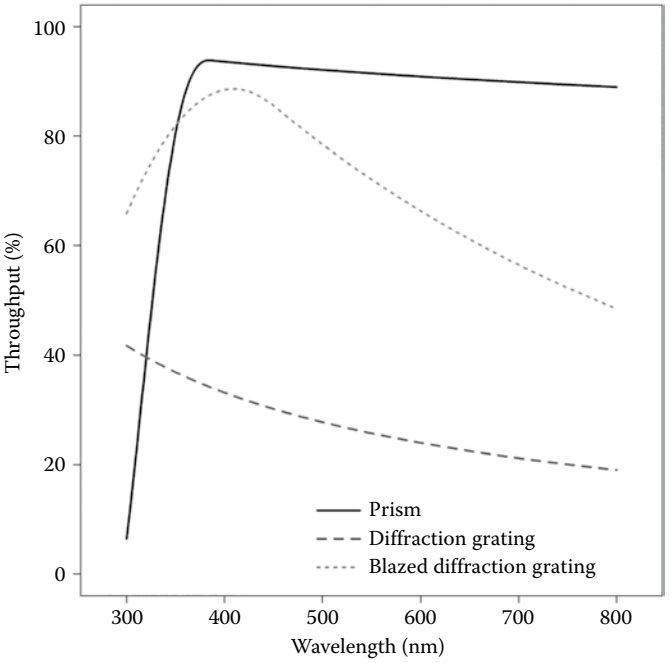


FIGURE 2.5 Throughput efficiency of prism and diffraction grating.

and can be high in prism-based systems, dispersing blue wavelengths less than red wavelengths. Hence, the difference in center wavelength between neighboring bands is not constant over the spectrum. For the prism-based EnMap spectrometer, the average spectral sampling distance is 6.5 nm, but varies to ~4.8 nm in the blue region and 8.2 nm at 1000 nm (Kaufman et al. 2015). Geometrical distortion in the spectral axis can be compensated for, at the cost of added components, weight, and complexity.

Refractive wavelength dispersive elements utilize a diffraction grating, consisting of periodic angled ridges that emanate light from each slit, causing additive and destructive interference resulting in structural dispersion. Changing the dimension of the periodic structure controls the amount of dispersion and energy diffracted. Hence, more control and higher resolving powers are achievable in comparison with transmissive optics, especially in the NIR. Concentrating most of the diffracted energy for a given wavelength is a technique called blazing. Blazing tends to improve the overall efficiency of the grating (Figure 2.5), yet concentrates the peak efficiency within a narrow range that descends rapidly in comparison with nonblazed gratings, where peak efficiency loss is minor over a longer wavelength range. Although light throughput is variable by wavelength, dispersion is constant, and the difference in center wavelength between neighboring bands is the same over the spectrum.

Plane and curved gratings are common; for instance, the Headwall photonics hyperspectral series, based on an Offner configuration with two concave mirrors and one convex grating (Figure 2.4). Incoming light from the entrance slit is guided

by the lower mirror to the reflection grating, where it is dispersed and guided by the upper mirror to the detector. The realization of an all-reflective optical system utilizing metallic materials achieves higher throughput because surface reflective losses that occur with transmissive optics are avoided. However, grating may exhibit more scattering than prisms, reducing the spectral purity. Furthermore, Offner-type spectrometers operate with a relatively low f-number ($\geq f/2$), accepting a long slit while maintaining a compact size and producing high image quality free of aberrations (Bannon and Thomas 2005). Reflection wavelength dispersion elements are well-suited to situations where high signal-to-noise ratio is desired.

2.3.2 Detector Array

Dispersed light is projected across a two-dimensional detector array, whose sole function is to quantify the intensity of the collected light by converting radiation energy into electrical signals. By this definition, the technology within a digital photographic camera or even a cell phone is suitable for a hyperspectral sensor. While hyperspectral sensors do employ detector arrays within the same family, the detector array quality is considerably heightened. Typically, detector arrays of machine vision or scientific grade are utilized. These detectors distinguish themselves by providing high performance, with frames per second and noise as specific key parameters. Higher frames per second are preferential for capturing dynamic motion. Low noise is preferential in all situations because noise introduces artifacts into the image, which results in a grainy appearance. Noisy data compound as the number of bands utilized increases. Considerable research effort has been spent on quantifying the signal-to-noise ratio (Weatherbee and Procino 2010), which is a key parameter for deciding between detector array options.

No detector array technology is superior across all metrics. Performance is often seated in physics limits that are unconcerned with budget size. Hard trade-offs are wrought between competing technologies, and there often exists a crossover point where a detector array that performs poorly in all other regards will display superior performance in one. The main detector array technologies are the silicon charged couple device (CCD), scientific complimentary metal-oxide semiconductor (CMOS), and InGaAs CMOS hybrid. The working principle of a CCD is light absorption by a photoelectron collection well, in contrast with CMOS, which utilizes a photodiode-amplifier pair for achieving photoelectric conversion (Bigas et al. 2006). A CCD has off-detector analog-to-digital conversion that is less parallel than the on-chip conversion in CMOS. Requiring all electrical signals to be converted into digital at a common port limits acquisition speed. Consequently, CMOS detector arrays are more flexible, and higher-end versions can be designed to have much lower noise than an equivalent CCD at the highest frame rates.

Because both CCD and CMOS utilize silicone as the means of absorbing photons, they are physically limited to an absorbance range regulated by its molecular structure. Photoconversion in silicone declines considerably beyond 800 nm (Figure 2.6), and is typically replaced by InGaAS arrays for NIR regions. InGaAs arrays bear more similarity to CMOS, with the main differing factor being the compound matrix utilized to absorb photons in the NIR. Although an InGaAs detector array

has superior performance in the NIR, array noise is considerably more sensitive to temperature. Cooling elements of various voracities often accompany InGaAs detectors and add to the total system weight.

Comparisons from various manufacturers of the performances of their detection arrays abound. It is challenging to determine the veracity of each manufacturer's claims. Considering the variability within families of detector arrays, coupled with the subtle technological improvements that aid performance, general rules are suspect. Sensitivity in the NIR regions is a good example. A brief Internet search will provide abundant examples of CCD and CMOS detector array comparisons touting superior performance, depending on the manufacturer's bias. Perhaps the only dependable conclusions are that CCDs permeate the market, and traditionally offer high performance and low noise. Saturation can be a problem with CCDs under high light conditions. The response of CCDs to light intensity is not linear, especially beyond 85% of the bit depth. As the market increases, CCD costs are spread over a larger client base, providing more performance for cost. However, newer generation CMOSs compete with CCDs and best them in frame rates, field of view (FOV) (spatial resolution), dynamic range, and noise at higher frame rates. Regardless of rules of thumb, the adopter should consult with the manufacturer to assess the trade-offs between offered models. Often, the deficiencies of one type of detector array can be compensated for via other components on the hyperspectral sensor, or the performance gained from selecting a higher-end model may not be relevant to the user's application.

Additional parameters influence real-world detector performance. For instance, the ability of the array to produce large voltage from a few photons is arguably as important as, if not more important than, the wavelength dispersion device in determining the final signal-to-noise ratio. The main detector characteristics regulating performance are quantum efficiency (QE), integration time, and noise. QE refers to the conversion of photons into electrons, which are included in the intensity of the incoming light signal. An ideal detector would have a QE of 1.0 across all wavelengths; in practice, detectors convert up to 80% of VNIR and 65% of NIR light, primarily governed by the substance used. Efficiency is lowered as some light is absorbed by the detector itself or the structural supports surrounding the detector, or the light hits upon a region of the detector that provides a weak or no electrical response. Conversion of photons to electrical signal depends upon the band structure of the detector and is thus wavelength sensitive. QE typically descends from its peak value until it reaches a wavelength cutoff, where a signal beyond the typical noise is unmeasurable. For VNIR detectors, the QE per wavelength often forms a bell-shaped curve, with declining efficiency toward the ultraviolet (UV) and NIR regions (Figure 2.6), while NIR detectors feature a more flattened response across a broad spectral range, declining sharply into the visible and further NIR.

Other operating parameters that have a large impact on performance are the interrelated characteristics of dynamic range, full well capacity, and bit depth. These values are related to the pixel physics. Currents produced by detectors are in the order of femtoamperes, which discourages accurate measurements. Detectors must be able to recognize slight amperage variations in order to accurately parse the analog signal into discrete packets of digital information. Analog signals are converted into digital

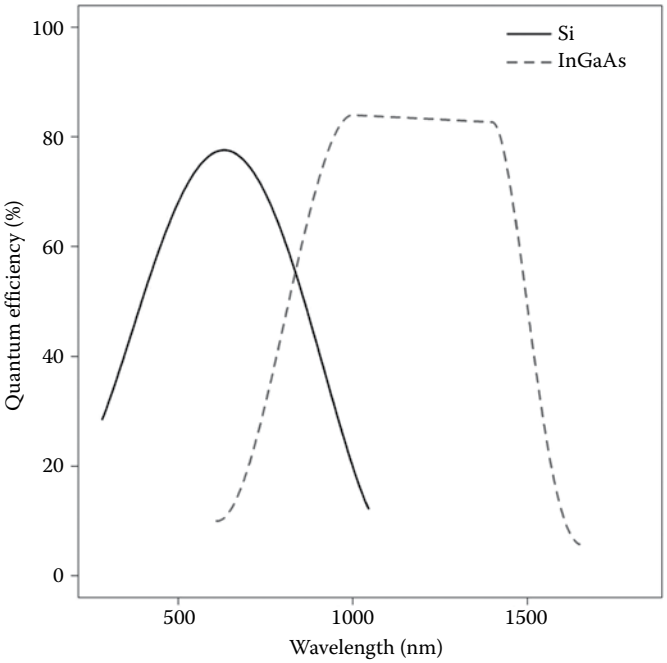


FIGURE 2.6 Theoretical quantum efficiency of Si- and InGaAs-based detector arrays.

number of a range given by the power to base 2, typically 2^8 (0–255) or 2^{12} (0–1023) levels. One approach to getting around this physical limitation is to allow a large charge to accumulate. The greater the maximum charge the detector can accumulate, the greater the amperage gap between digital numbers can be widened, or a larger range of digital numbers can be recorded based on the detector's electrical sensitivity threshold.

Full well capacity refers to the number of electrons each detector pixel can accumulate before saturation. A closely related parameter is dynamic range, which refers to the ratio between full well capacity and noise. If a sensor has a full well capacity of 22,000 electrons and a readout noise of 7, the dynamic range is 3142 : 1. This value provides a measure of how clearly a signal appears from the noise. Strong signals above the noise floor can thus be parsed into small divisions, increasing the permissible bit depth. For the above sensor, a bit depth of 10 (1024 grey levels) creates a conversion ratio (the number of electrons needed to distinguish between grey levels) of ~21.48 electrons, which is well above the detector's electrical sensitivity threshold. Conversely, a bit depth of 12 (4096 grey levels) has a conversion ratio of ~5.37, which is a more reasonable compromise. Generally, a larger bit depth is advantageous because it retains finer scale variations from the original image. Yet caution is warranted when interpreting the performance gain from any one parameter because many detector characteristics are interrelated. Often, the drawbacks of one parameter are partially compensated for by another, or have little impact on a specific application area. Premium performance metrics cost considerably more than products intended for diverse markets.

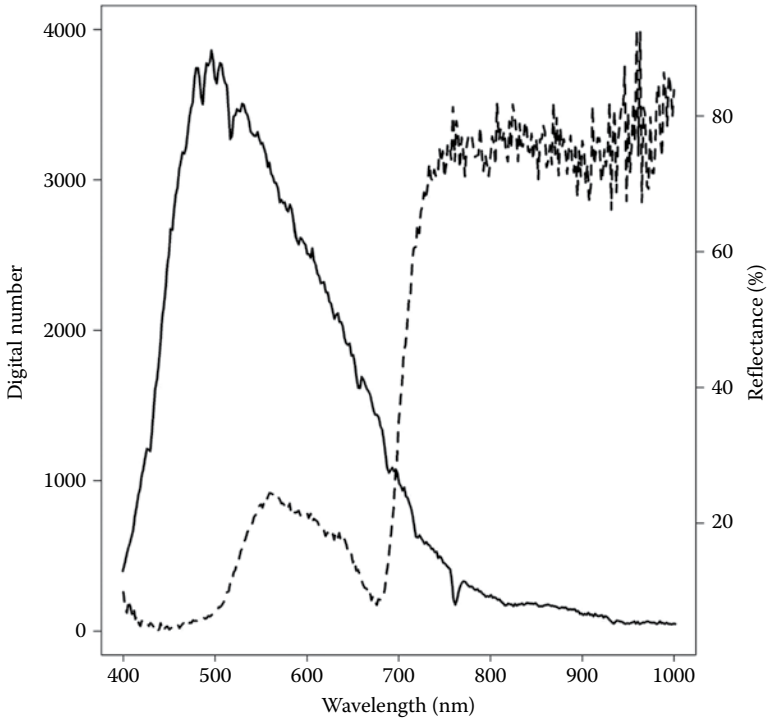


FIGURE 2.7 Radiance and reflectance of a green leaf, acquired by a Headwall Photonics Micro-Hyperspec A Series.

Overall, VNIR detectors can have substantial bias against sensitivity to NIR wavelengths, resulting in low signal to noise. Figure 2.7 shows a typical spectrum recorded by a Headwall Photonics Micro-Hyperspec A Series hyperspectral imager under clear sky conditions at solar noon of a white reference. Note that the spectrum drops off considerably beyond 800 nm.

2.3.3 Hyperspectral Sensor Performance Metrics

The main challenge of hyperspectral imaging is dividing the light passing through the entrance slit through a wavelength dispersion element and then onto an array of detectors of micrometer proportions. Misalignment errors of the projection of the ground sampling area onto the detector induces overlap between pixels. Specifically, spectral misalignment tilts the band center wavelength such that the poles of the detector, supposedly receiving light at 500 nm, may in actuality be receiving light 2–4 nm differently, depending on the degree of misalignment. Spatial misalignment can also occur, such that within the wavelength range (e.g., 400–1000 nm), shorter wavelengths may receive radiation from the proper ground area, whereas at longer wavelengths the radiation is received from the adjacent pixel. Aside from homogenous

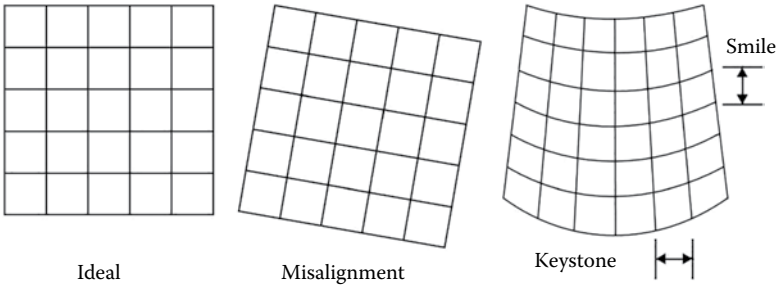


FIGURE 2.8 Issues with projection of dispersed light onto a detector array. (Left) ideal case of ordered alignment; (Middle) misalignment of the projection onto the array; (Right) distortion caused by smile and keystone.

misalignment problems, the majority of detectors suffer curved misalignment effects to a certain degree. Known as smile effects and keystone effects (Figure 2.8), these properties denote the variations in misalignment in the spectral and spatial domains. Smile effects are a curvature in the misalignment in the spectral domain, caused by imperfections in the optics or the dispersive element. Similarly, keystone effects introduce a bias in projection onto the spatial domain. Keystone effects are typically negligible at nadir and increase at the poles of the FOV. Adopters should inquire with manufacturers regarding their sensors' smile and keystone performances across the entire FOV and wavelengths, as opposed to a sole "averaged" value.

Assuming the incoming light is properly registered on the detector, the frame of the lens can shade incoming radiation with high incident angles. Known as a vignetting effect, this results in a radial light intensity falloff toward the image edges. Under homogenous lighting conditions, the light intensity received can vary by 30%–40% between the detector center and edge (Goldman 2010). Vignetting can be induced through a variety of causes that are explained by Goldman (2010) and Kim and Pollefeys (2008). Correcting vignetting by modeling the optical pathway or by image-based techniques (Yu 2004) is possible. Image-based techniques utilize an integrating sphere of known radiance output to acquire a gain and offset matrix for each pixel of the detector along both the spatial and spectral dimensions.

Adopters of hyperspectral UAVs incorporate a range of quality assurance diagnostics as part of their image acquisition workflow. Many undertake the bare minimum, such as dark current calibration, and simply utilize the manufacturer-provided calibration software modules. Others have a more robust processing chain, wherein they themselves take additional steps to quantify the precise radiometric response and noise introduced by electronic components and utilize custom software to correct artifacts in the incoming signal. Various types of quality assurance or error modeling and corrections procedures have been attempted, such as laboratory and field calibration (Liu et al. 2014), distortion correction (Gorsevski and Gessler 2009), point spread function (PSF) nonuniformity effects (Schläpfer et al. 2007), camera calibration (Aasen et al. 2015), and signal-to-noise modeling. These procedures range in complexity and should be considered in applications requiring a high degree of spectral and spatial accuracy.

2.4 CONCLUSIONS

As of 2017, the hyperspectral UAV market is poised to expand, providing a wealth of high spatial resolution hyperspectral data at hitherto unimaginable costs. Coupled with the ease of integrating commercial off-the-shelf hardware, adopters are offered a variety of turnkey and custom hyperspectral solutions suitable for UAV imagery acquisition. However, challenges remain when assessing technical specifications and understanding how they convert to real-world performance. Lacking a coalition of experts instituting rules for common practice, early adopters are squarely in the product testing and deployment phase. They pay an "early adopter tax," in that commercial offerings are expensive, possibly prone to malfunction, and crude. Early adopters are also placed in the role of the trendsetter, their choices adopted by the next generation. Hence, those thinking of adopting a hyperspectral UAV may bear some responsibility for developing core practices and pragmatic advice in their area of application. Only by understanding the technology underlying hyperspectral UAV systems can this burden be lightened.

REFERENCES

Aasen H, Burkart A, Bolten A, Bareth G. 2015. Generating 3D hyperspectral information with lightweight UAV snapshot cameras for vegetation monitoring: From camera calibration to quality assurance. *ISPRS Journal of Photogrammetry and Remote Sensing*. 108, 245–259.

Bannon D, Thomas R. 2005. Harsh environments dictate design of imaging spectrometer. *Laser Focus World*. 4(8): 93–95.

Baluja J, Diago M, Balda P, Zorer R, Meggio F, Morales F, Tardaguila J. 2012. Assessment of vineyard water status variability by thermal and multispectral imagery using an unmanned aerial vehicle (UAV). *Irrigation Science*. 30(1): 511–522.

Bigas M, Cabruja E, Forest J, Salvi J. 2006. Review of CMOS image sensors. *Microelectronics*. 37(1): 433–451.

Bhardwaj A, Sam L, Akanksha F, Martín-Torres FJ, Kumar R. 2016. UAVs as remote sensing platform in glaciology: Present applications and future prospects. *Remote Sensing of Environment*. 175(1): 196–204.

Chabot D, Bird D. 2012. Evaluation of an off-the-shelf unmanned aircraft system for surveying flocks of geese. *Waterbirds*. 35(1): 170–174.

Chaves M, Santos T, Souza C, Ortuno M, Rodrigues M, Lopes C, Maroco J, Pereira J. 2007. Deficit irrigation in grapevine improves water-use efficiency while controlling vigour and production quality. *Annals of Applied Biology*. 150, 237–252.

Comba L, Gay P, Primicerio J, Aimonino D. 2015. Vineyard detection from unmanned aerial systems images. *Computers and Electronics in Agriculture*. 114, 78–87.

Dayton D, Nolasco R, Myers M, Gonglewski J, Fertig G, Even D, Hill B. 2011. Maritime Hawaii hyperspectral measurements using a SWIR camera. In *Proceedings of SPIE 8186, Electro-Optical Remote Sensing, Photonic Technologies, and Applications V*, edited by G Kamerman, O Steinvall, K Lewis, G Bishop, J Gonglewski.

Goldman D. 2010 Vignette and exposure calibration and compensation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 32(12): 2276–2288.

Gat N, Scriven G, Stramski D, Reynolds R. 2015. The AquaScan—A novel marine environments hyperspectral sensor. In *Meeting of the Military Sensing Symposia (MSS) Passive Sensors*, Gaithersburg, MD, 15 September 2015.

Gilles D, Billing R. 2015. Deployment and performance of a UAV for crop spraying. *Chemical Engineering Transactions*. 44(1): 307–312.

Gorsevski P, Gessler P. 2009. The design and the development of a hyperspectral and multispectral airborne mapping system. *ISPRS Journal of Photogrammetry and Remote Sensing*. 64, 184–192.

Hakala T, Suomalainen J, Peltoniemi J. 2010. Acquisition of bidirectional reflectance factor dataset using a micro unmanned aerial vehicle and a consumer camera. *Remote Sensing*. 2, 819–832.

Herwitz S, Johnson L, Dunagan S, Higgins R, Sullivan D, Zheng J, Lobitz B et al. 2004. Imaging from an unmanned aerial vehicle: Agricultural surveillance and decision support. *Computers and Electronics in Agriculture*. 44(1): 49–61.

Honkavaara E, Kaivosoja J, Mäkynen J, Pellikka I, Pesonen L, Saari H, Salo H, Hakala T, Marklelin L, Rosnell T. 2012. Hyperspectral reflectance signatures and point clouds for precision agriculture by light weight UAV imaging system. In *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Volume I-7. XXII ISPRS Congress*, Melbourne, Australia, 25 August–01 September 2012.

Hruska R, Mitchell J, Anderson M, Glenn N. 2012. Radiometric and geometric analysis of hyperspectral imagery acquired from an unmanned aerial vehicle. *Remote Sensing*. 4(12): 2736–2752.

Kaufman H, Sang B, Storch T, Segl K, Foerster S, Guanter L, Erhard M et al. 2015. Environmental mapping and analysis program—A German hyperspectral. In *Optical Payloads for Space Missions*, edited by Q Shen-En. John Wiley & Sons, Ltd.

Kim S, Pollefeys M. 2008. Robust radiometric calibration and vignetting correction. *IEEE Transactions on Pattern Analysis and Machine Intelligence Archive*. 30(4): 562–576.

Li H, Zhang H, Zhang B, Chen Z, Yang M, Zhang Y. 2015. A method suitable for vicarious calibration of a UAV hyperspectral remote sensor. *IEEE Journal of Selected Topics in Applied Earth Observation and Remote Sensing*. 8(6): 3209–3223.

Lin Y, Puttonen E, Hyyppä J. 2013. Investigation of tree spectral reflectance characteristics using a mobile terrestrial line spectrometer and laser scanner. *Sensors (Switzerland)*. 13(7): 9305–9320.

Liu Y, Wang T, Ma L, Wang N. 2014. Spectral calibration of hyperspectral data observed from a hyperspectrometer loaded on an unmanned aerial vehicle platform. *IEEE Journal of Selected Topics in Applied Earth Observation and Remote Sensing*. 7(6): 2630–2638.

Patterson M, Brescia A. 2010. Operation of small sensor payloads on tactical sized unmanned air vehicles. *Aeronautical Journal*. 114, 427–436.

Rossini M, Fava F, Cogliati S, Amaducci S, Colombo R. 2013. Assessing canopy PRI from airborne imagery to map water stress in maize. *ISPRS Journal of Photogrammetry and Remote Sensing*. 86, 168–177.

Richter R. 2005. Hyperspectral sensors for military applications. In *Emerging EO Phenomenology* (pp. KN1-1 – KN1-4). Meeting Proceedings RTO-MP-SET-094, Keynote 1. Neuilly-sur-Seine, France: RTO. Available from: http://www.rto.nato.int/abstracts.asp.

Salamí E, Barrado C, Pastor E. 2014. UAV flight experiments applied to the remote sensing of vegetated areas. *Remote Sensing*. 6, 11051–11081.

Schläpfer D, Nieke J, Itten K. 2007. Spatial PSF non-uniformity effects in airborne pushbroom imaging spectrometry data. *IEEE Transactions on Geoscience and Remote Sensing*. 45(2): 458–468.

Suomalainen J, Roosjen P, Bartholomeus H, Clevers J. 2015. Reflectance anisotropy measurements using a pushbroom spectrometer mounted on UAV and a laboratory goniometer—Preliminary results. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XL-1/W4, 2015 International Conference on Unmanned Aerial Vehicles in Geomatics*, Toronto, Canada, 30 August –02 September 2015.

Swanson R, Kehoe M, Smith C, Moon T, Bousquet R, Brown S, Lykke K, Maciejewski P, Barnard K. 2007. Compact anamorphic imaging spectrometer. In *Meeting of the Military Sensing Symposia (MSS) Specialty Group on Camouflage, Concealment & Deception; Passive Sensors; Detectors; and Materials*, Vol. 1, February 5–9, 2007, SENSIAC Military Sensing Information Analysis Center, Georgia Institute of Technology.

Weatherbee O, Procino W. 2010. Signal-to-noise ratio (SNR) discussion. Resonon whitepaper. http://www.resonon.com/whitepapers/SNR%20Discussion_v4.pdf

Yu W. 2004. Practical anti-vignetting methods for digital cameras. *IEEE Transactions on Consumer Electronics*. 50(4): 975–983.

Zarco-Tejada P, González-Dugo V, Berni J. 2012. Fluorescence, temperature and narrow-band indices acquired from a UAV platform for water stress detection using a micro-hyperspectral imager and a thermal camera. *Remote Sensing of Environment*. 117(15): 322–337.

Zhang M, Hugh H, Liu J. 2016. Cooperative tracking a moving target using multiple fixed-wing UAVs. *Intelligent Robot Systems*. 81, 505–529.

3 Building a UAV-Hyperspectral System II
Hyperspectral Sensor Considerations and Data Preprocessing

Cameron Proctor

CONTENTS

3.1 TECHNICAL CONSIDERATIONS

Despite the remote sensing community having amassed a large amount of expertise in hyperspectral data processing, few have the opportunity to venture into the engineering realm. Technical problems from integrating electronic and optical components or contrasting manufacturer products in order to make purchasing decisions occur when the adopter's knowledge of hyperspectral hardware is limited. The industry's relative lack of consumer reviews and reluctance to provide technical specifications beyond tombstone data leaves adopters with few means to discern between instruments. Often, subtle technical parameters count more toward the performance of the final system, and choices that appear simplistic can have drastic consequences. Adopters can be swayed by such notions as a turnkey system or a lower price, neither of which is necessarily likely to produce a scientific instrument that fits the user's requirements.

Adopters often fail to realize the mismatch between their needs and the sensor until the application phase, by which time considerable time, money, and resource expenditure have locked them in with a suboptimal system. The preventative is to delve deep into understanding the inner workings of a hyperspectral unmanned aerial vehicle (UAV).

3.2 RELATIONSHIP BETWEEN SENSOR AND UAV

Although it is tempting to envision a sole hyperspectral UAV for all application areas, on a practical level sensor hardware and flight characteristics are interrelated and work together to constrain operating parameters (Patterson and Brescia 2010). For example, AVIRIS spatial resolution varies depending on the altitude of the aircraft, and is typically captured at two tiers based on the aircraft platform. The larger NASA/Ames Research Center ER-2 operates at 20 km altitude, 11 km swath width, and 20 m resolution, in contrast with Twin Otter International's lower-altitude turboprop, which operates at 4 km altitude, 4 km swath width, and 4 m resolution (Greena et al. 1998). On account of the increasing field of view (FOV) with altitude, the plane must vary either its flight speed or its sensing interval in order to compensate, hence the ER-2 operates at 734 km/hr (as opposed to the Twin Otters' 222 km/hr). The fact that the AVIRIS sensor had to be ported to a different aircraft altogether to achieve higher spatial resolution alludes to the considerably narrow operating range regulated by the hardware.

In the UAV market, adopters can select from a wide range of optics, imager frame rates, and UAV platforms that dictate the operating parameters. The endless combinations make it challenging for manufacturers to transform the technical specifications of their hardware into real-world performance. Often, responsibility lies with the adopter. In most cases, adopters are interested in the size of the swath width in order to obtain an aerial coverage similar to that of airborne or satellite systems. The determinants of swath width are the angular field of view and the flight altitude (Figure 3.1), according to

$$\text{Swath} = (2 \times H) \times \tan\left(\frac{\theta_{\text{FOV}}}{2}\right)$$

where H is the altitude in meters and θ_{Fov} is the angular field of view in degrees.

The swath width equation demonstrates that altitude has a linear relationship with swath width, whereas FOV approaches linearity. For a 25° FOV sensor at 50 m altitude, tripling the flight height increases the swath width from 22.16 to 66.5 m, in contrast with tripling the FOV, which increases it to 76.7 m (~3.4 times).

Swath width is a prime determinant of the area coverage a UAV hyperspectral is capable of scanning. However, pushbroom scanning by definition involves two spatial dimensions: cross track (orthogonal to the flight direction, i.e., swath width) and along track (parallel to the flight direction). Distances traversed in the along-track dimension depend on the speed of the UAV and the flight time spent scanning. A 20-minute flight at 200 m altitude using a 48.7° lens with a fixed-wing UAV at 22 m/s covers 4.7 km^2 (181 m cross-track × 26,400 m). Scanned areas, in practice, are much smaller because UAV endurance is spent on mobilizing to the site and

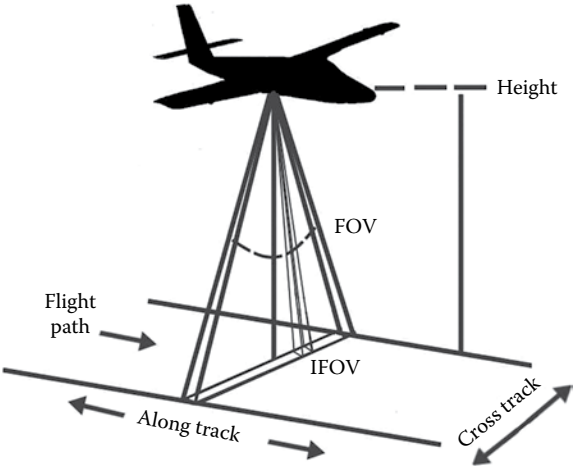


FIGURE 3.1 Typical operation of a pushbroom UAV in nadir view.

in-flight maneuvers. Common practice further dictates an overlap of 30%–60% between scan lines, reducing the coverage gain per additional scan line (Figure 3.2). Scan line overlap reduces the likelihood of gaps due to divergence from the planned UAV position and sensor orientation.

Arguably, the greatest parameter of interest to hyperspectral UAV adopters is the ground sampling distance. The smaller the ground sampling distance, the finer the detailed elements of objects that can be resolved. Ideally, any object of interest in the image should be at least 4 pixels in size to be detectable. A ratio of ground sampling distance to object size of 10 : 1 is preferred for object-based classification methods. In a pushbroom configuration, each pixel in the detector array is sensitive to radiation given by the lens' instantaneous field of view (IFOV). At nadir, the

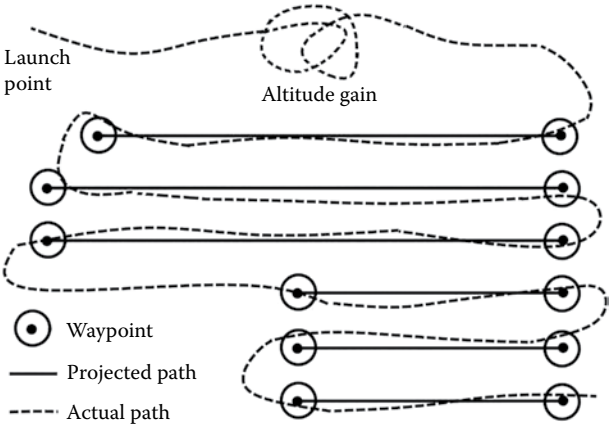


FIGURE 3.2 Preflight mission planning showing set waypoints versus actual path of UAV.

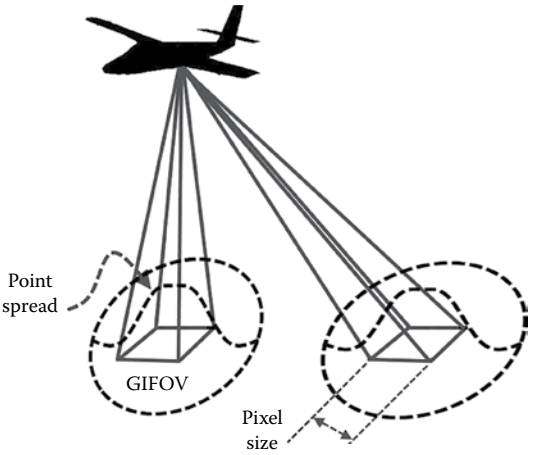


FIGURE 3.3 Difference between ground sampling distance, GIFOV, and pixel size for a hyperspectral imager.

IFOV is projected as a circle along the ground [ground IFOV (GIFOV)] (Figure 3.3), according to

$$\text{GIFOV}_{\text{nadir}} = 2 \times H \times \tan\left(\frac{\theta_{\text{IFOV}}}{2}\right)$$

where θ_{IFOV} is the instantaneous field of view.

Note that the GIFOV at nadir is greater than the area sensed and allocated to 1 pixel. However, the contribution of radiation to the final value from within the GIFOV is not homogenous. The contribution peaks near the center and decreases around the edges of the GIFOV; it further varies by observing conditions and as a function of the wavelength. Known as the *point spread function*, complicated measurements or modeling are often required to provide a reasonable estimate of this quantity. A Gaussian point spread function is assumed; however, the validity of this assumption over the detector array with wavelength is questionable. Often, the point spread function will exhibit some anisotropy and non-Gaussian behavior. The takeaway is that the ground sampled by the detector extends beyond the ground attributed to a pixel. Thus, each detector pixel receives some contamination from the adjacent pixels (cross track) and between scan lines (along track). Although an image may look sharp, radiation from the ground surrounding the pixel is included without being visually distinguishable.

The IFOV is an angular quantity. At nadir, the projection of the IFOV forms a circle on the ground. Off-nadir, the IFOV projection becomes more elliptical because it intersects at increasing angles. The view angle effect increases the area viewed on the ground by

$$\Delta X(\theta) = H(\tan(\theta + \theta_{\text{IFOV}}) - \tan(\theta - \theta_{\text{IFOV}}))$$

where ΔX is the cross-track area scanned on the ground and θ is the angle between the nadir and the view angle (typically the center of the IFOV of the off-nadir pixel). The

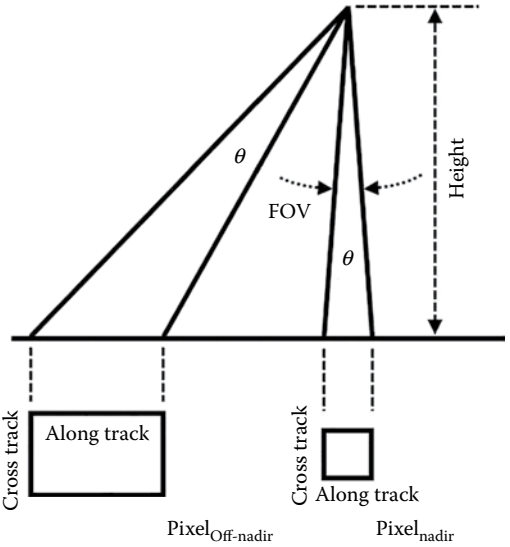


FIGURE 3.4 Example of off-nadir FOV and associated GIFOV.

elliptical GIFOV at off-nadir angles increases overlap. In addition, the area from which signal is received and the path length through the atmosphere are increased. For a UAV at 200 m with an IFOV of 0.902 mrad, the GIFOVs at nadir, 12.17°, and 24.4° are 0.36, 0.38, and 0.435 m, respectively (Figure 3.4). Hence, the sampling density (number of samples per ground area) is reduced at the poles of the scan line. View angle effects can be enhanced by changes in sensor orientation due to the roll of the UAV.

Without movement of the UAV, the sensor will acquire data across the entire scan line. Forward motion by the UAV elongates the GIFOV into an oval shape. However, most data processing systems do not have the capabilities to represent the sampling area as an array of overlapping oval shapes. For simplification, the GIFOV is allocated into a nonoverlapping pixel array whose *X* and *Y* dimensions attempt to preserve the geometric fidelity of the sampling area. Compelling the GIFOV into a pixel necessitates treating the cross- and along-track dimensions independently, or distortions can occur during geocorrection (reallocating image pixels to their correct geographic positions) and geocoding (assigning to each pixel an estimation of its correct geographic coordinates). Forcing pixels elongated in the along-track direction into a square representation will cause gaps between scan lines, while contracted pixels will overlap.

To achieve square pixels, it is common practice to predetermine the necessary combination of the proper operating parameters of the UAV and sensor before purchasing hardware. Many manufacturers offer tools or expertise to help adopters make proper purchasing decisions. It is highly recommended that adopters make use of these offerings. Ignoring GIFOV, a simplified approximation for cross- and along-track ground sampling distance can be determined. Generally, flight altitude and FOV determine the cross-track pixel resolution, while UAV ground speed and sensor integration time determine the along-track pixel resolution.

$$\mathrm{GSD}_{ct} = (2 \times H) \times \tan\left(\frac{\theta_{\mathrm{FOV}}}{2}\right) / \text{Number of spatial pixels}$$

$$\mathrm{GSD}_{at} = v\Delta t = v \times 1/\mathrm{fps}$$

where v is the UAV ground speed, Δt is the sampling time interval, fps is the frames per second, and GSD_{ct} and GSD_{at} are the ground sampling distance in the cross-track and along-track directions, respectively. In this approximation, GSD_{ct} relates to the ground portions of the GIFOV that contribute strongly to the value recorded by the detector. GSD_{at} relates to the length of IFOV along the ground, also known as the *dwell time*. It is important to distinguish between the UAV velocity and the ground speed because wind can considerably add to or detract from the total ground area traversed.

Adopters should perform their due diligence and consider the implications of each hardware choice. For example, wide angular FOV lenses force a higher flight altitude, which can reduce ground sampling distance. A UAV with a 3.6 km/hr cruise speed and an 11.49 frames per second pushbroom imager with 1004 spatial bands would have a GSD_{at} of 0.087 m (1.0 m/s × 0.087 second). For a 25.0403° angular FOV lens, the flight altitude would have to be ~196 m to achieve a square ground sampling distance

$$0.087 = \left(2 \times H \times \tan\left(\frac{\theta_{\mathrm{FOV}}}{2}\right)\right) / (\text{Number of spatial pixels})$$

$$\frac{0.087 \times 1004}{2 \times \tan\left(25.0403/2\right)} = H$$

For a square ground sampling distance, the relationship between the required frames per second and altitude decreases as a power function (Figure 3.5). Descending from the altitude ceiling (which is strongly variable between fixed-wing and multirotor UAVs) quickly increases the required frames per second. Descending from 30 to 20 m increases the frames per second by 1.5×. For a given UAV ground speed, the FOV is negatively correlated with the required frames per second because the longer cross-track distance is balanced by increasing the scanning interval. The opposite is true for the UAV ground speed. The frames per second are increased in order to compensate for the larger ground area covered in the scanning interval.

Sensor frame rates are adjustable, up to a maximum rate determined by the detector quality. AVIRIS and CASI offer frame rate ranges between 10 and 100, up to 333 fps, and maximum UAV fps of 25–100 are not uncommon. Other parameters are less adjustable. In particular, the speed range of a UAV can be quite limited. Fixed wings must operate at high speeds to remain aloft, and multirotors have difficulty achieving moderate speeds. In many cases, the easiest adjustable parameter is the altitude.

Achieving square pixels is beneficial because it simplifies hurdles that can occur in postprocessing. In reality, achieving the conditions necessary for square pixels

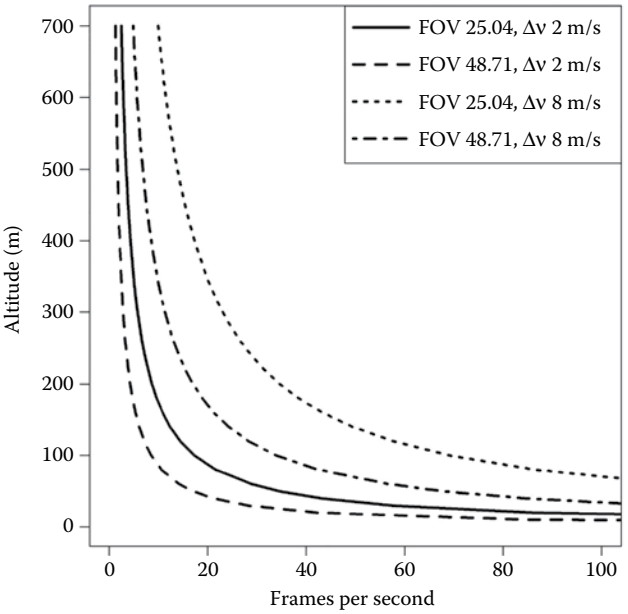


FIGURE 3.5 Required frames per second and altitude to achieve square pixels for various lens FOVs and UAV speeds.

($GSD_{at} = GSD_{ct}$) is unlikely under real flight conditions. Even under the steadiest of conditions, the UAV will experience unpredictable changes in altitude and speed (Figure 3.6). Minor variations can be corrected, assuming that the quality of the UAV's position and sensor orientation are recorded with sufficient fidelity. UAVs record navigational data via global positioning systems (GPS) and an inertial navigation system, which are complementary. Both have strengths and weaknesses. The inertial navigation system records orientation data at a frequency of ~100 Hz or higher, whereas GPS normally records at one to two orders of magnitude lower (1–10 Hz). With GPS, the positional errors are minor, but tracking high dynamic movements is difficult, resulting in leaps of position (Figure 3.6). In such cases, a spline is fit to the GPS data in order to provide a smoothed estimate of the UAV location. Conversely, the inertial navigation system is very accurate over short periods, but the number of errors increases with time. Integrating the GPS and inertial navigation system observations further reduces navigation errors. The deeper the coupling, the more accuracy can be obtained from it, although there is greater risk of outliers impacting the accuracy. Research into integrated GPS and inertial navigation systems is a quickly developing field that offers the potential to utilize cheaper sensors without loss of accuracy.

Once the orientation, location, and altitude of the sensor at the time of acquisition of each scan line have been compiled, the image undergoes orthorectification (Figure 3.7). Essentially, the centerline of the GPS positional information is used to draw a line along the ground, which is then elevated according to the flight altitude. The sensor FOV is then projected at each step along the line onto the terrain, according to the sensor orientation. Because the area scanned by the projected FOV varies with the terrain, a

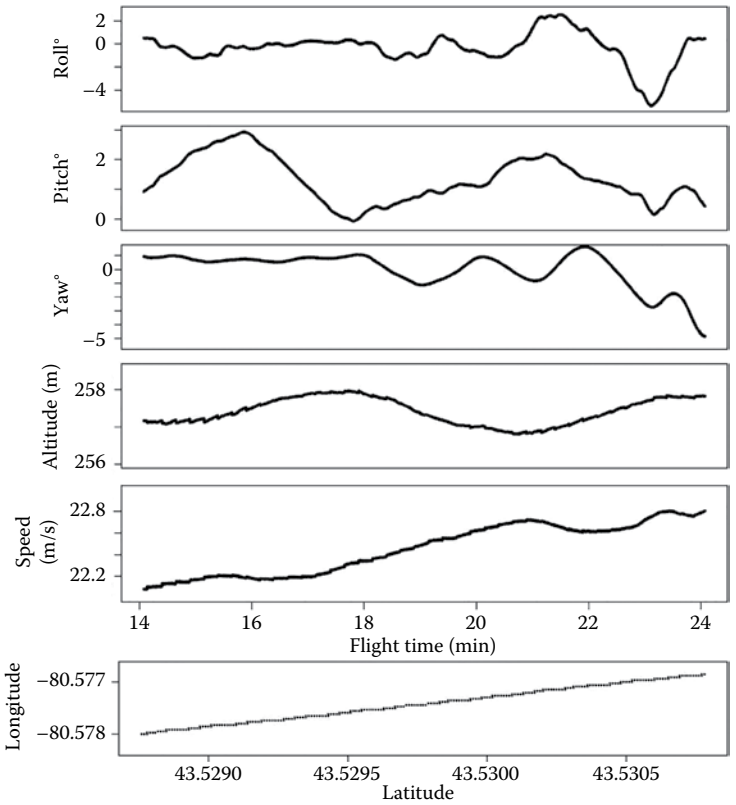


FIGURE 3.6 Inertial measurement unit-recorded positional, orientation, and flight speed data, recorded for a straight-line flight under winds less than 5 km/hr.

digital elevation model is involved. The resultant orthoimage is corrected such that the effects of the sensor orientation, camera attitude, and relief of the terrain have been removed, so that it can be matched to a map at the proper scale.

Modules for orthorectification are often available from the sensor manufacturer for a small fee. However, typically these are pushbutton solutions with minimal available options. Unlike with traditional remote sensing software, many of these modules are black box. Hence, it is difficult to determine the efficacy of the orthorectification approach and the algorithms used. Luckily, for typical cases in the agriculture and forestry sectors, their performance is more than adequate. Manufacturer software also provides tools for imagery viewing, radiance and/or reflectance conversion, and even preflight planning. Many postprocessing functionalities are also available from software, such as ENVI and RESE's PARGE.

3.3 COMPONENTS OF A HYPERSPECTRAL UAV

Although commercial off-the-shelf integrated hyperspectral UAV systems are available, it is advisable that users familiarize themselves with the hardware

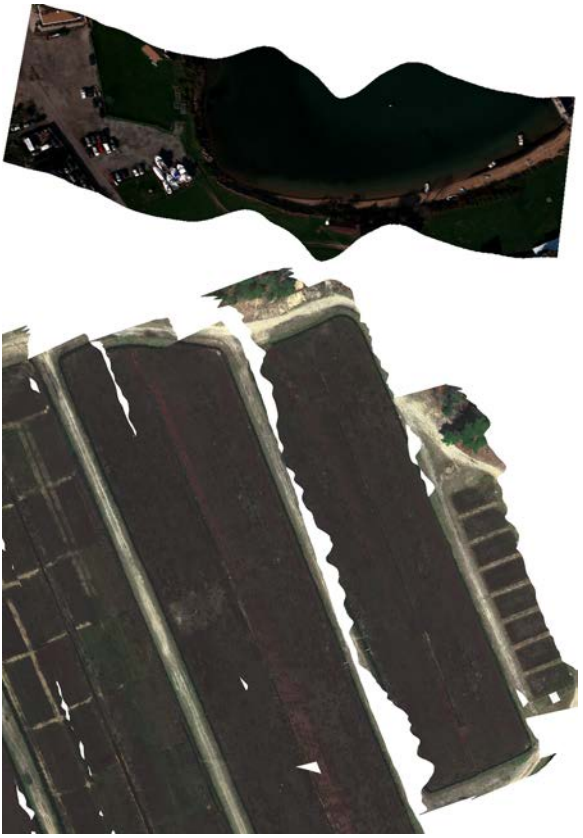


FIGURE 3.7 (Top) Orthorectified scan line of Block Island, Rhode Island; (Bottom) Orthorectified scan lines of cranberry bogs in Massachusetts. Note that scan lines do not overlap in some cases, or overlap slightly but the movement of the UAV creates gaps between scan lines.

performance and limitations before making purchasing decisions. Often, budgetary constraints will elicit trade-offs in UAV endurance, feasible scannable area, sensor performance metrics, or all of the above. Adopters should have a clear understanding of these trade-offs in relation to their system requirements and must-haves. The aforementioned obligations are heightened with custom-designed systems, which the majority of research community projects involve. Typically, some degree of customization is required, especially for a research community such as this that pushes boundaries. Hence, producing a system that meets the user's performance goals will likely require, at minimum, dealing with three parties: the sensor manufacturer, the UAV manufacturer, and a third-party engineering company providing integration services. Because hyperspectral UAVs are an atypical business case, product and service providers are unlikely to have a history of working with one another. Those undertaking custom systems at the cutting edge of commercial viability must have an appreciation for long lead times and tempered expectations.

For users, a UAV's main appeal is the automatic piloting, affording many the ability to collect imagery with a minimum of flight training. For the relevant transportation safety authorities, however, this attribute is a drawback that must be compensated for. Hence, a considerable portion of UAV components are dedicated to providing pilot information (instrumentation) and manual control capabilities. For safety reasons, a human operator must be able to take manual control at any time. However, current UAV models tend to avoid this capability and will fly into dangerous situations unawares. Operating a UAV in autonomous mode requires the onboard electronics to know the UAV's current position and motion elements (rotation and movement), and to provide proper adjustment of the craft's flight controllers in order to achieve a higher-order maneuver (e.g., hover, climb, fly straight 250 m). Recent advancements in autopilots and GPS have made these devices more cost effective. Even the actions to perform in case of communication failure are automated (e.g., auto-return to base). Often, the preferred operation of a UAV in automated mode is to preset a number of waypoints that direct the UAV over the area to be scanned, for which the autopilot devises a flight plan and executes the maneuvers necessary to pass through each waypoint. To that end, the following hardware and software components are essential for any system (Figure 3.8):

- Frame: Provides mechanical support for all components.
- Propeller: Electronic DC motor that spins a shaped blade at >1000 rpm.
- Electronic speed controller: Regulated connection of power between the battery and the motors, using pulse width modulation. Pulses of full power are provided to the motor as a square waveform. At 50% duty cycle, the motor is on for half the time.

FIGURE 3.8 Typical components of a hyperspectral UAV.

- Autopilot: Tasked with knowing the craft's position and its motion elements, and providing flight control commands. Translates commands or motion tasks into control of the rotors or flight control surfaces.
- Radio transmitter: Manual or automatic controller transmission to the craft. If additional channels are available, this can be used to send commands to the sensor controller.
- Radio receiver: Intercepts radio transmitter commands.
- Power supply: Battery and power distribution components.
- Hyperspectral payload: The hyperspectral sensor, frame grabber, data processing, and data storage components.
- Sensor lens: Often a fixed-focus lens with minimal ability to vary the focus. Determines the field of view and aperture.
- Inertial measurement unit (IMU): A device that records the movement and rotation of the platform (i.e., six degrees of freedom).
- GPS: An antenna that receives positional information from a series of orbiting satellites. Often works in conjunction with the IMU.
- Ground control station: A software application running on an on-site computer, which has bidirectional communication with the UAV. It often displays real-time data of the UAV, acts as a virtual cockpit, and is used for mission planning.
- Gimbal (optional): A device to stabilize the sensor in a fixed position, regardless of the craft's rotation.

3.3.1 Inertial Measurement Unit

An IMU is a self-contained unit containing three accelerometers, magnetometers, and/or gyroscopes, with each unit tasked with measuring force in one axis (each is placed orthogonally to the others). Measured forces are considered acceleration, which is used to determine velocity and position via dead reckoning. Due to the IMU integrating the signal over a finite time interval, it records the "average" acceleration, assuming this value occurred for the entire time frame. However, this assumption is poor and leads to accumulated errors in position, causing the IMU's assumed location to drift from the UAV's actual position. Sensor fusion of the acceleration data with positional data via GPS improves tracking accuracy. An IMU is hard-mounted to the sensor in order to record how it moves through space. Specifically, the IMU makes a best-guess estimate of the direction the sensor was pointing during the frame capture. The position of the UAV and the sensor orientation denote the ground location sampled per frame, which is essential for geocorrecting and orthorectification. Poor IMU performance leads to crooked or "wiggly" elements in the final imagery. Swapping out the low-cost, microelectromechanical system (MEMS) accelerometers and gyroscopes can be a tempting proposition because fiberoptic gyroscopes (FOGs) can offer 20× accuracy; however, base costs for a MEMS IMU are ∼$5000, and for a FOG base costs can be as high as $15,000–$50,000.

3.4 DATA PROCESSING AND STORAGE

A hyperspectral imager generates more than 1 GB of data per minute. The volume of data requires a dedicated, high-speed processing resource for fast data capture; essentially, a stripped-down mini-PC with high data transfer input (such as Camera Link or GigE Vision), depending on the sensor. Solid-state hard drives have faster read/write times, and +500 GB slim versions are readily available. The mini-PC runs the manufacturer software for data capture and camera control. Preflight workflow should involve tuning the sensor data collection parameters for the day's solar conditions. In addition, the performance of the sensor spectrally and spatially should be confirmed prior to flight. Thus, access to the data processing unit is necessary. It is of little value for the data processing unit to have a range of monitor outputs and ports for human interface devices. These elements merely consume space and add weight. A preferential solution is to remotely connect a laptop and use it to control the data processing unit.

3.5 SELECTING A UAV CRAFT

Daly (2009) of Jane's Information Group lists approximately 190 different commercial UAV models. UAVs can range considerably in size and form (Figure 3.9), and encompass hybrid platforms, such as vertical takeoff and landing craft with wings and rotors. Blimps, paragliders, kites, and other exotic aerial vehicles have all been utilized, with much success, for remote sensing purposes. Everaerts (2008) presented one of the first inventories of UAV platforms and sensors used in remote sensing, and was followed by a number of authors who have developed reviews for certain application areas.

These include environmental research (Hardin et al. 2010), wildlife research (Jones et al. 2006), forests (Dunford et al. 2009), and vegetated areas (Salamí et al. 2014). While exotic UAVs are useful for certain applications, the mainstays of the hyperspectral UAV world are multirotor and fixed-wing UAVs. The only exception is helicopters, which are extremely suitable to hyperspectral applications because they have high payload capacity. Central Michigan University recently purchased a 6-ft. helicopter and hyperspectral camera, and AeroScout offers hyperspectral imaging service with a Specim's AisaKESTREL. However, hyperspectral imagery acquisition via helicopter has seen infrequent adoption by the remote sensing community because helicopters are typically larger, less automated, and more mechanically complex, increasing upkeep costs.

Currently there is no generally accepted classification system for UAVs, despite many attempts in the defense (unmanned aircraft system access to national airspace) and civilian sectors (Herwitz et al. 2004, Watts et al. 2012). As a general rule, classification systems distinguish UAVs according to size and flight altitude. In terms of hyperspectral UAVs suitable for nonmilitary budgets, the range is fairly narrow because much of the civilian market does not offer the ideal payload-carrying capacity. Hence, for hyperspectral UAVs, a small UAV approaches the larger end of commonly sold systems. Table 2.1 provides a classification system for hyperspectral UAVs, divided into three classes: (1) low altitude, less than 5 kg; (2) low altitude,

FIGURE 3.9 Examples of fixed-wing and octocopter hyperspectral UAVs.

5–25 kg; and (3) medium altitude, >25 kg. Most current hyperspectral UAV adopters operate a UAV at Class 1, mainly due to cost. Note that multirotors typically cost less than fixed-wing designs, but offer stricter payload capacities and lower flight times.

The discussion that generates the most heated opinions is that of fixed-wing versus multirotor UAVs. Regardless of the merits and downsides offered by each UAV type, the decision is not clear cut. Only via careful examination of one's requirements, and foreknowledge of the locations one intends to survey, can a decision be made. Users with site access constraints or in heavily naturalized areas may opt for low takeoff and landing requirements in lieu of surveyable areas. Conversely, users intending to survey large tracts of land will need to invest in heavy-duty systems that will require logistical support. A relatively lightweight system, HYMSY (Suomalainen et al. 2014), consisting of a custom Specim 2.0 kg hyperspectral pushbroom system integrated into an Aerialtronics Altura AT8 v1 octocopter UAV, was able to achieve 2–10 ha of surveying per flight. Heavier-duty UAVs include those operated by Zarco-Tejada et al. (2012), who flew a Headwall Photonics Micro-Hyperspec (3 kg hyperspectral payload) on a fixed-wing Elimco E300 Viewer UAV with 1.5 hr endurance for 1000–2500 ha. The added endurance comes at a financial cost because the E300 Viewer is a Class 2a

UAV whose cost reaches the higher end. The E300 Viewer UAV has a 5 m wingspan and weighs 13.5 kg on takeoff. By weight, the E300 Viewer is outside the hand-launched fixed-wing class, requiring a specialized launcher that adds to the final cost.

Researchers can opt to start with smaller systems in order to demonstrate proof of concept or explore the frontiers of UAV applications. However, few phenomena of interest operate at spatial scales in the hundreds of meters. The ultimate goal of UAVs is to utilize the technology at traditional spatial scales, measured in square kilometers. Applications such as agriculture, forestry, mineralogy, and others seek data products at large spatial scales; hence, hyperspectral UAV adopters invested in research in these areas may require survey-grade equipment or larger platforms. At the higher end, the choice of UAV may consume an equal or greater amount of capital than the hyperspectral payload.

3.5.1 Multirotor UAVs

The most simplistic design possible, a multirotor offers high versatility. Consisting of a core nexus of navigation electronics and control circuitry, rotors are physically distanced from the nexus by power-transferring metal spars. Multirotors come in quad, hexa, and octo configurations, with one or two rotors for each arm. The rotors operate in coaxial rotation, in which the propellers operate in opposite directions. Unlike helicopters, where flight control is achieved by varying the pitch of the blades, platform motion is achieved by varying the propeller speed. Hovering is possible by applying equal thrust to all rotors, with half turning clockwise and half turning counterclockwise. Pitch and roll is achieved by applying more thrust to one rotor and less to its diametrically opposed rotor. Multirotors lack avionic elements that provide lift; hence, they rely on a constant source of power to provide the thrust necessary to counteract gravity. Consequently, endurance is considerably lower, and there are few recovery options upon power failure. Multirotors compensate for their poor power-to-lift ratio through ease of use. Launchable from virtually anywhere, multirotors do not require landing strips, and are often deployed directly on site. Hence, the higher power drain can be partially compensated for by near-immediate surveying of the area of interest.

Because rotors provide the lift, there is considerably less need to modulate the exterior of the craft to minimize wind resistance. The larger planar profile provides a greater interface for wind and turbulence, which exposes the craft to greater instability in high winds. However, the openness of multirotors provides benefits for sensor integration because space constraints are relaxed. Sensors are easy to mount beneath the central nexus, with room to spare. Often, this added space is taken advantage of with the addition of a gimbal. A gimbal physically disconnects the sensor package from the movements of the craft. If the craft were to roll due to high winds, the gimbal would sense the movement and would operate servers to offset the craft's movements in three dimensions. Outside of vibration dampening and stabilization, camera controllers allow manipulation of the camera field of view while in flight. Video production companies often utilize these systems to provide aesthetically pleasing shots; however, their use in surveying fields is less prevalent. Sturdy legs allow the craft to land without placing the weight of the multirotor on the sensor.

Sensors are placed below the center nexus in order to maintain the center of gravity through the midpoint of the unit. It is critically important that the center of gravity be through the midpoint to ensure equal weight distribution to each rotor. Unbalanced systems can place continued strain on certain rotors, which can be magnified by flight maneuvers. Since the thrust output of the rotors often approaches maximum performance, additional load runs a severe risk of burnout. Rotor failure is an ever-present risk. Advanced quadcopter systems have software to detect and correct for engine loss temporarily as the unit lugs home. Hexa- and octocopters are not immune to flight control loss by sudden rotor failure, but the extra rotors have a more well-distributed thrust profile, allowing continued flight with one rotor inoperable. The extra rotors provide more lift, at the expense of added weight and faster battery drain. Anyone with experience with batteries knows that the output drops precipitously when drained. Declining power output induces a feedback loop because the multirotor insists on greater power to compensate for declining altitude. Careful attention to battery level, warning devices, or other fail-safe mechanisms are crucial for ensuring that the multirotor does not mysteriously fall from the sky.

3.5.2 Fixed-Wing UAVs

Fixed-wing UAVs approximate the design of a traditional airplane. However, relaxing the size of the body and payload compartments permits a wide range of unique aerodynamic forms. The only commonality to all fixed-wing UAVs is a core body and two wings. Tailless platforms, delta wings, and propellers mounted at the front and back are design choices soliciting compromises in aerodynamics (fuel efficiency), payload capacity (size and weight), redundancy (multiple control surfaces), resistance to atypical operating conditions (stall and spin control), and flight control (maneuverability). Discussion of the relative merits of each design choice are beyond the scope of this chapter.

Regardless of design choices, all fixed-wing UAVs gain a substantial portion of lift from their wings. Wings are shaped into an airfoil to provide a high lift-to-drag ratio. As air flows over them, driven by the airplane's forward motion, lift is generated. Beyond the stall speed of the fixed wing, altitude climbs; hence, most fixed-wing UAVs require an initial boost of speed or runway to gather the speed necessary to generate lift. Forward propulsive force is provided by one or more propellers, typically mounted through the UAV's central axis. These propellers convert rotary motion into forward motion of the plane, thus forcing air movement over the wings. When under flight, hinged control surfaces (ailerons, elevators, flaps, brakes, and/or rudders) mounted on the wings or tail fins can shift from flush with the wing to deflect the flow of air. The force is unbalanced, and thus shifts the UAV around three axes of rotation: yaw, roll, and pitch. Forcing the elevators upward pushes the fixed-wing UAV's nose vertically; changes in the vertical orientation of the nose are denoted pitch. Yaw changes the direction the nose is pointing horizontally, and occurs by adjusting the rudder or ailerons. In roll, the direction of the nose is unchanged, but the fixed-wing UAV's body rotates. Increasing the lift on one wing and decreasing it on the other causes roll.

The fixed-wing UAV's form is conducive to remaining aloft, as long as forward thrust is provided and turning maneuvers or climbs are moderated. At steep gradients, when the wings are not perpendicular against gravity, the lift provided by the wings is insufficient, and the fixed-wing UAV descends. More troublingly, turbulence, wind, or air current can push the fixed-wing UAV beyond its recovery point, causing a stall, spin, or even loss of flight control. Hence, flight controllers are designed to only attempt maneuvers within a clear safety margin. The inability of a fixed-wing plane to adhere to linear waypoints is crucially important to note. Unlike a multirotor, which can move independently in each dimension, a fixed-wing UAV primarily moves forward and is nudged along the other two dimensions. Consider a fixed-wing UAV gaining altitude: Climbing maneuvers are moderated by the thrust limits and the pitch that can be maintained without inducing a stall. Gains in altitude are relatively minor compared to the forward movement, so the fixed-wing UAV must execute a bank in order to return home before radio reception is lost. The resultant maneuver is a spiral upward to the operating altitude.

Takeoff and landing for a fixed-wing plane are the two most dangerous maneuvers. In both cases, the area must be free of obstructions, and preferably soft. A 500 m patch of flat ground is not an unreasonable requirement for landing, depending on the UAV size class. The landing strip should also be oriented to minimize cross wind, preferably descending while flying into the reduced velocity. Unfortunately, landings are often not autonomous, but an acquired skill. Thus, the easiest controlling perspective is to view the descending fixed-wing UAV nearly head on. The operator must be capable of assessing the velocity of the fixed-wing UAV relative to its rate of descent, and make adjustments accordingly. The difficulty in finding a suitable landing strip has led to many alternative UAV landing devices, such as nets, which are often employed in off-base military operations. Other options include parachute recovery systems. When the fixed-wing UAV reaches a low altitude, a parachute is deployed from a top compartment, enabling a descent at low speed. However, the descent is not controlled, and wind will shift the landing location considerably far from the spot of deployment; hopefully not into water, trees, or other obstacles detrimental to electrical components. Wind will also not stop once the UAV hits the ground, and drag damage can be substantial. Quick recovery or a parachute detachment system are options. Both belly landings and parachute recovery entail risks. In belly landing, the underside must be protected by either structural support, a payload sunken into the interior, and/or a retractable payload shield. In parachute landing, the UAV may land in an inopportune location or on an inopportune component, and the parachute deployment may interact with the propeller or not deploy properly, providing poor support.

For takeoff, a hand launch is likely insufficient to provide enough initial velocity to surpass the stall speed, especially for platforms weighing greater than 5–10 kg. Traditionally, a bungee of a pneumatic catapult is utilized to accelerate the craft along a rail to 10–20 m/s velocity. The safety of the catapult is of grave concern due to the forces involved. Rupture of a pressure vessel or snapping of a bungee cord can send catapult material in random directions at high velocity. Hard hats and eye protection should be worn throughout the UAV flight, and nonessential personnel should be offset at a safe distance.

As alluded to earlier, fixed-wing UAVs maintain lift under loss of control and power failure. Under these circumstances, the fixed-wing plane becomes an uncontrolled glider. Only in the most fortunate of circumstances will a gliding UAV be recoverable with minimal damage. An added benefit of a low profile against the wind is flight stability. A fixed-wing UAV has superior handling of high wind conditions and turbulence on account of its minimal profile against the wind. The downside is that the interior payload bay is constricted in volume. Enabling all components to fit becomes a game of Tetris. On a practical level, the situation is even more dire, considering the need to insert and remove components for maintenance. Space is at a premium in fixed-wing UAVs. Accounting for the space requirements of all "hard" components may still yield complications when the physical locations of wiring and connectors are added.

3.6 SPECIAL FLIGHT OPERATIONS CERTIFICATE INSURANCE

Around the world, government agencies regulate travel according to the complexity and liability of the mode of transport. Air travel has some of the most stringent regulations, and UAVs are no exception. It is the responsibility of the UAV operator to be aware of and comply with their local regulations. In Canada, UAVs are regulated federally by Transport Canada, which outlines the methods and requirements to gain authorization to fly. Many government agencies distinguish between recreational UAV flight (i.e., hobbyists) and commercial UAV flight. Commercial UAV flights are defined as having financial implications; work, research, and all flights for purposes beyond the joy of flying typically fall under the realm of commercial. Furthermore, regulations may vary by UAV weight class, operator (individual vs. commercial company), intention (manufacturer research and development), or other criteria. UAV operators must prove that they meet or exceed the regulations in order to petition for authorization, which is granted case by case. If operational approval is granted, the operator is provided with a certificate documenting their successful application and outlining all conditions. The operator agrees to abide by the details outlined in their application; it is not a blanket permission. As such, applications are considerably information dense, outlining the type of UAV to be flown, its safety features, the personnel involved, their training and experience, the plan for securing the site, preparations against hazards, response to possible complications, contingency plans, avoidance of obstacles, insurance details, geographical locations, times of day, dates requested, and any other metrics relevant to safe flight operations. UAV operators not in compliance with conditions of the certificate, such as not abiding by the set flight dates or not ensuring that the UAV safety devices are in working order, will be charged with penalties for noncompliance, even in the event of a successful flight.

Although the UAV regulation regime is constantly changing, there are standard requirements that are likely to remain for the foreseeable future. Specifically, all UAV operations must be within visual line of sight at all times. The operators must be positioned at a vantage (preferably high) that offers an unrestricted view of the airspace the operator intends the UAV to operate within and beyond, in case of loss of control. Views of the UAV are not to be obstructed by trees. Furthermore, the UAV cannot be operated within a distance buffer from airports, built-up areas, controlled

airdromes, forest fire areas, moving vehicles, and other restricted locations. The UAV must operate under a certain altitude threshold and cannot exceed set speed limits. The operators must possess a certain amount of liability insurance, and the operator must have undergone training.

The regulations related to hyperspectral UAVs (nonrecreational UAVs between 2 and 25 kg in weight) were relaxed by Transport Canada in November 2014, relieving the operator of the requirement to apply for a Special Flight Operations Certificate (SFOC) if the operator can meet all 58 conditions outlined under the Exemption from Sections 602.41 and 603.66 of the Canadian Aviation Regulations. If the operator meets these conditions and submits a notification to Transport Canada prior to the commencement of operations, outlining their contact information, UAV model, description of operation, and geographical boundaries of operation, they do not need to apply for an SFOC.

3.7 POSTCOLLECTION PROCESSING TO REFLECTANCE

All sensors, whether multispectral or hyperspectral, measure incoming light as radiance. However, radiance depends on the solar output, viewing geometry, and atmospheric conditions, whereas reflectance is a property closely associated with the concentration and spectral properties of the compound(s) within the sensor's field of view. It is assumed that the sensor manufacturer has provided the radiometric calibration files necessary to convert the sensor digital number into radiance by wavelength. Without normalization against a known source, it is impossible to convert the radiance due to the considerable variations in wavelength dispersion device and detector efficiency.

In order to convert the radiance recorded by the hyperspectral UAV into reflectance data, preprocessing is required. Approaches to preprocessing radiance into reflectance fall under two categories: (1) measurement or simulation of the downwelling radiance, or (2) transformation of recorded values into a known reflectance. These approaches are the same as with the multispectral system, although the spectral resolution requirements are heightened. Interestingly, the user's ability to discern an improper radiance-to-reflectance conversion is heightened for hyperspectral data because the spectral resolution permits the shape of the full spectral curve of known objects to be resolved. The human eye-brain combination is unparalleled at distinguishing subtle irregularities, and can be quite useful when comparing reflectance conversion results against known spectra. Perhaps the simplest means of conversion is to record the spectra of a white Lambertian calibrated reference during flight, alongside pre- or postflight measurement of dark current. Dark current can easily be collected by taking an image snapshot with a lens cap on prior to flight. Reflectance can simply be determined according to the following equation

$$\text{Reflectance} = \frac{\text{Digital Number} - \text{Dark Current}}{\text{White Reference Measurement} - \text{Dark Current}}$$

Practically, the ability to record a white reference in flight is constrained by manufacturer size limits. The gold standard in calibrated white reference, Spectralon

material, costs upward of thousands of dollars per square foot. In order to record at least one pure pixel of a calibrated white reference, the area coverage must be at least 4× the ground sampling resolution. Acquiring a white reference of this size at reasonable cost is challenging, save for the narrowest, lowest-altitude FOV sensors. Ground sampling of white references prior to flight can be performed because the differences in atmospheric absorption between the surface and flight height are often minor. Care should be taken to avoid shadows or glare from surrounding objects.

Lower-cost materials can be utilized at the cost of reflectance purity (i.e., reflectance lower than 95%) and non-Lambertian behavior (Hruska et al. 2012). A variety of calibrated tarps and targets of known reflectivity less than 100% have been utilized (Li et al. 2015). For instance, the original white reference, $BaSO_4$, can be combined with flat white paint to produce a large target area of near 100% reflectance (Knighton and Bugbee 2000). An added benefit of a less-than-100% reflective material is that the sensor detection range can be focused on the range of the objects under investigation. If the disparity between the calibrated target and the objects is large, the sensor must be adjusted to detect both, as opposed to detection of subtle variations within the object of interest. Typically, the sensor integration time or aperture is adjusted such that the brightest object(s) only reaches 85% saturation of the sensor.

Alternatives to correcting against a known target are dark object subtraction, internal average relative reflectance, log residual, and flat field correction. In practice, achieving good results via these algorithms is difficult, and depends mainly on the quality of the objects within the image. Overall, these methods should be avoided in favor of empirical line calibration or radiative transfer modeling, if possible. Empirical line calibration has the benefit of user recording of spectra under the same conditions as the flight acquisition. Using a high-precision instrument, the reflectance of noncalibrated sources in the final imagery can be recorded for the purpose of aligning the imagery radiance. However, the gold standard for atmospheric correction remains radiative transfer modeling. Software packages such as ENVI offer the quick atmospheric correction (QUAC) and fast line-of-sight atmospheric analysis of spectral hypercubes (FLAASH) algorithms, which support hyperspectral sensors such as HyMap, AVIRIS, HYDICE, Hyperion, Probe-1, CASI, and AISA. FLASSH is based on MODTRAN4 calculations, which use the viewing and solar angles and the mean surface elevation of the measurement, assuming a certain model atmosphere, aerosol type, and visible range. These one-touch approaches offer excellent performance for known sensors.

The drawback of commercial-grade atmospheric correctors is that they are often limited to specific-use cases and sensors of known properties. Hyperspectral UAV sensors are so recent that they are not explicitly addressed by commercial software. Given the rigid nature of commercial software, it is challenging to customize the atmospheric correction module for custom sensors. Many modules require the input data to be pre-converted to radiance. Furthermore, the center wavelength and full width at half maximum (FWHM) must be defined in order to resample the MODTRAN solar spectrum to the wavelength sensitivities of the sensor. The algorithms must further be capable of simulating atmospheric absorption at various sensor heights. Assuming the absorption properties of the whole atmosphere would inject significant systemic bias into the final results. Ideally, the effects of sensor height would be modeled at near-meter sensitivity for UAVs.

3.8 CONCLUSIONS

Insiders tend to characterize the level of specificity provided by hyperspectral UAVs as useful, but caution that hyperspectral sensors will not replace other sensors in the future. The added costs of hyperspectral sensors, in terms of complexity of use, weight, and cost, can only be justified in a limited range of applications. Constrained by these factors, hyperspectral UAV proponents are poised to be early adopters in the foreseeable future. Greater developments are expected on the low-weight quadracopter integrated with a RGB- or near-infrared (NIR)-sensitive modified RGB camera. Commercial off-the-shelf technologies for hyperspectral UAVs will continue to develop at a slow pace, offering greater choices of hardware and even fully integrated systems targeting average users. Users with unique requirements will still need to select and integrate electronic components, and even develop software. Considering that adopters will continue to be the minority, a critical mass of enthusiasm and development may never materialize.

REFERENCES

Daly M. 2009. *Jane's Unmanned Aerial Vehicles and Targets*. Alexandria, VA: Jane's Information Group.

Dunford R, Michel K, Gagnage M, Piégay H, Trémelo M. 2009. Potential and constraints of unmanned aerial vehicle technology for the characterization of Mediterranean riparian forest. *International Journal of Remote Sensing.* 30: 4915–4935.

Everaerts J. 2008. The use of unmanned aerial vehicles (UAVs) for remote sensing and mapping. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences.* Vol. XXXVII. Part B1.

Greena R, Eastwood M, Sarture C, Chrien T, Aronsson M, Chippendale B, Faust J et al. 1998. Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (AVIRIS). *Remote Sensing of Environment.* 65(3): 227–248.

Hardin P, Hardin T. 2010. Small-scale remotely piloted vehicles in environmental research. *Geography Compass.* 4(9): 1297–1311.

Herwitz S, Johnson L, Dunagan S, Higgins R, Sullivan D, Zheng J, Lobitz B et al. 2004. Imaging from an unmanned aerial vehicle: Agricultural surveillance and decision support. *Computers and Electronics in Agriculture.* 44(1): 49–61.

Hruska R, Mitchell J, Anderson M, Glenn N. 2012. Radiometric and geometric analysis of hyperspectral imagery acquired from an unmanned aerial vehicle. *Remote Sensing.* 4: 2736–2752.

Jones G, Pearlstine L, Percival H. 2006. An assessment of small unmanned aerial vehicles for wildlife research. *Wildlife Society Bulletin.* 34(3): 750–758.

Knighton N, Bugbee B. 2000. *A Mixture of BASO4 and White Paint is a Low Cost Substitute Reflectance Standard for Spectralon*. Utah State University.

Li H, Zhang H, Zhang B, Chen Z, Yang M, Zhang Y. 2015. A method suitable for vicarious calibration of a UAV hyperspectral remote sensor. *IEEE Journal of Selected Topics in Applied Earth Observation and Remote Sensing.* 8(6): 3209–3223.

Patterson M, Brescia A. 2010. Operation of small sensor payloads on tactical sized unmanned air vehicles. *Aeronautical Journal.* 114: 427–436.

Salamí E, Barrado C, Pastor E. 2014. UAV flight experiments applied to the remote sensing of vegetated areas. *Remote Sensing.* 6(11): 11051–11081.

Suomalainen J, Anders N, Iqbal S, Roerink G, Franke J, Wenting P, Hünniger D, Bartholomeus H, Becker R, Kooistra L. 2014. A lightweight hyperspectral mapping system and photogrammetric processing chain for unmanned aerial vehicles. *Remote Sensing.* 6(11): 11013–11030.

Watts A, Ambrosia V, Hinkley E. 2012. Unmanned aircraft systems in remote sensing and scientific research: Classification and considerations of use. *Remote Sensing.* 4: 1671–1692.

Zarco-Tejada P, González-Dugo V, Berni J. 2012. Fluorescence, temperature and narrow-band indices acquired from a UAV platform for water stress detection using a micro-hyperspectral imager and a thermal camera. *Remote Sensing of Environment.* 117(15): 322–337.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

4 LiDAR and Spectral Data Integration for Coastal Wetland Assessment

Kunwar K. Singh, Lindsey Smart, and Gang Chen

CONTENTS

4.1 INTRODUCTION

Coastal environments are unique, with highly diverse natural communities that provide several benefits to both humans and wildlife. These natural communities are increasingly threatened by human population growth, amenity migration, unsustainable land-use practices, and the increasing demand for natural resources. Additional factors, such as direct inundation from sea-level rise and saltwater intrusion, compound the threats to coastal environments (Gornitz 1995; Hackney and Yelverton 1990; Pearsall and Poulter 2005; Poulter et al. 2009). Inundation from sea-level rise is anticipated to have a significant impact on coastal environments (Hauer et al. 2016; Karegar et al. 2016; Moorhead and Brinson 1995). Saltwater intrusion, another concern of sea-level rise, is the movement of saltwater into traditionally freshwater areas through both natural and artificial surface water conduits, such as streams, rivers, ditches, or canals (Ardon et al. 2013). This saltwater intrusion has resulted in a phenomenon referred to as marsh migration, which is the landward movement of salt marsh into areas that were once primarily fresh forested wetlands, leaving behind what many have come to refer to as "ghost forests" (areas of standing

dead trees within newly formed salt marshes) (Pearsall and Poulter 2005). This transition from forest to marsh is highly variable both spatially and temporally, and is dependent upon local factors such as vertical accretion, subsidence rates, and other anthropogenic stressors. There is significant uncertainty regarding the implications of these impacts on the continued persistence of coastal wetland ecosystems and their ability to continue to provide the benefits upon which human communities rely. The spatiotemporal variability of these processes makes it difficult to develop sound land-use management practices and policies (Enwright et al. 2016). This requires an adaptive land-use policy and management decisions, which depend on accurate methodologies to quantify ecosystems (Hudak et al. 2012; Singh et al. 2016). Consequently, quantifying fine-scale spatial patterns of coastal wetlands, particularly those wetlands at risk from saltwater intrusion, is crucial for understanding the continued persistence of coastal ecosystem services.

Integration of data from multiple remote sensing platforms is a potential solution to quantifying coastal wetland. Different platforms can complement one another and fill data gaps; for example, fusion of LiDAR and spectral imagery (Chen et al. 2004; Donoghue and Watt 2006; Hudak et al. 2002; Li et al. 2011; Singh et al. 2012). Spectral data capture the 2D spectral characteristics of the Earth's surface, whereas LiDAR measures the 3D arrangement of the Earth's surface in point cloud or continuous wave format; therefore, complementing the limitations of spectral data with several advantages. First, LiDAR is a source of high-resolution and highly accurate point cloud data. With a small footprint system, LiDAR has the capacity to reach a level of accuracy as high as that of conventional field surveys. Second, the intensity component of LiDAR data is useful to either pan sharpen spectral data or provide additional information for the segmentation of features in the scene (Campos-Taberner et al. 2016; Singh et al. 2010). Third, environment and atmosphere have little impact on LiDAR data, unlike spectral imagery, allowing LiDAR data collection under a wide range of environmental conditions (Jensen 2007). Despite its ability to capture vertical structural components along with the above advantages, LiDAR alone is not sufficient for estimating the overall condition of ecosystems.

This makes data integration essential for utilizing the complementary characteristics of both data sources for extraction and classification of Earth objects (Zhang and Lin 2017). Data integration has been explored in natural systems for mapping forest structural attributes, plant species distribution, and forest fuel estimation (Alonzo et al. 2014; Chust et al. 2008; Elaksher 2008; Garcia et al. 2011; Swatantran et al. 2011). In urban systems, spectral and LiDAR data fusion has been applied to land-cover mapping, 3D modeling, and building footprint extraction (Awrangjeb et al. 2010; Singh et al. 2012; Sohn and Dowman 2007). Multiple factors, such as extent, targeted application, spatial and spectral resolution, temporal match, and fusion methods, play a decisive role in the integration of spectral and LiDAR data (Zhang and Lin 2017). For example, an evaluation of trade-offs between data volume and thematic map accuracy in Singh et al. (2012)'s study suggests that a spatial resolution of 5 m for LiDAR surface models offers a balance between computational efficiency and classification accuracy for mapping land cover over large and spatially heterogeneous regions.

Numerous methods have been proposed for integrating these data sources. Applications tend to be data driven, therefore dictating the data characteristics, processing, and integration methods. For example, estimation of vegetation parameters such as tree species, height, and forest types relies on spatial and spectral resolution of spectral data and LiDAR point spacing. Since field data are collected at the plot level, multivariate and machine learning analyses are often applied to model vegetation parameters. Zhang and Xie (2012) used a neural network approach with tree-level survey data for detecting individual trees, estimating tree metrics, and identifying their species types using LiDAR with hyperspectral imagery. Pixel-based and object-based methods are often used to map land cover. Haala and Brenner (1999) applied pixel-based techniques to multispectral and LiDAR-derived digital surface models to identify buildings and trees. Gamba and Houshmand (2002) analyzed synthetic aperture radar (SAR), LiDAR, and aerial imagery for extraction of land cover. Likewise, Chen et al. (2009) employed an object-oriented classification using multispectral pan sharpened QuickBird imagery and LiDAR data over an urban area.

We integrate LiDAR and spectral data to map coastal wetland land-cover types. The main research objectives of this study were to (1) quantify the important vertical and horizontal structural metrics for different wetland compositional types, (2) assess the ability to differentiate wetland compositional types based on LiDAR and spectral remote sensing metrics using the random forest algorithm, and (3) evaluate model performance.

4.2 METHODS

4.2.1 Study Area

Goose Creek State Park is located in Beaufort County, North Carolina (Figure 4.1). It covers 6.77 km^2 just off the Pamlico Sound in North Carolina's Outer Coastal Plain. The park contains extensive marshes, inlets, and creeks on the northern side of the sound. It has been the site of timber production, commercial fishing, and small-scale subsistence farming. Lumber companies acquired extensive tracts of land along the creeks and harvested vast stands of old growth bald cypress and longleaf pine. Much of the land that is now part of the park was once clear-cut for timber. Evidence of the timber industry, such as remnants of piers and loading docks up and down the creek, remains at the park today. There are marshes at the shoreline that transition to riverine and depressional swamp forests further from the shoreline at higher elevations.

4.2.2 Data Sets Used

Ecological training and validation data were obtained from North Carolina Division of Coastal Management's (NC DCM's) refined version of the National Wetlands Inventory (NWI) data set. Wetlands were developed with a minimum polygonal mapping unit of 1 acre, using NWI data, a hydrography line data set, county soils surveys, and 30 m satellite imagery (North Carolina Department of Environmental Quality, http://deq.nc.gov/about/divisions/coastal-management). In addition to the refined NWI data set, we also used LiDAR and aerial color infrared (CIR) imagery collected by the National Agriculture Imagery Program (NAIP) for the analysis (Table 4.1).

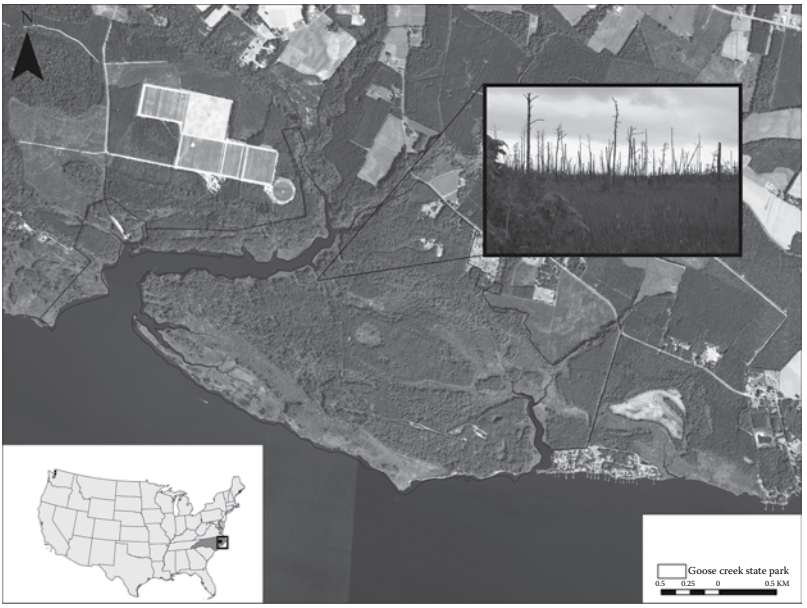


FIGURE 4.1 Study area of Goose Creek State Park with an inset image of a typical salt-affected forested wetland.

TABLE 4.1
Data Sets Used in the Analysis

Data Set	Date	Source	Derived Metrics
LiDAR	January to February 2014	North Carolina Floodplain Mapping Program	Vegetation and terrain metrics
Color infrared imagery	October 2014	NAIP, USDA Farm Service Agency	Normalized difference vegetation index
NWI data	1999–2010	North Carolina Division of Coastal Management	Ecological training and validation data

4.2.3 LiDAR Data Processing

Integrating LiDAR and spectral data requires preprocessing individual data sets to fully utilize their complimentary information. LiDAR data preprocessing includes (1) removal of artifacts (noise points), (2) classification of LiDAR points into ground and nonground, and (3) data quality check. Removal of artifacts is recommended to identify and eliminate noise points that are caused by transmission lines, atmospheric aerosols, birds, or low-flying aircraft. These points are usually outside the range of realistic elevations.

We used airborne mapping LiDAR that was acquired in LAS file format with points classified as vegetation and ground (Table 4.2) to develop the digital elevation

TABLE 4.2
LiDAR Data Specifications

Content	Specification
Survey data	January 2014
Program	North Carolina Floodplain Mapping Program
Altitude above ground	1676 m
Multiple returns	Four returns per pulse
Swath width	1025 m
$RMSE_z$[a]	11.7 cm
Average post spacing	0.7 m
Average point density	2.0/m^2

[a] Root mean square error in z.

model (DEM) and normalized heights and metrics. LiDAR data processing was completed using GRASS 7.0.1 (Neteler et al. 2012). The ground point locations were imported, and the multiscale curvature-based classification algorithm (v.lidar.mcc) was run over the ground points to remove any spurious classifications of vegetation as "ground," which can often happen in areas with dense wetland vegetation (Evans and Hudak 2007). Next, we imported the ground points using v.surf.rst at 3 m spatial resolution to create the DEM. Multiple classes with different vegetation types were merged using v.in.lidar and v.patch, respectively, to create a merged canopy point layer. We exported the merged point layer to a text file using v.out.ascii. The merged canopy file was brought back in using r.in.xyz, calculating each of the desired LiDAR metrics at a 3 m resolution. The spatial resolution was determined by selecting the finest resolution possible to detect fine-scale variability in vegetation, while also maintaining the ability to characterize vegetation canopy. A resolution of 3 m was deemed satisfactory because the number of vegetation returns for each 3 m cell ranged from 0 to 76, with an overall average of 14 vegetation returns per cell. Then, we subtracted the interpolated DEM from the height metric to normalize estimates of height (Figure 4.2).

We derived metrics representing terrain characteristics from the LiDAR data. These included flow accumulation, topographic wetness index (TWI), and distance to shoreline (Table 4.3). Flow accumulation represents the accumulated weight of all cells flowing into each downslope cell in the output raster. The elevation model is a prerequisite for flow accumulation and TWI. TWI is a function of the slope and upslope contributing area and the flow direction. The TWI is defined as

$$\ln\left(\frac{a}{\tan b}\right)$$

where a is the local upslope area draining through a certain point per unit contour length, and tan b is the local slope (in radians).

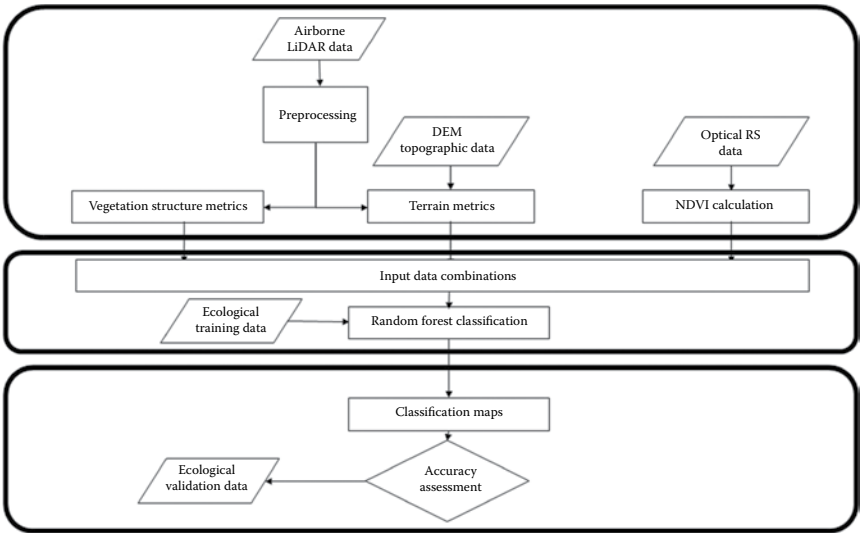


FIGURE 4.2 Overview of the data preparation and classification process.

We generated the distance to shoreline in two parts. First, we digitized the shoreline using the elevation data as reference to identify the shoreline contour. Then, a simple Euclidean distance from this linear shoreline feature was derived as the final metric. LiDAR-derived vegetation metrics were calculated by binning the laser return points at the selected cell size resolution according to the univariate statistics of interest, and were used to characterize vegetation canopy structure. The ability to detect soil moisture or potential inundation is an important abiotic factor since it controls wetland extent and function. Some research has explored the applicability of LiDAR intensity values to detect soil saturation. Therefore, we imported intensity values at the 3 m resolution. An enhanced Lee filter was applied to the raw intensity values in an effort to reduce speckle but retain edges or transition zones in the intensity data. The enhanced Lee filter was passed over the intensity values multiple times, with increasing window size each time 5, 7, and 9 number of cells (Huang et al. 2014).

4.2.4 Spectral Data Processing

Spectral data preprocessing routines are necessary to improve the spectral quality of imagery for accurate outcomes and for maximizing information through data integration. Routines include geometric and radiometric corrections, and the transformation of the image to a specific map projection system. Errors in the relative position of pixels in imagery caused by the imaging system and imaging conditions are addressed through geometric corrections. Radiometric correction addresses errors in the imagery that occur due to sensor sensitivity, sun angle and topography, and atmospheric interference (Jensen 2007). The CIR images were provided at 1 m resolution and were resampled to 3 m resolution using a bilinear cubic interpolation to match the resolution of the LiDAR data. We performed orthorectification by selecting invariant features within the study area, specifically road centerlines and parking,

TABLE 4.3
Random Forest Models, Input Variables, and Total Accuracy

Sources	Variables	Random Forest Models								
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**Final**
	Distance to shoreline			x	x	x			x	x
Spectral-data-derived metrics	Normalized difference vegetation index (NDVI)	x				x	x	x	x	x
LiDAR-derived vegetation metrics	Minimum vegetation height		x		x		x	x	x	
	Total vegetation returns		x		x		x	x	x	x
	Coefficient of variation of vegetation height		x		x		x	x	x	x
	Maximum vegetation height		x		x		x	x	x	
	Mean vegetation height		x		x		x	x	x	x
	Standard deviation of vegetation heights		x		x		x	x	x	x
	Range of vegetation heights		x		x		x	x	x	x
	Variance of vegetation heights		x		x		x	x	x	x
	Skewness of vegetation heights		x		x		x	x	x	
	25th percentile heights		x		x		x	x	x	x
	50th percentile heights		x		x		x	x	x	x
	75th percentile heights		x		x		x	x	x	x
	Mean intensity		x		x		x	x	x	x
LiDAR-derived terrain metrics	Elevation			x	x	x		x	x	x
	Flow accumulation			x	x	x		x	x	x
	Topographic wetness index			x	x	x		x	x	
	Accuracy (%)	42.1	51.3	50.7	61.2	55.4	53.3	56.4	61.8	61.9

and then using these features as tie points in both the LiDAR data and the spectral data to ensure proper alignment of the two data sources. We then used CIR images to calculate the normalized difference vegetation index (NDVI). NDVI utilizes the near-infrared spectral band and the red spectral band, important spectral bands for identifying characteristics of vegetation greenness and health. Finally, we classified different input layer combinations using the random forest algorithm.

4.2.5 Data Integration and Analysis

Classification algorithms that are frequently used to integrate these data include random forests, gradient boosting, classification and regression trees, and object-based image analysis, among others (Chen and Hay 2011; Jakubowski et al. 2013). We used the random forest algorithm to map land-cover classes of coastal wetlands (Table 4.4). The random forest algorithm is a decision tree classification and has been applied successfully in ecological research and land-cover mapping (Breiman 2001). It is especially suited for classification of multisource remote sensing data because

TABLE 4.4
Target Land-Cover Classes and Descriptions

Land-Cover Types	Descriptions
Bottomland hardwood forest	Riverine forested or occasionally scrub/shrub communities usually occurring in floodplains that are semipermanently to seasonally flooded. In bottomland hardwoods, typical species include oaks, sweet gum, river birch, and occasionally pines.
Depressional swamp forest	Very poorly drained nonriverine forested or occasionally scrub/shrub communities that are semipermanently or temporarily flooded.
Estuarine forest	A forested wetland community subject to occasional flooding by tides, including wind tides.
Estuarine scrub/shrub	Any scrub/shrub community subject to occasional flooding by tides including wind tides.
Headwater swamp	Forested systems along the upper reaches of first-order streams.
Riverine swamp forest	Riverine forested communities usually occurring in floodplains that are semipermanently to seasonally flooded. In swamp forest systems, typical species include cypress, black gum, water tupelo, and red maple.
Salt-affected swamp forest	Riverine forested communities as described above, but influenced by saltwater intrusion—either acutely or via long-term inundation.
Salt/brackish marsh	Any salt marsh or other marsh subject to regular of occasional flooding by tides, including wind tides as long as this flooding does not include hurricane or tropical storm waters.
Upland	Nonwetland areas.

it is insensitive to nonnormal and noisy data. It also fits a predetermined number of classification trees, and then combines the predictions across all trees. The majority rule across all decision trees determines the accuracy of output classification. The number of input variables is user defined. We performed data manipulation and analysis in R statistical software (R Core Team 2013). For the classification, we used the randomForest R statistical software package (Lawrence et al. 2006). The National Wetlands Inventory data set was used to identify and select training samples. A bootstrapped sample of the original training data was used to train the model. Test set accuracy was determined using cross-validation of the remaining training samples (out-of-the-bag samples). Variable importance was estimated by randomly permuting the value of out-of-the-bag samples for variables. The error increases as the particular variable is removed from the model, determining the importance of that specific variable. Samples that were not used in the training were used as test samples to perform accuracy assessment.

Preliminary exploratory data analysis was performed to initially identify important variables and avoid overfitting the model as well as to reduce collinearity among predictor variables. One such exploratory test was the Kolmogorov-Smirnov (KS) test, a nonparametric test of the equality of 1D probability distributions

(Smirnov 1948). This test can be used to compare a sample with a reference probability distribution or to compare two samples. The KS statistic quantifies a distance between the empirical distribution function of the sample and the cumulative distribution function of the reference distribution, or between the empirical distribution functions of two samples. The magnitude of the KS statistic helped in identifying a subset of variables from all the input variables. Cumulative distribution functions were drawn for these variables for each of the wetland types being modeled to visualize differences across types.

4.3 RESULTS

4.3.1 Exploratory Data Analysis

Exploratory data analysis and KS tests showed that there were several spectral and structural attributes that could be used to distinguish between wetland compositional types. The KS statistics suggested that all of the potential predictor variables were significantly different between wetland types. The cumulative distribution functions for each potential predictor variable and compositional type were evaluated for discernable relationships between variables and type (Figure 4.3). For example, the cumulative distribution functions for the distance to shoreline variable differ across

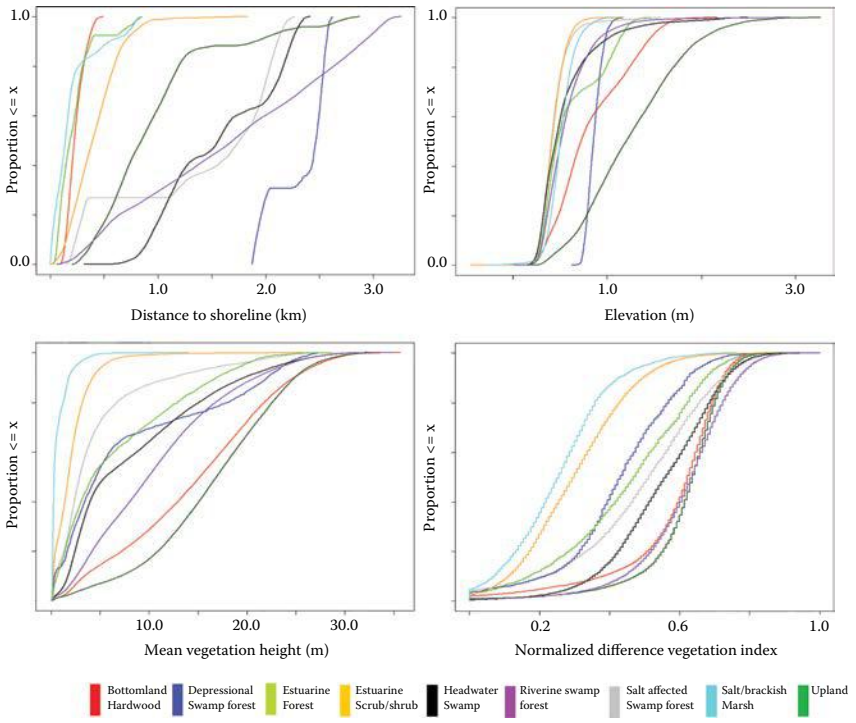


FIGURE 4.3 Cumulative distribution functions for wetland types and variables.

wetland types (Figure 4.3). The salt or brackish marsh, which is associated with areas very close to shore, is significantly different from the upland or bottomland hardwood forests, which tend to be located at a larger distance from the shore. Differences such as these are also visible in the NDVI, minimum heights, and maximum heights predictor variables.

4.3.2 Random Forest Classification

We observed significant variability in mapping error rates across wetland types. While estuarine scrub/shrub and riverine swamp forest classes had the lowest error (i.e., less than 30% error), salt-affected riverine swamp forest and estuarine forest showed very high error rates (~80% error and over) (Table 4.5). The most important variable for the classification was distance from the shoreline, followed by the terrain, LiDAR, and spectral variables. Of the LiDAR variables, the total number of vegetation returns (as a measure of vegetation density), the range in vegetation heights, and the mean vegetation heights contributed the most (Figure 4.4).

We observed the highest model performance when LiDAR, multispectral remote sensing, and terrain metrics were combined (final model), with an increase of 6% accuracy over the spectral and terrain model (i.e., Model 5) (Table 4.3). The LiDAR model alone (Model 2) performed the best of the three single data source models (Model 1, spectral imagery alone; Model 3, terrain alone). One hundred permutations of the final random forest fitted model were run to analyze the permutation importance of all predictor variables in the model (Figure 4.5). The contribution of predictor variables toward the model overall were analyzed by evaluating the mean decrease in accuracy with the removal of each variable. The predictor variable importance was also evaluated for each land-cover classification. Permutation results suggested that distance to shoreline, elevation, NDVI, and total vegetation returns were consistently the most important to overall accuracy in all permutations. This was statistically significant as represented in Figure 4.5. Comparison of the importance of the predictor variables for the highest and the lowest accuracy class reveals that

TABLE 4.5
Land-Cover Classifications, Associated Areas, and Class Errors

Vegetation Type	Area (acres [km^2])	Class Error
Bottomland hardwood forest	38.1 (0.15)	0.66
Depressional swamp forest	2.1 (0.01)	0.52
Estuarine forest	1.19 (0.01)	0.97
Estuarine scrub/shrub	162.8 (0.66)	0.28
Headwater swamp	48.24 (0.20)	0.74
Riverine swamp forest	534.93 (2.16)	0.21
Salt-affected swamp forest	1.66 (0.01)	0.76
Salt/brackish marsh	21.85 (0.088)	0.33
Upland	527.38 (2.13)	Not available

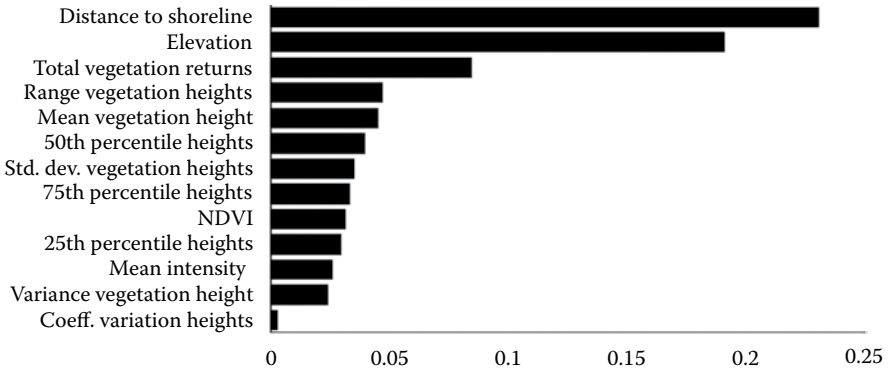


FIGURE 4.4 Variable importance for predictor variables used in the random forest classifications.

the lowest accuracy class relied on distance to shoreline and elevation along with mean intensity for prediction accuracy. This differed significantly from the highest accuracy class, whose accuracy was determined most consistently by LiDAR-derived structural metrics. Figure 4.6 shows a map of the best model.

4.4 DISCUSSION

According to the random forest outputs, the forested wetland types and salt-affected swamp forests were the most difficult to classify. Variability in species composition and spatial distribution may have caused difficulties in the classification of wetland types. Such difficulty in mapping forested wetlands has been cited in other analyses in similar coastal areas within North Carolina (Allen et al. 2013). When the classification was run on the six broader NWI wetland classifications present in the study area (e.g., estuarine shrub/scrub, palustrine forest wetland), accuracy of the most parsimonious model reached 82%. In the random forest classification of the NC DCM wetlands, the distance to shoreline and elevation variables were most important. This is understandable, since distance from the shoreline as well as elevation patterns are both important drivers of wetland extent and function. Although the LiDAR data were collected in January, during leaf-off conditions, to optimize the mapping of ground elevation, highly useful information on canopy structure can still be extracted from the nonground returns. Several studies have shown that leaf-off LiDAR data can be used to characterize vegetation canopy and structure (Sexton et al. 2009; Smart et al. 2012). Although the exact height of the canopy may be influenced by the season (due to leaf-on and leaf-off differences), the overall structure and variability remains identifiable even in leaf-off LiDAR data. Therefore, we concluded that the importance of the elevation and distance to shoreline variables were due to the inherently dynamic nature of wetlands that can appear both spectrally and structurally similar but may differ hydrologically and in their classification due to slight changes in elevation, inundation patterns, and distance to water.

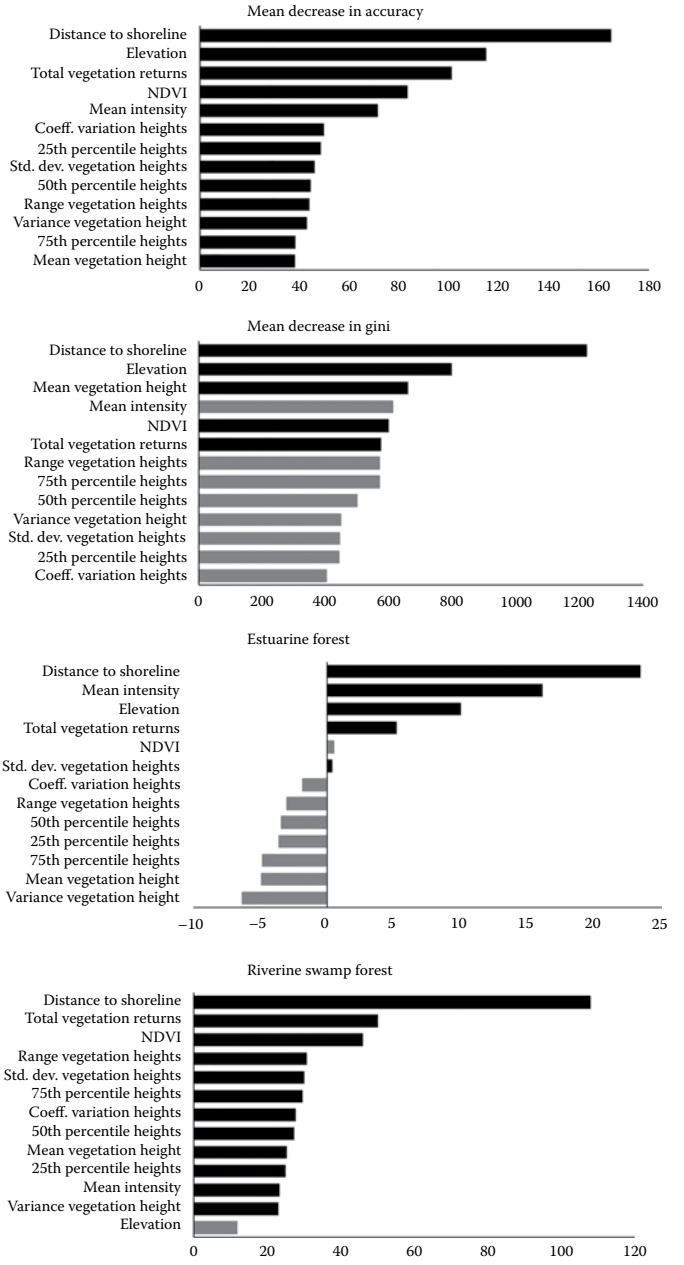


FIGURE 4.5 Permutation importance of predictor variables after 100 permutations of the random forest fitted model. Aside from the mean decrease in Gini and overall accuracy, shown in the figure are predictor variable importance for one of the highest accuracy classes (riverine swamp forest) and one of the lowest accuracy classes (estuarine forest). Black color shows statistically significant across all permutations, whereas gray signifies nonsignificant impacts on overall accuracy.

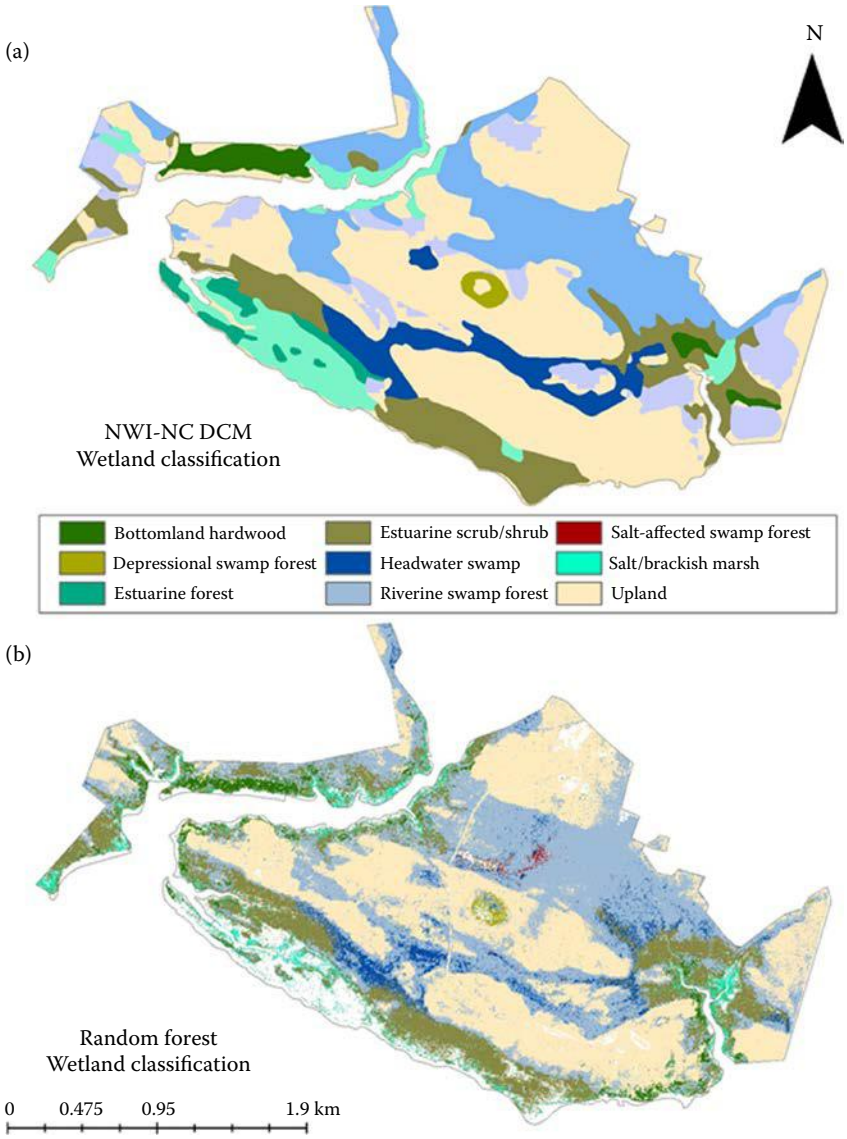


FIGURE 4.6 Costal wetland maps. (a) NC DCM—NWI wetland classification. Polygons from NC DCM wetland classification provided the target land cover for the random forest classification. (b) A map from the top-performing random forest classification model. White areas within the study area bounds represent areas of insufficient data for prediction. *Note*: Salt-affected swamp forests only appear in the random forest classification because this is not a land-cover class present in the NC DCM wetlands. Ancillary training data were made available and were used to supplement the NC DCM wetlands so that the algorithm could be used to explore the ability to identify salt-impacted wetlands.

The random forest method was selected because it is a nonparametric technique that can handle both continuous and categorical independent variables. Random forest runs efficiently on large databases, can handle many input variables without variable deletion, and maintains high predictive power even when a significant proportion of the data is missing. It works well with large and noisy remote sensing data sets. It has also been shown to reduce overfitting and achieve higher accuracies compared to traditional classification trees (Breiman 2001). The training data are the key to the performance of the random forest method. In future analyses, the training data can be improved to include acquisitions of more recent field vegetation data. The data points used to train the model in this study were drawn from polygonal data sets with a relatively coarse minimum mapping unit. The inclusion of more accurate and thorough field-based analyses is considered a source of potential improvement, and may improve overall accuracies if the model were to be trained on a different data set. In addition, accuracy may be improved with an ensemble modeling approach. Potentially, a combination of LiDAR and another active remote sensing technology, SAR, could be included. This may improve the ability to differentiate the marshes that appear spectrally similar but differ in elevation (i.e., high and low marsh). Results from initial studies combining these active remote sensing technologies prove promising (Allen 2014; Allen et al. 2013).

It was expected that the addition of LiDAR-derived vegetation metrics would increase the accuracy of the classification significantly, but results suggested it increased the accuracy by 6%. LiDAR was able to distinguish between different wetland types that had characteristic patterns of vertical structure (i.e., differentiating between marsh, scrub/shrub, and forest), particularly in cases where the types might have appeared spectrally similar. However, there are also cases wherein wetland types (i.e., estuarine forest and riverine forest) appear both spectrally and structurally similar, and the main drivers of species differentiation are related to terrain factors (i.e., distance to shoreline and elevation). Despite the relatively small contribution of LiDAR in this particular study, because of the complexity and diversity of coastal wetland composition, a multidata source approach (one that combines terrain, spectral, and vertical measures) appears to be the most appropriate.

4.5 CONCLUSION

The ability to classify and discern wetland types is necessary in order to quantify changes in wetland ecosystems through time. The vegetation extent and habitat classification maps show the utility of combining LiDAR and spectral imagery in the mapping of coastal wetlands. Analyses of the classifications indicate that elevation (along with distance to shoreline) is by far the most important variable for mapping wetland extent and function. Vegetation metrics also contributed to the overall classification accuracy. Results of this research suggest that these remote sensing variables hold promise for assessing wetland composition and differentiating between wetland types. In coastal ecosystems, further analysis is needed to understand the relationships between wetland composition and saltwater intrusion. Future research will use multitemporal LiDAR data to detect changes in wetland ecosystems resulting from saltwater intrusion using spectral data sources and field data. Results from this

analysis would provide a better understanding of the complex interactions between wetland ecosystems and sea level rise induced saltwater intrusion.

4.5.1 Opportunities and Challenges

In order to exploit all the potential offered by the new generations of LiDAR and spectral sensors to map coastal wetlands, we must address several issues. The first issue is temporal disparity. It is a rare occurrence that all data used in an analysis are collected at the same time. Temporal resolution of a satellite sensor may vary from hours to days, whereas in the case of airborne sensors it may range from days to years, depending on mission planning. Landscapes are not static—they show high spatiotemporal variability due to processes such as forest growth and urbanization. This dynamic nature can lead to erroneous results from data registration and integration. Identification of such disparities and then detrending them (e.g., removing an aspect from the data that is causing some kind of distortion) will require research. The second issue is data registration. Registration of LiDAR and spectral images data with low possible errors is necessary to harness the full potential these data offer. Multiple approaches are available for data registration; however, new generations of sensors require the development of new registration algorithms to achieve accurate registration (Zhang and Lin 2017). The third issue is the limited availability of standardized data processing approaches. Despite promising prospects, the limited availability of multisensor data and the lack of standardized data processing algorithms hamper their widespread integration in analyses. For example, standardized processing and analysis approaches for multitemporal LiDAR and return intensity would help spur research (Eitel et al. 2016) and might improve data integration for Earth and ecological sciences. The fourth issue is the increasing need for high-performance computing. Fusion studies are typically limited to small spatial extents due to LiDAR's high cost, small footprint, and large data volume. While cost will decrease, we expect increases in LiDAR data volumes with the expansion of the spectral and temporal dimensions of LiDAR data (Eitel et al. 2016). The increasingly high dimension of LiDAR and spectral data and the complexity of the processing algorithms require the design of adequate algorithms and architectures. The final issue is the development of appropriate and accurate mapping algorithms. Data fusion offers great potential to map and model environmental problems. This requires an improved understanding of the physics of LiDAR and spectral data in order to design optimal classification and mapping algorithms. With very high point spacing and resolutions, one needs an in-depth understanding of the involved physics of the sensors for accurate calibration and mapping of environmental conditions and processes.

REFERENCES

Allen, T.R. 2014. Advances in remote sensing of coastal wetlands: LiDAR, SAR, and object-oriented case studies from North Carolina. In *Remote Sensing and Modeling*, 405–428. Cham: Springer.

Allen, T.R., Wang, Y., & Gore, B. 2013. Coastal wetland mapping combining multi-date SAR and LiDAR. *Geocarto International*, 28, 616–631.

Alonzo, M., Bookhagen, B., & Roberts, D.A. 2014. Urban tree species mapping using hyperspectral and LiDAR data fusion. *Remote Sensing of Environment*, 148, 70–83.

Ardon, M., Morse, J.L., Colman, B.P., & Bernhardt, E.S. 2013. Drought-induced saltwater incursion leads to increased wetland nitrogen export. *Global Change Biology*, 19, 2976–2985

Awrangjeb, M., Ravanbakhsh, M., & Fraser, C.S. 2010. Automatic detection of residential buildings using LiDAR data and multispectral imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65, 457–467.

Breiman, L. 2001. Random forests. *Machine Learning*, 45, 5–32.

Campos-Taberner, M., Romero-Soriano, A., Gatta, C., Camps-Valls, G., Lagrange, A., Le Saux, B., Beaupere, A., Boulch, A., Chan-Hon-Tong, A., Herbin, S., Randrianarivo, H., Ferecatu, M., Shimoni, M., Moser, G., & Tuia, D. 2016. Processing of extremely high-resolution LiDAR and RGB data: Outcome of the 2015 IEEE GRSS data fusion contest—Part A: 2-D contest. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9, 5547–5559.

Chen, G., & Hay, G.J. 2011. An airborne LiDAR sampling strategy to model forest canopy height from Quickbird imagery and GEOBIA. *Remote Sensing of Environment*, 115, 1532–1542.

Chen, X.X., Vierling, L., Rowell, E., & DeFelice, T. 2004. Using LiDAR and effective LAI data to evaluate IKONOS and Landsat 7 ETM+ vegetation cover estimates in a ponderosa pine forest. *Remote Sensing of Environment*, 91, 14–26.

Chen, Y.H., Su, W., Li, J., & Sun, Z.P. 2009. Hierarchical object oriented classification using very high resolution imagery and LiDAR data over urban areas. *Advances in Space Research*, 43, 1101–1110.

Chust, G., Galparsoro, I., Borja, A., Franco, J., & Uriarte, A. 2008. Coastal and estuarine habitat mapping, using LiDAR height and intensity and multi-spectral imagery. *Estuarine Coastal and Shelf Science*, 78, 633–643.

Donoghue, D.N.M., & Watt, P.J. 2006. Using LiDAR to compare forest height estimates from IKONOS and Landsat ETM+ data in Sitka spruce plantation forests. *International Journal of Remote Sensing*, 27, 2161–2175.

Eitel, J.U.H., Höfle, B., Vierling, L.A., Abellán, A., Asner, G.P., Deems, J.S., Glennie, C.L., Joerg, P.C., LeWinter, A.L., Magney, T.S., Mandlburger, G., Morton, D.C., Müller, J., & Vierling, K.T. 2016. Beyond 3-D: The new spectrum of LiDAR applications for earth and ecological sciences. *Remote Sensing of Environment*, 186, 372–392.

Elaksher, A.F. 2008. Fusion of hyperspectral images and LiDAR-based DEMs for coastal mapping. *Optics and Lasers in Engineering*, 46, 493–498.

Enwright, N.M., Griffith, K.T., & Osland, M.J. 2016. Barriers to and opportunities for landward migration of coastal wetlands with sea-level rise. *Frontiers in Ecology and the Environment*, 14, 307–316.

Evans, J.S., & Hudak, A.T. 2007. A multiscale curvature algorithm for classifying discrete return LiDAR in forested environments. *IEEE Transactions on Geoscience and Remote Sensing*, 45, 1029–1038.

Gamba, P., & Houshmand, B. 2002. Joint analysis of SAR, LiDAR and aerial imagery for simultaneous extraction of land cover, DTM and 3D shape of buildings. *International Journal of Remote Sensing*, 23, 4439–4450.

Garcia, M., Riano, D., Chuvieco, E., Salas, J., & Danson, F.M. 2011. Multispectral and LiDAR data fusion for fuel type mapping using support vector machine and decision rules. *Remote Sensing of Environment*, 115, 1369–1379.

Gornitz, V. 1995. Sea-level rise—A review of recent past and near-future trends. *Earth Surface Processes and Landforms*, 20, 7–20.

Haala, N., & Brenner, C. 1999. Extraction of buildings and trees in urban environments. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54, 130–137.

Hackney, C.T., & Yelverton, G.F. 1990. Effects of human activities and sea level rise on wetland ecosystems in the Cape Fear River Estuary, North Carolina, USA. In *Wetland Ecology and Management: Case Studies*, 55–61. Dordrecht: Springer.

Hauer, M.E., Evans, J.M., & Mishra, D.R. 2016. Millions projected to be at risk from sea-level rise in the continental United States. *Nature Climate Change*, 6, 691–695.

Huang, C., Peng, Y., Lang, M., Yeo, I.Y., & McCarty, G. 2014. Wetland inundation mapping and change monitoring using Landsat and airborne LiDAR data. *Remote Sensing of Environment*, 141, 231–242.

Hudak, A.T., Lefsky, M.A., Cohen, W.B., & Berterretche, M. 2002. Integration of LiDAR and Landsat ETM plus data for estimating and mapping forest canopy height. *Remote Sensing of Environment*, 82, 397–416.

Hudak, A.T., Strand, E.K., Vierling, L.A., Byrne, J.C., Eitel, J.U.H., Martinuzzi, S., & Falkowski, M.J. 2012. Quantifying aboveground forest carbon pools and fluxes from repeat LiDAR surveys. *Remote Sensing of Environment*, 123, 25–40.

Jakubowski, M.K., Li, W.K., Guo, Q.H., & Kelly, M. 2013. Delineating individual trees from LiDAR data: A comparison of vector- and raster-based segmentation approaches. *Remote Sensing*, 5, 4163–4186.

Jensen, J.R. 2007. *Remote sensing of the environment: An Earth resource perspective*. (2nd ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

Karegar, M.A., Dixon, T.H., & Engelhart, S.E. 2016. Subsidence along the Atlantic Coast of North America: Insights from GPS and late Holocene relative sea level data. *Geophysical Research Letters*, 43, 3126–3133.

Lawrence, R.L., Wood, S.D., & Sheley, R.L. 2006. Mapping invasive plants using hyperspectral imagery and Breiman Cutler classifications (randomForest). *Remote Sensing of Environment*, 100, 356–362.

Li, A.M., Huang, C.Q., Sun, G.Q., Shi, H., Toney, C., Zhu, Z.L., Rollins, M.G., Goward, S.N., & Masek, J.G. 2011. Modeling the height of young forests regenerating from recent disturbances in Mississippi using Landsat and ICESat data. *Remote Sensing of Environment*, 115, 1837–1849.

Moorhead, K.K., & Brinson, M.M. 1995. Response of wetlands to rising sea level in the lower coastal plain of North Carolina. *Ecological Applications*, 5, 261–271.

Neteler, M., Bowman, M.H., Landa, M., & Metz, M. 2012. GRASS GIS: A multi-purpose open source GIS. *Environmental Modelling & Software*, 31, 124–130.

Pearsall, S., & Poulter, B. 2005. Adapting coastal lowlands to rising seas. *Principles of Conservation Biology*, 3, 366–370.

Poulter, B., Feldman, R.L., Brinson, M.M., Horton, B.P., Orbach, M.K., Pearsall, S.H., Reyes, E., Riggs, S.R., & Whitehead, J.C. 2009. Sea-level rise research and dialogue in North Carolina: Creating windows for policy change. *Ocean & Coastal Management*, 52, 147–153.

R Core Team. 2013. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Sexton, J.O., Bax, T., Siqueira, P., Swenson, J.J., & Hensley, S. 2009. A comparison of lidar, radar, and field measurements of canopy height in pine and hardwood forests of southeastern North America. *Forest Ecology and Management*, 257, 1136–1147.

Singh, K.K., Chen, G., Vogler, J.B., & Meentemeyer, R.K. 2016. When big data are too much: Effects of LiDAR returns and point density on estimation of forest biomass. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9, 3210–3218.

Singh, K.K., Vogler, J.B., Meng, Q., & Meentemeyer, R.K. 2010. Mapping Land Use Patterns in an Urbanizing Landscape Using LiDAR Intensity Data. In *Optical Remote Sensing of the Environment*, OMC2. Tucson, AZ: Optical Society of America.

Singh, K.K., Vogler, J.B., Shoemaker, D.A., & Meentemeyer, R.K. 2012. LiDAR-Landsat data fusion for large-area assessment of urban land cover: Balancing spatial resolution,

data volume and mapping accuracy. *ISPRS Journal of Photogrammetry and Remote Sensing*, 74, 110–121.

Smart, L.S., Swenson, J.J., Christensen, N.L., & Sexton, J.O. 2012. Three-dimensional characterization of pine forest type and red-cockaded woodpecker habitat by small-footprint, discrete-return lidar. *Forest Ecology and Management*, 281, 100–110.

Smirnov, N. 1948. Table for estimating the goodness of fit of empirical distributions. *Annals of Mathematical Statistics*, 19, 279–279.

Sohn, G., & Dowman, I. 2007. Data fusion of high-resolution satellite imagery and LiDAR data for automatic building extraction. *ISPRS Journal of Photogrammetry and Remote Sensing*, 62, 43–63.

Swatantran, A., Dubayah, R., Roberts, D., Hofton, M., & Blair, J.B. 2011. Mapping biomass and stress in the Sierra Nevada using LiDAR and hyperspectral data fusion. *Remote Sensing of Environment*, 115, 2917–2930.

Zhang, C.Y., & Xie, Z.X. 2012. Combining object-based texture measures with a neural network for vegetation mapping in the Everglades from hyperspectral imagery. *Remote Sensing of Environment*, 124, 310–320.

Zhang, J., & Lin, X. 2017. Advances in fusion of optical imagery and LiDAR point cloud applied to photogrammetry and remote sensing. *International Journal of Image and Data Fusion*, 8, 1–31.

5 Multiview Image Matching for 3D Earth Surface Reconstruction

Chuiqing Zeng and Jinfei Wang

CONTENTS

5.1 INTRODUCTION

Image matching is a key topic in photogrammetry and computer vision, with many applications in obtaining 3D surface information (e.g., Tsai, 1983). Given a set of images of the Earth's surface, multiview image matching in photogrammetry is used to estimate the 3D shape of the corresponding surface. The representation of the 3D Earth's surface, referred to as *3D modeling* or *3D reconstruction* in this chapter, describes the Earth's surface in various formats, including the digital elevation model (DEM), the point cloud, and the triangular irregular network (TIN). Point clouds use a collection of 3D points to describe the Earth's surface, whereas a TIN relies on a triangular network of vertices to express the surface. DEM is a raster file with elevation as the value of raster cells; thus, strictly speaking, DEM is merely a 2.5D representation of the Earth's surface. For convenience and consistency, all of the above Earth surface representation methods will be referred to herein as *3D representation*.

Image matching and 3D reconstruction is a challenging technique in photogrammetry because it is an inverse, ill-posed problem (Szeliski, 2010) from 2D images to a 3D surface. Classic stereo-image matching of two images to determine distance from a camera to an object was first discussed many decades ago (e.g., Gennory, 1980; Gruen, 1985; Marr & Poggio, 1979; Matthies et al., 1989). Multiview image matching (called *multi-image matching* in this chapter) is rooted in classic binocular stereo-image matching, but the matching employs more than two images. Multiple images generate redundant image information, which is expected to mitigate problems of classic stereo matching (Baltsavias et al., 2008), such as occlusions, distinct object discontinuities, and repetitive objects. In contrast, the redundant information also brings in challenges of integrating multiple images during the matching, and improving the matching efficiency for many images. Studies have demonstrated the qualitative and quantitative advantages of multi-image matching over a single stereo pair (e.g., Tack et al., 2012c).

Recently, enormous unmanned aerial vehicle (UAV) surveys (Ai et al., 2015; Xu et al., 2014; Zeng et al., 2017a,b) and freely available multisource data from different levels of government and organizations (e.g., City of Surrey, 2014) provide an unprecedented chance for multi-image matching. With the improvement of computation capabilities, multi-image matching has rapidly developed over the last three decades and has evolved to be more powerful, coping with different matching primitives at different scales (e.g., Okutomi & Kanade, 1993; Paparoditis et al., 2000; Toldo et al., 2013; Tsai, 1983; Zhang, 2005). Multi-image matching, however, is still abstruse due to its sophisticated mathematics, the complicated matching techniques, and the technical terminology used. It is important to clarify and standardize

multi-image matching steps in photogrammetry and to analyze the performance of various methods. As a result, the objectives of this chapter are to:

1. Review existing multi-image matching methods, summarize the core principles of multi-image matching, and propose a concise and fundamental framework to standardize and generalize multi-image matching in photogrammetry;
2. Elaborate on advanced techniques that enhance the multi-image matching steps and enrich matching strategies, and discuss other peripheral factors that affect the performance of multiview image matching; and
3. Discuss the limitations, challenges, and opportunities of current multi-image matching methods to explore the future directions of development.

This chapter is organized as follows: Section 5.2 introduces the fundamental framework of multiview image matching, from input images to output 3D representations; Section 5.3 introduces advanced techniques for multi-image matching to improve the matching accuracy and accelerate the computation speed; Section 5.4 summarizes the existing multi-image matching algorithms, analyzing their strengths and limitations; and Section 5.5 further discusses recent developments in multiview image matching and points out the possible trends in photogrammetric multiview image matching. Section 5.6 concludes this review and highlights the contribution of this work.

5.2 FUNDAMENTAL FRAMEWORK OF MULTI-IMAGE MATCHING AND 3D RECONSTRUCTION

With multiple images as input, the reconstruction of the 3D Earth's surface is a complicated process and requires a series of steps (Zeng et al., 2013). Before multi-images can be matched, preprocessing is usually required, such as radiometric correction and image enhancement, image orientation and bundle adjustment, and matching primitive (i.e., points, edges, and break lines) extraction (e.g., Sedaghat et al., 2012; Zhang & Gruen, 2006). Then image matching is used to find the corresponding matching primitives in multi-images. The relative position changes of corresponding matching primitives in images, referred to as parallax (Lillesand et al., 2008) or disparity (Szeliski, 2010), are proportional to the camera-to-ground distance (image depth). When camera position is known, the ground elevation of the matching primitive is also determined since image depth is determined. After image matching, further postprocessing generally includes point cloud to regular grid interpolation, elevation filtering and noise removal, geocoding output 3D models, and mosaic 3D models, among others.

5.2.1 Image Orientation

Image matching involves two separate coordinates. The 2D coordinates measuring the position of a pixel in the images are the image space, whereas the coordinate system of the 3D Earth's surface is the object space (Lillesand et al., 2008). Sometimes, image space is extended to an auxiliary 3D space by adding a dimension along the

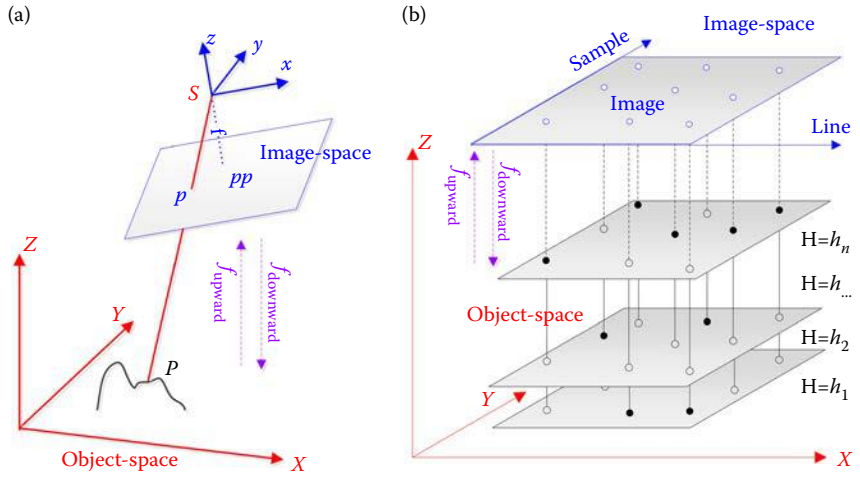


FIGURE 5.1 Two sensor model types to connect the image space and the object space. In (a), points in the image space are connected to the object space based on projective geometry. In (b), evenly distributed points in the image are mapped to a 3D grid with many elevation layers (h_1, h_2, h, …, h_n) in the object space then RFM coefficients are computed based on these points' connection from image space to object space. (From Tao, C. V., & Hu, Y. 2001. *Photogrammetric Engineering and Remote Sensing*, *67*(12), 1347–1357.)

optical axis, as in Figure 5.1a. Before multi-images can be matched, a functional relationship (Tao & Hu, 2001) between the image space and the object space must be established. This process is referred to as *image orientation*. In photogrammetry, such image-to-ground connection can be expressed using a rigorous physical sensor model or an empirical sensor model, such as the rational function model (RFM) (Di et al., 2003; Grodecki & Dial, 2003; Tao & Hu, 2001). Different types of rigorous and empirical sensor models and their performances are discussed in Habib et al. (2007).

Based on the collinearity equations in the rigorous physical sensor model, the transformation between a point in image space $p(x, y)$ and the ground point in object space $P(X, Y, Z)$ can be built, as illustrated in Figure 5.1a and described as follows (Di et al., 2003; Lillesand et al., 2008):

$$x = f_{\text{upward}}(X,Y,Z) = x_0 - f\frac{a_{11}(X-X_s)+a_{12}(Y-Y_s)+a_{13}(Z-Z_s)}{a_{31}(X-X_s)+a_{32}(Y-Y_s)+a_{33}(Z-Z_s)}, \quad (5.1)$$

$$y = f_{\text{upward}}(X,Y,Z) = y_0 - f\frac{a_{21}(X-X_s)+a_{22}(Y-Y_s)+a_{23}(Z-Z_s)}{a_{31}(X-X_s)+a_{32}(Y-Y_s)+a_{33}(Z-Z_s)}, \quad (5.2)$$

$$\mathrm{X} = f_{\text{downward}}(x,y,Z) = X_s + (Z-Z_s)\frac{a_{11}(x-x_0)+a_{21}(y-y_0)-a_{31}f}{a_{13}(x-x_0)+a_{23}(y-y_0)-a_{33}f}, \quad (5.3)$$

$$Y = f_{\text{downward}}(x,y,Z) = Y_s + (Z-Z_s)\frac{a_{12}(x-x_0)+a_{22}(y-y_0)-a_{32}f}{a_{13}(x-x_0)+a_{23}(y-y_0)-a_{33}f} \quad (5.4)$$

where $S(Xs, Ys, Zs)$ is the perspective center described in object space, f is the camera constant, pp $(x_0, y_0, -f)$ is the principal point in image space, and a_{ij} are the elements of the rotation matrix with the three angles $(\omega, \varphi, \kappa)$ (Moffitt & Mikhail, 1980). In contrast, empirical sensor models use an approximate scene-to-ground mathematical relationship without a rigorous function. RFMs (Grodecki, 2001) are well-known empirical models due to their popularity in commercial high-resolution optical satellite images. RFM uses a ratio of two polynomials to build the connection between image space and object space, as illustrated in Figure 5.1b. RFMs are fitted with the assistance of the rigorous sensor model. An RFM is described as follows (Di et al., 2003; Grodecki, 2001):

$$x = f_{\text{upward}}(X,Y,Z) = \frac{P_1(X,Y,Z)}{P_2(X,Y,Z)}, \quad y = f_{\text{upward}}(X,Y,Z) = \frac{P_3(X,Y,Z)}{P_4(X,Y,Z)}, \tag{5.5}$$

$$X = f_{\text{downward}}(x,y,Z) = \frac{P_5(x,y,Z)}{P_6(x,y,Z)}, \quad Y = f_{\text{downward}}(x,y,Z) = \frac{P_7(x,y,Z)}{P_8(x,y,Z)}, \tag{5.6}$$

$$P_l(X,Y,Z) = \sum_{m1}^{i=0} \sum_{m2}^{j=0} \sum_{m3}^{k=0} a_{ijk}^{l} X^i Y^j Z^k \tag{5.7}$$

where (x, y) are normalized pixel coordinates in image space (offset and scaled to $[-1, +1]$), and (X, Y, Z) are normalized coordinates in ground space. The polynomials P_l $(l = 1,2, \ldots, 8)$ have the form in Equation 5.7. Usually, the powers of object-space coordinates are constrained by $0 \leq m1 \leq 3$, $0 \leq m2 \leq 3$, $0 \leq m3 \leq 3$, and $m1 + m2 + m3 \leq 3$. Each $P(X,Y,Z)$ or $P(x,y,Z)$ is a third-order, 20-term polynomial.

The rigorous physical sensor model is mainly applied to frame and linear array cameras, whereas the RFMs are more generalized and can be used in many other sensor types, such as scanned line sensors. The rigorous model generally provides a more accurate result than that of the RFM, whereas the RFM is easier to implement and less computationally intensive than the rigorous model. The impacts of different sensor types on multi-image matching will be further discussed. Ground control points (GCPs) are widely used to improve the sensor model accuracy (e.g., Tack et al., 2012a; Tao & Hu, 2001; Toutin et al., 2012).

5.2.2 Candidate Search

After establishing the connection between each image and the ground, multi-image matching is used to search for the most consistent primitives that produce the highest consistency among all images under the image-ground geometry. There are two types of multi-image matching, based on the geometry in which the depth is searched: (a) image-space matching and (b) object-space matching. The image-space matching methods use the built image orientation to search for corresponding primitives among images at different viewpoints (e.g., Baltsavias, 1991; Gruen & Baltsavias, 1988; Toldo et al., 2013; Zhang, 2005). In contrast, the object-space matching methods iterate each cell in the object space and assess the mutual agreement between their

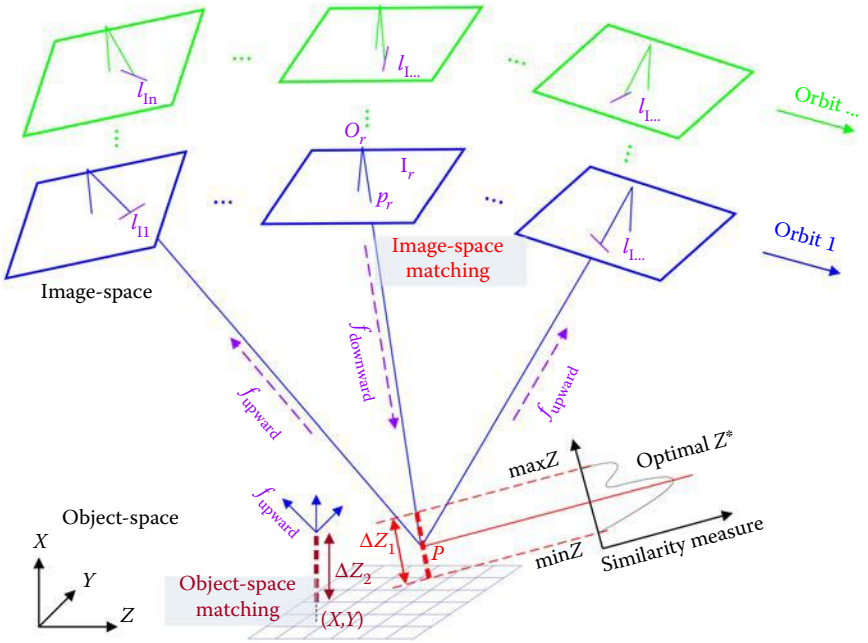


FIGURE 5.2 Core principles of the two types of multi-image matching, modified. Image-space matching searches for and matches corresponding primitives in images, whereas object-space matching estimates the elevation of each object-space cell by matching its projections in images. (From Paparoditis, N., Thom, C., & Jibrini, H. 2000. Surface reconstruction in urban areas from multiple views with aerial digital frame cameras. *Paper presented at the ISPRS Congress Amsterdam*, Amsterdam, Netherlands.)

projections on images (e.g., Noh et al., 2012; Paparoditis et al., 2000) to estimate the optimal elevation.

As illustrated in Figure 5.2, n images of a ground scene from multiple orbits can be matched in two ways. Image-space matching predefines a reference image I_r, with its imaging perspective center O_r, and a ground point $P(X, Y, Z)$ that projects in the I_r at point p_r that also needs to be matched. P is searched along the reference ray $\overrightarrow{O_r\, p_r}$ at a search range ΔZ_1. Line segments $l_{I1}, \ldots, l_{In}$ in the n images are the trajectory of possible P projections in search range ΔZ_1. The downward sensor model of I_r is used to estimate the initial position of P based on p_r and the searching Z, according to Equations 5.3, 5.4, and 5.6. Based on the initial position of P, the corresponding point in each image (except the reference image) is determined by the upward sensor model of that image, according to Equations 5.1, 5.2, and 5.5. Specifically, for the point $p_r(x_r, y_r)$ about to be matched, the corresponding point of p_r in another image i is the function of ground elevation Z:

$$p_i^Z = \begin{pmatrix} x_i \\ y_i \end{pmatrix} = f_{\text{upward}}[f_{\text{downward}}(x_r, y_r, Z)], \quad i = 1,2,3,\ldots,r-1,r+1,\ldots,n \qquad (5.8)$$

where n is the number of image views, and r is the reference image. The optimal position of P reflects the highest consistency (*similarity measure*) among corresponding points

of P in images. A similarity measure is defined among all the image corresponding points to search for the optimal elevation Z^*:

$$Z^* = \arg\max_{Z} f_s(\{p_r, p_i^Z\}), \quad Z \in \Delta Z_1 \tag{5.9}$$

where f_s is the similarity function later defined in Equation 5.12, and the optimal Z^* is the Z that maximizes the similarity measure (or minimizes the distance measure).

In contrast, object-space matching iterates all grid cells within the ground boundary and estimates the optimal elevation for each grid cell. In most cases, ground is implicitly defined as a regular mesh with certain columns and rows, known as grid cells (Szeliski, 2010). Specifically, for each grid cell, horizontal coordinates X_0 and Y_0 are predefined, while an optimal Z^* is searched for in a search range ΔZ_2. The upward sensor model is applied to each of the n images. Similarly to the image-space matching case, each Z corresponds to a point in each of the n images

$$p_i^Z = \begin{pmatrix} x_i \\ y_i \end{pmatrix} = f_{\text{upward}}(X_0, Y_0, Z), \quad i = 1,2,3\ldots,n \tag{5.10}$$

Thus, an optimal Z^* for that grid cell is searched for in a given Z range ΔZ_2, as illustrated in Figure 5.2 and defined as

$$Z^* = \arg\max_{Z} f_s(\{p_i^Z\}), \quad Z \in \Delta Z_2 \tag{5.11}$$

In image-space matching, X, Y, and Z coordinates are dependent. Specifically, a certain Z determines the X and Y using the downward sensor model, according to Equations 5.3, 5.4, and 5.6. Object-space coordinates merely work as intermediate parameters to bridge image-space coordinates from the reference image to the search images. The trajectory of the corresponding point in a search image according to the changing Z is the line l_{Ii}, called the epipolar line. In contrast, object-space matching iterates through known (X, Y) cells and is driven by the unknown Z. In object-space matching, there is no reference image; thus, all images are equally important.

5.2.3 Similarity Measures

Because matching candidates among multi-images are connected based on ground-image geometry, a similarity measure is a value or score to indicate the consistency of matching candidates among multi-images. Specifically, the similarity measure f_s in Equations 5.9 and 5.11 can be further expanded as

$$f_s(\{p_i\}) = f_s(\vec{V}_{p_1}, \vec{V}_{p_2}, \vec{V}_{p_3}, \ldots, \vec{V}_{p_m}), m \leq n \tag{5.12}$$

where p_i is one of the corresponding points, m is the number of neighboring images sharing a view of the matching point, and $\vec{V}$ is a feature vector (Zeng et al., 2014)

TABLE 5.1
Popular f_s for Two Image Blocks, Extracted and Improved

Similarity Measure Types	Measure Names	Characteristics
Parametric	Sum of the squared/absolute differences (SSD/SAD)	Assumes brightness constancy among images, sensitive to outliers
	NCC	Invariant to changes in image gain and bias, blur depth discontinuities
Nonparametric	Rank	Invariant under all radiometric changes that preserve the order of pixel intensity
	Census	
	Ordinal	
Mutual information (MI)	Joint moment method (JMM) (Tsai, 1983)	Less sensitive to noise and specularity changes than correlation method
	Hierarchical MI (HMI) (Furukawa & Hernández, 2015)	Highly invariant to transformation that does not change the amount of information in images

Source: Hirschmuller, H., & Scharstein, D. 2009. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *31*(9), 1582–1599. doi:10.1109/TPAMI.2008.221.

to express a measurable variant for point p_i. For instance, $\vec{V}$ can be a 2D matching window centered at point p_i, or it can be a 1D epipolar line segment crossing p_i. The similarity measure $f_s(\{p_r, p_i\})$ in Equation 5.9 is a special case of Equation 5.11, with one fixed (reference) image.

Regarding the definition of f_s, photo consistency is the essential and most popular similarity measure in optical image matching. There are many similarity measures listed in Table 5.1, such as pixel-based window matching and sum of squared differences, normalized cross-correlation (NCC) (Hannah, 1974), gradient-based measures (Scharstein, 1994), image variance and entropy (Zitnick et al., 2004), and hierarchical mutual information (Hirschmuller, 2008). According to Hirschmuller & Scharstein (2009), similarity measures (or cost or distance measures in the opposite manner) roughly fall into three categories: (a) parametric measures based on the magnitude of image intensities, (b) nonparametric measures based on the local order of image intensities, and (c) mutual information measures based on relationships between images. Detailed summaries of similarity measures are given in Brown et al. (2003) and Hirschmuller & Scharstein (2009).

Similarity measures used in classic binocular stereo-image matching are designed to compare two primitives (e.g., pixels, vectors, or image blocks). In Equation 5.12, multi-image similarity measure $f_s(\{p_i\})$ involves more than two vectors; thus, strategies are required to integrate similarity scores from many neighboring image pairs into a single similarity score. For instance, the most naïve strategy is choosing the highest similarity score (e.g., Moravec, 1979) from all two-image similarities, where a higher matching rate is assumed to be more reliable. Different strategies have been developed to integrate a two-vector f_s into a multivector f_s, as required in Equation 5.12 and listed in Table 5.2.

TABLE 5.2
Strategies for Converting Multiple Similarity Values to a Single Similarity Value

Strategies	Expression	Examples
Maximum	$f_s(\vec{V}_{p1}, \vec{V}_{p2}, \vec{V}_{p3}, \ldots, \vec{V}_{pm}) = \max_{i,j} f_s(\vec{V}_{pi}, \vec{V}_{pj})$, $i, j = 1,2,3,\ldots, m, i \neq j$	Winner takes all (Hernández & Schmitt, 2004)
Mean or sum	$f_s(\vec{V}_{p1}, \vec{V}_{p2}, \vec{V}_{p3}, \ldots, \vec{V}_{pm}) = \frac{1}{N} \operatorname{sum}_{i,j} f_s(\vec{V}_{pi}, \vec{V}_{pj})$, N is the number of image pairs	Sum of normalized cross-correlation (SNCC) (Zhang & Gruen, 2006)
Weighted mean	$f_s(\vec{V}_{p1}, \vec{V}_{p2}, \vec{V}_{p3}, \ldots, \vec{V}_{pm}) = \frac{1}{N} \operatorname{sum}_{i,j} W_{ij} \cdot f_s(\vec{V}_{pi}, \vec{V}_{pj})$, W_{ij} is the adjusting weight that favors local maxima	Robust voting (Vogiatzis et al., 2005)
Global optimization	f_s is determined by a combination of pixel similarity and neighboring smoothness globally; see Section 5.3.4 for details	Markov random field (Campbell et al., 2008)

Using the maximum of two-image similarity values to represent multi-image similarity is a straightforward method proven to be effective (Hernández & Schmitt, 2004), but a correct maximum similarity relies on a unique scene inside the matching window, which requires a large support window. The strategy of using the mean or sum of the two-image similarity values is reported to perform well in repetitive texture areas (Zhang & Gruen, 2006). These simple statistic methods, however, are all affected by similarity value outliers, such as similarity values at occlusion or specular reflectance surfaces. Weighted mean was designed to suppress such outlier noises by allocating weights based on local maxima (Vogiatzis et al., 2005). To completely get rid of the effect of server occlusions and failure matching values, global optimization considers the neighboring consistency to assign correct similarity values.

5.2.4 3D Reconstruction

The optimal elevation is Z^* in either image-space or object-space matching, according to Equations 5.9 and 5.11. The candidate elevation Z is increased from min Z to max Z using a constant increment, with a corresponding p_i^Z at each Z. A similarity measure profile between min Z and max Z, as illustrated in Figure 5.2, is used to decide the optimal elevation Z^*.

In an ideal case, the similarity profile shows a unimodal shape, with a clear and sharp maximum similarity value to solve Equation 5.9 or 5.11. In some cases, however, there is more than one peak for the similarity profile, and advanced strategies have been developed in the literature. In Zhang & Gruen (2006), two criteria were used to decide whether it is a successful match: a dominant peak with the local maxima a certain percentage greater than the rest of the peaks, and consistency between forward and inverse matching when different images are used as reference images. Failure to meet either of these criteria leads to multiple match candidates for a single point, or even an unsuccessful match and abandoned points. Multiple candidates for a single point can be globally optimized using the probability relaxation technique

(Christmas et al., 1995) or Markov random field (MRF) (Toldo et al., 2013). In Noh et al. (2012), the absolute peak value threshold and the threshold between the largest and second largest peaks were used to search for a successful match.

There are many data formats for representing the 3D Earth's surface, such as the digital surface model (DSM), the TIN, and 3D point clouds. In object-space matching, each grid cell with a known X and Y and an optimal Z^* after multi-image matching generates a 2.5D DSM; the iteration over all grid cells within the image boundary generates elevations as attributes for each grid. In the image-space matching case, a pixel or feature point in the reference image creates an accurate ground point after matching; iteration over all image pixels or detected feature points creates point clouds inside the image boundary. Point clouds can be further processed to produce a TIN or DSM. A TIN is usually generated after matching different primitives, such as points and edges (Wu et al., 2012). There are also other ground definition methods in multi-image matching, such as the inclined planar surface patch (Jiang, 2004) as well as the voxel in a space-sweep method (Collins, 1996) and a volume graphic cut matching method (Vogiatzis et al., 2005). Currently, DSM is still the most popular 3D Earth surface representation for satellites and aerial images acquired via narrow baseline.

Different output formats have different characteristics (Table 5.3); thus, each is suitable for different applications. From the perspective of data rendering, TIN can render a flat terrain efficiently with a relatively smaller file size, whereas point clouds and DSMs suffer from large-scale terrain rendering with larger file sizes. Moreover, terrain rendered by point clouds usually has pits and spikes. From the perspective of data manipulation, 3D point clouds can be easily split or merged, whereas within a similar TIN manipulation there is usually a topology consistency issue, and thus manifoldness (Furukawa & Hernández, 2015). From the perspective of data compatibility, DSMs representing elevation on regular grids can easily be integrated with other raster layers (e.g., land use or land cover) for analysis.

5.3 ADVANCED TECHNIQUES TO REFINE AND ACCELERATE MULTI-IMAGE MATCHING

In Section 5.2, the core framework of multi-image matching was introduced, but an effective and efficient multi-image matching method to process large-scale geographic images requires many advanced techniques. Based on the core framework, this

TABLE 5.3
Comparison of 3D Reconstruction Representation Formats

Output Format	Data Size	Rendering Speed/ Quality	Data Manipulation	Data Compatibility
DSM	Medium	Medium/fair	Normal	Good
Point clouds	Large	Slow/poor	Easy	Poor
TIN	Small	Quick/good	Difficult	Fair

section further elaborates on strategies to convert a theoretic multi-image matching framework into realistic and stable algorithms.

5.3.1 Visibility Determination and Occlusions

In Equation 5.12, the similarity is calculated among m images that share a view of the ground point in the n total images. Triple-view satellite images (e.g., Zhang & Gruen, 2006) used in multi-image matching methods usually assume that m equals n because satellite images are taken under narrow baseline and the ground scene extent in each image is approximately the same. In recent low-altitude UAV imagery matching (e.g., Ai et al., 2015; Tong et al., 2015), however, such an assumption is not satisfied because a ground point may merely be visible in a small set of images where m is smaller than n. Therefore, it is critical to search for appropriate sets of images that share the approximate ground extent; then, each set of images are separately matched and merged.

For multi-images acquired with accurate locations and camera parameters, images can be triangulated and bundle adjusted (Szeliski, 2010), and then visibility can be estimated based on the images-to-ground geometry. For images taken with inaccurate locations [i.e., low-accuracy global positioning system (GPS)] or unordered images without locations, structure from motion (SfM) (Szeliski, 2010) methods are used to simultaneously recover image orientation, and thus estimate the visibility of ground points. Meanwhile, SfM can coarsely match feature points in the multi-images, thus generating sparse 3D point clouds, which serve as preliminary input for further dense matching that generates a 3D model for the entire scene. Readers are referred to Furukawa & Hernández (2015) for more detailed algorithms about visibility search.

Some ground objects may not be visible in several or all of the multi-images, which leads to the occlusion issue. In Brown et al. (2003), three types of occlusion processing methods are discussed: methods that detect occlusion, methods that reduce sensitivity to occlusion, and methods that model the occlusion geometry. Occlusion can be detected via discontinuities in the depth map, inconsistent forward and inverse matching between images, and opposite order with adjacent matches, among other methods. Occlusions can also be lessened by modifying the similarity measure computation. For instance, the similarity measure can be adjusted to tolerate occlusion, or the search window can be resized, reshaped, or diffused to obtain the best match and minimize the effects of occlusions. Furthermore, occlusions can be modeled and included in the matching procedure, usually with dynamic programming. All occlusion mitigation techniques in the recent literature fall into these three categories. For example, in Hirschmuller (2008), occlusions are first detected by a left/right matching consistency check, and then assigned as the second lowest background value. In Noh et al. (2012), the occluded matching pairs do not contribute to the SNCC, and the occluded part of an epipolar line search window is detected and discarded in the similarity measure.

5.3.2 Matching Primitives

Matching primitives are the basic units for image matching. Common matching primitives include grid cells, point, and edge in the image space (e.g., Matthies et al., 1989), or grid cells, mesh, patch, and voxel in the object space (e.g., Baltsavias, 1991;

Collins, 1996). Different matching primitives have complementary abilities to match images with various image intensities, textures, and terrains. Recent developments in stereo-image matching (Aguilar et al., 2014; Noh et al., 2012; Stentoumis et al., 2014; Wu et al., 2012) mainly focus on the combination of different matching primitives for more effective and stable methods. This chapter discusses some of the critical matching primitives in multi-image matching, while more details about stereo matching primitives can be found in Gruen (2012).

5.3.2.1 Grid Cells

Using grid cells as matching primitives guarantees that image matching is conducted uniformly for the entire image. In Figure 5.2, object-space matching uses object-space grids as matching units, whereas image-space matching is driven by points in the reference image; such points can be pixels evenly distributed over the entire reference image. The advantage of the grid cell in image matching is to generate point clouds covering the entire image boundary, including problematic matching areas (i.e., low-texture areas), and to facilitate global optimization and hierarchical matching strategies. Conversely, the disadvantage of grid cell matching is a possible low matching coefficient, which can cause multiple corresponding candidates for one grid cell or matching failures.

5.3.2.2 Feature Points

Feature points that are detected according to image intensity gradients are not required to be evenly distributed, in contrast to grid cells. Feature points are also usually not in the center of a pixel. There are many feature point detection algorithms, such as the Förstner operator (Förstner, 1986), the Harris operator (Harris & Stephens, 1988), the SIFT operator (Lowe, 2004), and the DAISY descriptor (Tola et al., 2010). Different feature point detection methods have different characteristics; thus, they are selected by different matching methods for different purposes. For instance, a modified Förstner operator may be employed for feature point detection (Zhang, 2005) due to its reduced computational cost and easy threshold determination. The Harris operator may be chosen (Sedaghat et al., 2012) because of its repeatability rate and information content compared with other operators.

The advantage of feature points for image matching is the accuracy and reliability during matching, since feature points are usually located in texture-rich areas. Feature points have a higher possibility of successful matching, compared with grid cell points. The disadvantage of feature points is their imbalanced distribution, which may cause large areas without matched points.

5.3.2.3 Edges

Compared with grid cells and feature points, edge and line matching is more complicated. An edge often has no constant elevation, thus the matching of edges is eventually broken into matching of a point series that describes the edge (i.e., end point and midpoint). The corresponding edges in multi-images may look different due to perspective projection distortion, or broken edges because of occlusion, shadow, noises, or edge detection method deficiencies (Zhang, 2005). There are many edge detection methods, such as the Canny (Canny, 1986) and EDSION (Meer & Georgescu, 2001) operators.

Edge matching is an extension of point matching, with different strategies to define edge similarity measures. For example, in Schmid & Zisserman (2000), the mean of the NCCs for all the discrete pixels on the edges are calculated. In Zhang (2005), with a search window divided into two (left and right) parts, the consistency of these two parts' intensity and the contrast between multi-images are the criteria for edge matching. In Wu et al. (2012), a shiftable self-adaptive line cross-correlation (SSLCC) method first identified the optimal matching location in a search window, using a shiftable (moving) window, and then expanded the window size at the optimal location until it reached the maximum matching correlation.

The advantage of edge matching is to preserve elevation discontinuity along ground objects. It has also been demonstrated to be valuable for matching images with poor texture (Wu et al., 2012). The disadvantage is the difficulty of matching corresponding edges in different formats among multi-images, such as broken edges or projective distorted edges. Moreover, it is challenging to effectively integrate the matching results of edges and other point primitives into a final ground elevation product.

5.3.2.4 Patches

From the perspective of human interpretation, scenes (images) are decomposed into a series of objects. Objects can be further separated into small rectangular patches covering the surfaces. A patch is a 3D rectangle with its center and normal plane (Furukawa & Ponce, 2010), which is an extension of point primitives. To match patches, points evenly distributed on a patch are projected onto each image and then sampled on each image (Furukawa & Ponce, 2010). Similarity measures are constructed based on the group of projected points among images, similar to a 2D extension of edge matching.

Patch matching methods estimate the ground 3D surface and aspect simultaneously. Patch matching is closer to human interpretation in comparison with point primitives and is reported to better tolerate outliers or obstacles and rely less on topology assumptions and initialization (Furukawa & Ponce, 2010). A popular patch-based method, Patch-based Multi-View Stereo (PMVS) (Furukawa, 2014), is used in many studies.

5.3.3 Parameter Adjustment

During multi-image matching, there are many parameters that must be adaptively selected in order to make the matching effective, efficient, and robust. These include the reference image, the search range, the search window size, and the threshold of successful matching. Discussion of other relevant but optional parameters in stereo-pair image matching can be found in Zhang & Miller (1997).

5.3.3.1 Reference Image

In image-space matching, a reference image is selected before image matching is initiated. In most cases (e.g., Zhang, 2005), the nadir-view or near-nadir-view image is selected as a reference due to fewer occlusions. In object-space matching, although no reference image is nominated, a reference image is required for some similarity measures (i.e., SNCC), defined in Equation 5.6. In Noh et al. (2012), the reference

image changed at each ground cell: for a given grid cell, its reference image is chosen as the one that has the shortest distance between the grid cell and the image nadir position.

5.3.3.2 Search Range

Figure 5.2 and Equations 5.4 and 5.6 mentioned the search range ΔZ, from min Z to max Z, during the matching. The selection of an appropriate search range faces a dilemma between accuracy and efficiency: a short search range may fail in high terrain variation (i.e., tall building roofs or deep valleys), whereas a large search range increases the computational complexity and the chance of erroneous matching. In Paparoditis et al. (2000) and Zhang (2005), search ranges depend on *a priori* knowledge of the scene. In Jiang (2004) and Noh et al. (2012), the search range was adaptively determined by a TIN. The minimum and maximum elevations of all the matched points that comprise a point's neighboring triangular patches are employed to determine the search range of that point. Such a search range directly based on a point's neighbored TIN vertices, however, is incorrect if the current working point is a local maximum or minimum.

5.3.3.3 Search Window Size

A search window is used as the working zone to define similarity between multi-images in local-based matching methods. On one hand, the search window size should be large enough to carry sufficient intensity variation for reliable matching; on the other hand, an excessively large search window increases the computation time and projection distortion. In Kanade & Okutomi (1994), a window was selected adaptively by evaluating the local variations of the intensity and the disparity. Some studies select a fixed window size according to their study site after experiments (Paparoditis et al., 2000; Tack et al., 2012b; Takeuchi, 2005). In Zhang (2005), the search window size was experimentally set based on the change of the SNCC peak. In Gwinner et al. (2009), the size of the search window was adjusted by an epipolar constraint. Finally, in Zhu et al. (2015), the support window was based on an approximate 3D support plane, described by a depth and two per-pixel depth offsets. More discussion about various window size selection strategies can be found in Zhu et al. (2015). The search window should be adaptive at different terrain types: specifically, window size increases in flat terrain since the ground texture is dull, while window size decreases in rough and texture-rich areas to avoid terrain discontinuity and projection distortion (Tack et al., 2012c). In contrast, a matching result is necessary to evaluate how flat the terrain is in order to choose the window size to perform the matching. The relation between window size and terrain type is a chicken and egg problem.

5.3.3.4 Threshold of Successful Matching Similarity

Multiview images of the same ground scene are not identical due to projective distortion, terrain relief, and image noise; thus, a perfect similarity cannot be achieved even in the case of successful matching. A threshold that distinguishes successful and failed matches is essential. In Paparoditis et al. (2000), the authors claimed the difficulty of defining a reliability criterion and a satisfying rejection

threshold automatically; thus, a threshold of correlation scores was manually set. In Zhang (2005), a minimum acceptance threshold was set for SNCC values. In Noh et al. (2012), a combination of thresholds of the highest peak and the contrast between the highest and the second highest peak were set at the same time. In Gwinner et al. (2009), the distribution of the correlation coefficients for all of the matched interest points was employed to determine the cross-correlation coefficient threshold.

5.3.4 Global Optimization

In the core principles of multi-image matching, only single primitive matching and window-based similarity computation are discussed. The global connection between points, such as the spatial consistency, is important to refine the matching result. Imposing global consistency disambiguates multiple candidates and avoids mismatches caused by repetitive texture structures and surface discontinuities (Zhang & Gruen, 2006).

5.3.4.1 Multicandidate Optimization

If the similarity profile (illustrated in Figure 5.2) for a matching point has no clear and sharp peak, multicandidates are assigned to the point. In order to remove ambiguities in a global context, two popular optimization methods are proposed in the literature: the probability relaxation technique (Christmas et al., 1995; Zhang & Gruen, 2006; Zitnick & Webb, 1996) and MRF optimization (Campbell et al., 2008; Toldo et al., 2013). In probability relaxation, a smoothness constraint is employed to iteratively search for correct candidates. The smoothness constraint assumes that the neighboring point has similar terrain heights, except when there are matched edges between them. Given two neighboring points (i and h) in image space and their corresponding matching candidates (j and k) in object space, their compatible coefficient C is defined as follows (Zhang & Gruen, 2006):

$$C(i,j;h,k) = \frac{T}{\beta d_{ih}} \mathrm{e}^{-\left(\frac{\Delta Z_{jk}}{\beta d_{ih}}\right)^2} \tag{5.13}$$

where d_{ih} is the distance between two points i and h in image space, ΔZ_{jk} is the elevation difference for points k and j in object space, β is the scaling coefficient that is empirically set to a constant, and T is the weighting factor to control the continuity of the terrain surface. Specifically, when there is a matched edge between points i and h, then T will be set as a very small value; otherwise, T is set as 1. According to Equation 5.13, a small distance in image space (d_{ih}) or object space (ΔZ_{jk}) leads to a large compatibility coefficient, and thus a high possibility of supporting the two points as a pair of successful matches. In MRF optimization, a discrete label MRF optimization assigns a label $\bar{k}_i \in \{1,\ldots,K,U\}$ to each point p_i, with 1 to K as the multiple matching candidates of the point; U is the failed matching status, which means none of the matching candidates are correct. The cost function to be minimized consists of two parts: the smoothness term and the data term. The cost of a labelling $\bar{\mathbf{k}} = \{\bar{k}_i\}$ is defined as (Campbell et al., 2008)

$$E(\bar{\mathbf{k}}) = \sum_i f_{\text{data}}(\bar{k}_i) + \sum_{(i,j)} f_{\text{smooth}}(\bar{k}_i, \bar{k}_j), \tag{5.14}$$

$$f_{\text{data}}(k_i = x) = \begin{cases} \lambda e^{-\beta \rho_i(x)}, & x \in \{1 \ldots K\} \\ C, & x = U \end{cases}, \tag{5.15}$$

$$f_{\text{smooth}}(k_i = x, k_j = y) = 2 \frac{\left| z_{i,x} - z_{j,y} \right|}{(z_{i,x} + z_{j,y})}, x \neq U, y \neq U \tag{5.16}$$

where (i, j) denote neighboring pixels in image space, ρ_i is the peak value of NCC scores, Z is the ground elevation in object space, λ is the weight of the data term, and β and C are constant values. A large peak value (ρ_i) or small elevation difference between neighbors (i, j) leads to a lower overall energy cost $E(\bar{\mathbf{k}})$, similarly to the probability relaxation method. The data term and smoothness term can be variants (Toldo et al., 2013), as soon as they define an effectively convergent minimization issue for the candidate disambiguation.

5.3.4.2 Least-Square Matching Refinement

The goal of least-square matching (LSM) is to solve a group of nonlinear observation equations that share some common unknown parameters, to minimize the sum of the squared differences, and to give optimized common parameters. The least-square matching of stereo images can be applied to multi-image matching (Baltsavias, 1991) or multipoint surface reconstruction (Wrobel, 1991) in object space. A general least-squares observation equation is given as (Baltsavias, 1991; Zhang, 2005)

$$I_0(x, y) - e_i(x, y) = I_i(T_g(x, y)) \tag{5.17}$$

where I_0 is the intensity of a point required to be matched in the reference image, e_i is the true error function, i is the index of the matching image, and $T_g(x, y)$ is the geometric transformation between images, such as the $f_{\text{upward}} f_{\text{downward}}(x, y)$ in Equation 5.3 or a simple affine transformation. Two types of observation equations are established: radiometric constraints (i.e., the intensity consistence between images) and geometric constraints (i.e., collinearity conditions).

The LSM refinement improves the matching accuracy by further exploiting prior information and constraints; LSM claims to offer subpixel matching accuracy (Gruen & Baltsavias, 1986). However, LSM is a highly nonlinear process and therefore requires very good approximate values (Gruen, 2012); thus, primitive matching in image space or object space are prerequisites as input for LSM.

5.3.4.3 Energy Minimization

Many global methods (especially in the computer vision community) are formulated in an energy minimization framework to refine the matching result. An energy function for image matching usually includes two terms: the *data term*, which describes the successfulness of similarity for each point, and the *smoothness term*, which encodes

the smoothness assumptions and neighboring constraints. The objective of such an energy function is to find a solution for Z (surface elevations) that minimizes a global energy function (Pierrot-Deseilligny & Paparoditis, 2006)

$$E(Z) = \sum A(X,Y,Z) + w \cdot F(\vec{G}(Z)), \tag{5.18}$$

$$\text{e.g.,} \quad A(X,Y,Z) = 1 - \text{SNCC}(X,Y,Z) \tag{5.19}$$

where Z is the unknown to be optimized and A is the data term to measure the similarity among projections of point (X, Y, Z). An example data term is given in Equation 5.19. The smoothness term is $F(\vec{G}(Z))$ depending on the variations of Z, such as the 8-neighborhood elevation difference, and w is the weight to balance the data term and the smoothness term.

Different strategies have been developed to solve such a local minimum problem based on regularization and Markov random fields, including continuation, simulated annealing, highest confidence first, and mean field annealing methods (Szeliski, 2010). More efficient methods, such as maximum flow and graph cut (e.g., Ishikawa, 2003; Roy & Cox, 1998) have been recently developed. In some sense, the MRF optimization in Equation 5.14 is also an energy minimization function, although the target of Equation 5.14 is to remove ambiguity in multicandidates of points.

5.3.5 Multiscale Matching Strategies

Multiscale image matching, also referred to as hierarchical image matching, multiresolution matching, or matching through image pyramid, is an important strategy for robust, accurate, and efficient image matching. The matching results at a coarser scale are less sensitive to error sources, such as occlusions, shadows, and projective distortion. Thus, the propagation of disparity (or elevation), from the coarse image level to the fine image level, can improve the matching accuracy and decrease the computational cost. Image pyramid construction specifically for image matching was elaborated upon in Baltsavias (1991).

Multiscale image matching involves a few steps. First, based on the original images, an image pyramid is created over n levels, and given a certain scale factor (kernel size) and a downsampling method. The downsampling method is varied in literature; examples include adaptive smoothing (Chen & Medioni, 1990), mean value (Jiang, 2004), wavelet transform (Yuan & Ming, 2009), and Gaussian filter (Noh et al., 2012). Second, starting from the top level (nth level), images are matched at each intermediate level and the matching result is inherited as *a priori* knowledge in the next intermediate level. For example, a TIN constructed after matching in this level can be utilized to determine minimum and maximum search elevations in the next level (Jiang, 2004; Noh et al., 2012; Wu et al., 2012). Finally, the matching process reaches the original image level (0th level) and generates the output. Multiscale strategy is critical for multi-image matching because the computing time soars when many images are involved in the matching.

5.4 SOFTWARE, ACCURACY, AND EFFICIENCY

5.4.1 Available Multi-Image Matching Software

There are many open source or commercial multi-image matching software packages available for 3D modeling of the Earth's surface:

- Clustering Views for Multi-view Stereo (CMVS) (Furukawa, 2014), or the previous PMVS, which follows a multistep approach that does not need any initial approximation of the surface: Although mainly designed for multiview images captured by consumer cameras rather than remote sensors, CMVS is an open source matching approach that matches multi-images simultaneously rather than matching images pair by pair. There are other packages that can take the CMVS output and produce georeferenced point clouds, such as the SfM_georef (James & Robson, 2012). There are also multiple ongoing efforts to piece together existing tools for a streamlined multi-image preprocessing, matching, and product generation, such as the OpenDroneMap (GitHub, 2015).
- Ames Stereo Pipeline (ASP) (Broxton & Edwards, 2008; NASA, 2014; NeoGeographyToolkit, 2014): ASP is a free, open source suite of automated geodesy and stereogrammetry tools designed to process planetary imagery. It models the entire image formation process and finds the optimal height for each DEM post, given every available orbital image. A recent alternative, Satellite Stereo Pipeline (s2p), was developed to refine the performance on large pushbroom images (de Franchis et al., 2014).
- MicMac (IGN, 2014): MicMac is an open source multi-image matching method using graphic-cut methods. The general strategy used by MicMac is a multiresolution (image pyramid) approach. At a given resolution, MicMac minimizes an energy function, combining the data term and *a priori* knowledge about the image pattern.
- Commercial software: Many types of commercial software are capable of matching multiview imagery. For example, in large photogrammetry software, SURE (Institute for Photogrammetry, 2014) matches a reference image to a set of adjacent images. The efficiency of processing is claimed to be increased by utilizing parallel processing and hierarchical optimization. The NGATE module of SOCET SET (BAE Systems, 2015) claims to address multisensor triangulation. The ERDAS Imagine Photogrammetry (formerly Leica Photogrammetry Suite) provides flexible options (sparse, dense, and semi-global matching [SGM] matching) to match multi-images (a pairwise method). In more computer-vision-oriented standalone software packages, such as PhotoScan (Agisoft, 2014) and Pix4Dmapper (Pix4D, 2015), a complete automated workflow is implemented to process multi-images and to generate various products (e.g., point clouds, digital terrain model [DTM], DSM, reconstructed 3D scenes).

In addition, multiview image matching methods in computer vision that are designed to match general images are widely available; readers are referred to Seitz et al. (2015).

5.4.2 Accuracy of Different Multi-Image Matching Methods

Reliable point clouds derived from multi-image matching have been reported in recent studies (e.g., Remondino et al., 2014; Seitz et al., 2006; Stumpf et al., 2015) and reflect the exciting progress in this subfield. These studies report point clouds derived from multi-image matching with accuracy close to that of LiDAR; terrestrial and airborne LiDAR scans have a point cloud accuracy of about 0.01 m (Abellán et al., 2014) and 0.15 m (Jaboyedoff et al., 2012), respectively. Table 5.4 is a list of selected recent studies on assessing multi-image matching accuracy.

Besides the qualitative and quantitative measures used in Table 5.4, program running time is another frequently used criterion for multi-image matching evaluation. But computing time relies on computer configuration, image size, and scene complexity; thus, it is inappropriate for cross-study comparison. There are newly developed strategies to address issues with the aforementioned methods, from

TABLE 5.4
List of Selected Accuracy Studies on Multi-Image Matching

Sources	Data Sets	Matching Methods	Accuracy Measure	
Seitz et al. (2006); Furukawa and Ponce (2010)	Two terrestrial data sets: 317 and 363 images	PMVS	Accuracy: 90% of the reconstruction is within 0.42–0.82 mm of the ground truth mesh (GTM)	Completeness: 94.3%–98.8% of the points on the GTM are within 1.25 mm of the reconstruction
Remondino et al. (2014)	Five terrestrial data sets: 6–12 images	SURE MicMac PMVS PhotoScan	Distance deviations from reconstructed object to GTM or the known shape are less than 0.75 mm	Qualitative analysis (profiles or cross sections): Compare the reconstructed shape with the known surface shape
Remondino et al. (2014)	Three aerial data sets: 6–10 images	SURE MicMac PMVS PhotoScan	Qualitative analysis (profiles): Compare the reconstructed shape with the known surface shape, such as roof shape	
Stumpf et al. (2015)	Three aerial data sets: 88–168 images	MicMac PMVS	Mean absolute error (MAE) between reconstructed points and LiDAR points: from 0.016 to 0.033 m	Root-mean-square error (RMSE) between reconstructed points and LiDAR points: from 0.027 to 0.056 m
Zhu et al. (2015)	Four terrestrial data sets: 16–48 images	PatchMatch+ MPGC	Accuracy: 90% of the reconstruction is within 0.38–0.48 mm of the GTM	Completeness: 95.4%–99.2% the points on the GTM are within 1.25 mm of the reconstruction

the perspectives of initial value sensitivity, convergence speed, and computational memory requirements (e.g., Li et al., 2015).

Regarding the multiview matching method performance, Stumpf et al. (2015) reported that MicMac provides higher accuracy, better coverage, and higher point density than that of PMVS. But the authors are inclined to agree that no ranking pertaining to accuracy is provided for current multiview matching methods, as stated in Remondino et al. (2014). This is because multi-image matching methods are still being improved and developed with new techniques; furthermore, the level of manual intervention, such as tie point selection, search window size adjustment, and successful matching correlation threshold setting, affects the performance comparison.

5.4.3 Efficiency and Robustness

Multi-image matching exploits redundant image information to improve matching accuracy; however, the computation costs are multiplied at the same time. Matching efficiency is critical for real-time matching applications (e.g., robot-aided navigation) and image matching for large areas (Ginzler & Hobi, 2015). Apart from graphic hardware acceleration, such as programmable graphics processing units (GPUs) for fast stereo-image matching (e.g., Vineet & Narayanan, 2008), software optimization is another feasible method to improve the performance of multi-image matching.

The nature of multiview image matching is suitable for parallel processing. For example, in methods that treat multi-images as stereo pairs, multiple stereo pairs can be processed simultaneously. Large-scale images can also be split into tiles first and processed in parallel for the same purpose. Moreover, many studies develop strategies to reduce computational complexity without sacrificing disparity accuracy. For instance, semiglobal matching performs a fast approximation by replacing the 2D energy function with the summary of 8- or 16-directional 1D paths to accelerate the matching but largely preserve the accuracy (Hirschmuller, 2008). In Rogmans et al. (2009), a flexible GPU-accelerated software model was developed to implement recent real-time algorithms focusing on local cost aggregation and image warping modules. In Yu et al. (2010), new algorithms and code optimization techniques were proposed to effectively map massive parallel platforms. In Arranz et al. (2012), new data and smoothing terms were defined, including a new parameter R in the model, in order to reduce the computational complexity of the stereo matching problem while maintaining the resolution in the disparity domain.

5.5 BROADER DISCUSSION AND NEW TRENDS

5.5.1 Frame Sensors versus Linear Array Sensors

Image sensors have developed from frame cameras in their early stages to recent scanline sensors on commercial satellites, especially high-resolution optical sensors. For a frame image (also referred to as a perspective image or a pinhole image), a rigorous sensor model is applied, and the possible corresponding points for a point in the reference image fall along a straight epipolar line in another search image. However, this is not the case for linear array sensors (also mentioned as

pushbroom sensors) because the imaging center changes for each scan line, which leads to curvilinear epipolar curves (e.g., Kim, 2000). Therefore, the classic epipolar constraint and stereorectification methods (Szeliski, 2010) designed for frame images cannot be directly applied to linear array images.

Epipolar geometry is critical for efficient and accurate multi-image matching, especially in large areas. As pointed out in Franchis et al. (2014), three different strategies are used to cope with nonstraight epipolar curves during image matching. Some studies propose to perform stereo matching by directly following the nonstraight epipolar curves (e.g., Hirschmuller, 2008; Hirschmüller et al., 2005). Other studies approximate the linear array sensor with an affine camera model (e.g., Morgan et al., 2006; Wang et al., 2011) or low-order polynomials (Zhang, 2005). Finally, epipolar image resampling can also be established by reassigning the generated conjugate epipolar curve pair points to satisfy the epipolar resampling image condition (e.g., Oh et al., 2010).

5.5.2 Multisensor and Multisource Image Matching

Images accumulated over time at the same area provide alternative image sources for matching between multitemporal, multisensor, multiresolution images with illumination, rotation, and scale differences. Multisource image matching takes advantage of using existing images to decrease project costs. However, it is also challenging because many implicit assumptions made for current image matching become invalid. For instance, the atmospheric and illumination conditions for along-track stereo-image triplets are almost identical. These images also have the same or very similar spatial resolution, and the same vegetation growth stage on the ground (Baltsavias et al., 2008), without major ground object change. Multisource image matching lacks consistency among images; thus, advanced matching methods (Jeong et al., 2015) are required for images acquired from a broad range of sensors.

In multisource image matching, image intensity is no longer a dominating matching feature, unless matching images can be adjusted toward similar spatial resolution and intensity distribution. Conversely, feature matching is usually more stable than intensity matching (Gruen, 2012) if images are obtained under different illumination and atmospheric conditions. A series of existing matching costs were evaluated in Hirschmuller and Scharstein (2009). The Census Transform (Zabih & Woodfill, 1994), which converts each pixel inside a moving window into a bit vector representing which neighbors are above or below the central pixel, was found to be robust against large-scale, nonstationary exposure and illumination changes. In Tack et al. (2012b) and Aguilar et al. (2014), cross-sensor optical images were matched after radiometric and geometric normalization of the multitemporal and multisensor imagery. In Sedaghat et al. (2012), multisource optical imagery was matched via a combination of techniques, including the SIFT algorithm, the Harris corner detector, the LSM approach, a robust control network construction technique, and a new target matching strategy based on distance and angle differences.

Image matching among *optical* and *microwave* radar images is more complicated than that among multisensor optical images. Theoretically, similarity measures exist between microwave and optical images because both images are projections of the

same ground scene and share certain common properties. Practically, optical and microwave images reveal different characteristics due to different image acquisition techniques. It is challenging to build a similarity measure similar to the ones applied to optical image matching because large geometric distortions and the displacement of ground features exist between optical and microwave images (OTB Development Team, 2014). For instance, in layover areas of synthetic-aperture radar (SAR) imagery, the order of ground objects is inverted and opposite to that in optical imagery; thus, window-based methods will fail to match optical and radar imagery because such methods rely on ground object consistency. Further effort is required to design effective and robust similarity measures between microwave and optical images.

5.5.3 Challenges, Limitations, and Opportunities

The progress of multi-image matching in recent years has included the integration of correlation algorithms, consistency measures, visibility models, shape knowledge, constraints, and minimization approaches in a multistep procedure (Remondino et al., 2014). Image matching, however, is still and will continue to be an imperfect solution because of its mathematic foundation (ill-posed problem) and inevitably failed ground objects (e.g., occlusions, transparent, and moving ground objects). There are many challenges and limitations, but also opportunities that accompany the development of multi-image matching at the current stage.

5.5.3.1 The Big Picture

Multi-image matching methods are still immature, and methods that modify and refine existing matching methods will be in the mainstream in the coming decade. Multi-image matching is a complicated variant of stereo-image matching; hence, stereo-image matching techniques will directly influence multi-image matching. New feature extraction algorithms, image matching methods, and optimization strategies will be designed. Also, the development of multi-image matching in computer vision is stimulated by easily accessible consumer cameras and massive images; technical breakthroughs of multi-image matching in computer vision will transfer to photogrammetry and apply to 3D reconstruction of the Earth's surface.

5.5.3.2 Computation Efficiency

Computing time is still a bottleneck for current multi-image matching. Less than 50 images of 640 by 480 pixel size usually requires at least half an hour to run matching and reconstruction (e.g., Furukawa & Ponce, 2010; Li et al., 2015). Dense image matching is so time consuming that it is currently impossible to apply in real-time applications such as virtual environment navigation. The development of computer hardware, such as the popular use of GPUs, may mitigate this issue. Improving computational efficiency also influences the development direction of multi-image matching methods.

5.5.3.3 Accuracy Assessment

Evaluating the multi-image matching results is critical for further method comparison and improvement. The difficulty of evaluating 3D products (e.g., point cloud, mesh)

after multi-image matching is twofold. On one hand, it is difficult to collect reliable ground truth in a 3D environment. A LiDAR point cloud can be dense and accurate but the LiDAR equipment is still cumbersome and expensive, especially for aerial surveys. Manual digitalization in a virtual 3D environment is labor intensive and time consuming. The 3D reference data built from construction maps are constrained by map availability and are not feasible in natural landscapes. The management of 3D reference data and the conversion between various data formats is another concern when highly accurate reference data are required. On the other hand, effective and standardized accuracy measures are lacking in a 3D environment. The evaluation is more complicated and computation heavy in 3D space than in 2D space (Zeng et al., 2014). Existing studies used various qualitative (visual effects, profile, or cross section) and quantitative (distance, point density) measures to evaluate accuracy. Concise, effective, standardized, and easy-to-compute measures are still necessary to assess not only the accuracy of the point cloud itself, but also its capability to maintain a ground object's geometric shape and topology.

5.5.3.4 Multisource Image Matching

The combination of multi-image matching and LiDAR scanning, especially in terrestrial surveys, is a promising field for accurate mapping and reconstruction. LiDAR point clouds provide ground control points and assistance for multi-image matching, whereas multi-image matching provides extra spectral (color) information and generates even denser point clouds than LiDAR (theoretically, 100 points/m^2 for 10 cm resolution images after matching). Such a combination may happen at the hardware level, using hybrid sensors to exploit the advantages. Multisource and cross-sensor image matching is another possible development trend for multi-image matching that will admittedly be challenging. The cooperation of image information from both optical and microwave sensors will vitalize multi-image matching, with a fundamental change in similarity measures. Current intensity-based similarity measures will be replaced by more generalized measures. Multisource image matching also stimulates new applications because of its ability to reuse multitemporal, multiresolution, and multisensor images.

5.6 CONCLUSIONS

In summary, this chapter has systematically elaborated on the fundamental issues regarding multi-image matching. Based on the newest progress in photogrammetry, a clear framework for the fundamental concepts of multi-image matching has been described. Multi-image matching from both image space and object space has been discussed in the context of modern sensor models. A series of equations related to the matching were listed to standardize the concepts and procedures.

Critical techniques for image orientation, matching candidate searching, similarity measures, and 3D surface representation were summarized. Based on the core principles, further discussion was provided on refining and accelerating the matching process, including matching parameter selection, matching primitives, global optimization, and occlusion processing. Finally, other issues relating to multi-image matching were discussed, covering the impact of the difference of frame sensors

and linear array sensor on image matching, the matching steps for multisource and cross-sensor images, the image matching differences between photogrammetry and computer vision, the matching efficiency and optimization for hardware and software, and currently available multi-image matching software. The future of multi-image matching will continue to enhance the current methods, improve computing efficiency, develop efficient and standardized accuracy assessment methods, and explore the possibilities of multisensor image matching.

ACKNOWLEDGMENTS

The work was partially funded by a Natural Sciences and Engineering Research Council (NSERC) Discovery Grant award to Dr. Jinfei Wang. We would also like to thank the editor and the anonymous reviewers for their constructive comments and feedback to improve the quality of this chapter.

REFERENCES

Abellán, A., Oppikofer, T., Jaboyedoff, M., Rosser, N. J., Lim, M., & Lato, M. J. 2014. Terrestrial laser scanning of rock slope instabilities. *Earth Surface Processes and Landforms*, 39(1), 80–97. doi:10.1002/esp.3493

Agisoft. 2014. PhotoScan. Retrieved from http://www.agisoft.ru/

Aguilar, M. A., Saldana, M. D., & Aguilar, F. J. 2014. Generation and quality assessment of stereo-extracted DSM from GeoEye-1 and WorldView-2 imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 52(2), 1259–1271. doi:10.1109/Tgrs.2013.2249521

Ai, M., Hu, Q., Li, J., Wang, M., Yuan, H., & Wang, S. 2015. A robust photogrammetric processing method of low-altitude UAV images. *Remote Sensing*, 7(3), 2302–2333. doi:10.3390/rs70302302

Arranz, Á., Sánchez, Á., & Alvar, M. 2012. Multiresolution energy minimisation framework for stereo matching. *IET Computer Vision*, 6(5), 425–434. doi:10.1049/iet-cvi.2010.0212

BAE Systems. 2015. *SOCET SET. Retrieved from* http://www.geospatialexploitationproducts.com/content/products/socet-set

Baltsavias, E., Gruen, A., Eisenbeiss, H., Zhang, L., & Waser, L. T. 2008. High-quality image matching and automated generation of 3D tree models. *International Journal of Remote Sensing*, 29(5), 1243–1259. doi:10.1080/01431160701736513

Baltsavias, E. P. 1991. Multiphoto geometrically constrained matching. *PhD* thesis, Federal Institute of Technology Zurich, Zurich, Switzerland.

Brown, M. Z., Burschka, D., & Hager, G. D. 2003. Advances in computational stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(8), 993–1008. doi:10.1109/TPAMI.2003.1217603

Broxton, M. J., & Edwards, L. J. 2008. The Ames Stereo Pipeline: Automated 3D surface reconstruction from orbital imagery. *Lunar and Planetary Science Conference*, Vol. 39, League City, Texas.

Campbell, N. D., Vogiatzis, G., Hern, C., & Cipolla, R. 2008. Using multiple hypotheses to improve depth-maps for multi-view stereo. Paper presented *at the Proceedings of the 10th European Conference on Computer Vision: Part I, Marseille*, France.

Canny, J. 1986. A computational approach to edge-detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6), 679–698.

Chen, J.-S., & Medioni, G. 1990. Parallel multiscale stereo matching using adaptive smoothing. In O. Faugeras (Ed.), *Computer Vision—ECCV 90*, Vol. 427, pp. 99–103. New York: Springer-Verlag.

Christmas, W. J., Kittler, J., & Petrou, M. 1995. Structural matching in computer vision using probabilistic relaxation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(8), 749–764. doi:10.1109/34.400565

City of Surrey. 2014. *Surrey open data*. Retrieved from http://www.surrey.ca/city-services/658.aspx

Collins, R. T. 1996. A space-sweep approach to true multi-image matching. Paper presented *at the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*.

de Franchis, C., Meinhardt-Llopis, E., Michel, J., Morel, J. M., & Facciolo, G. 2014. An automatic and modular stereo pipeline for pushbroom images. Paper presented *at the ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*.

Di, K., Ma, R., & Li, R. X. 2003. Rational functions and potential for rigorous sensor model recovery. *Photogrammetric Engineering and Remote Sensing*, 69(1), 33–41.

Förstner, W. 1986. A feature based correspondence algorithm for image matching. *International Archives of Photogrammetry and Remote Sensing*, 26(3), 150–166.

Franchis, C. D., Meinhardt-Llopis, E., Michely, J., Morel, J.-M., & Facciolo, G. 2014. On stereo-rectification of pushbroom images. Paper presented *at the International Conference on Image Processing (ICIP)*, Paris, France.

Furukawa, Y. 2014. *Clustering views for multi-view stereo (CMVS)*. Retrieved from http://www.di.ens.fr/cmvs/

Furukawa, Y., & Hernández, C. 2015. Multi-view stereo: A tutorial. *Foundations and Trends in Computer Graphics and Vision*, 9(1–2), 1–148. doi:10.1561/0600000052

Furukawa, Y., & Ponce, J. 2010. Accurate, dense, and robust multiview stereopsis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(8), 1362–1376. doi:10.1109/TPAMI.2009.161

Gennory, D. B. 1980. Modelling the environment of an exploring vehicle by means of stereo vision. *PhD* thesis, Standford University.

Ginzler, C., & Hobi, M. L. 2015. Countrywide stereo-image matching for updating digital surface models in the framework of the Swiss National Forest Inventory. *Remote Sensing*, 7(4), 4343-4370. doi:10.3390/rs70404343

GitHub. 2015. *OpenDroneMap Project*. Retrieved from https://github.com/OpenDroneMap/OpenDroneMap

Grodecki, J. 2001. Ikonos stereo feature extraction—RPC approach. Paper presented *at the Proceedings of the ASPRS Annual Conference*.

Grodecki, J., & Dial, G. 2003. Block adjustment of high-resolution satellite images described by rational polynomials. *Photogrammetric Engineering and Remote Sensing*, 69(1), 59–68.

Gruen, A. 1985. Adaptive least squares correlation: A powerful image matching technique. *South Africa Journal of Photogrammetry, Remote Sensing and Cartography*, 14, 175–187.

Gruen, A. 2012. Development and status of image matching in photogrammetry. *Photogrammetric Record*, 26(137), 36–57. doi:10.1111/j.1477-9730.2011.00671.x

Gruen, A. W., & Baltsavias, E. P. 1986. Adaptive least squares correlation with geometrical constraints. Paper presented *at the 1985 International Technical Symposium/Europe*.

Gruen, A. W., & Baltsavias, E. P. 1988. Geometrically constrained multiphoto matching. *Photogrammetric Engineering and Remote Sensing*, 54(5), 633–641.

Gwinner, K., Scholten, F., Spiegel, M., Schmidt, R., Giese, B., Oberst, J., Heipke, C., Jaumann, R., & Neukum, G. 2009. Derivation and validation of high-resolution digital terrain models from mars express HRSC data. *Photogrammetric Engineering and Remote Sensing*, 75(9), 1127–1142.

Habib, A., Shin, S. W., Kim, K., Kim, C., Bang, K. I., Kim, E. M., & Lee, D. C. 2007. Comprehensive analysis of sensor modeling alternatives for high resolution Imaging satellites. *Photogrammetric Engineering and Remote Sensing*, 73(11), 1241–1251.

Hannah, M. J. 1974. Computer matching of areas in stereo images. *PhD* thesis, Stanford University.

Harris, C., & Stephens, M. 1988. A combined corner and edge detection. Paper presented *at the Proceedings of The Fourth Alvey Vision Conference*.

Hernández, E. C., & Schmitt, F. 2004. Silhouette and stereo fusion for 3D object modeling. *Computer Vision and Image Understanding*, 96(3), 367–392. doi:10.1016/j.cviu.2004.03.016

Hirschmuller, H. 2008. Stereo processing by semiglobal matching and mutual information. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(2), 328–341. doi:10.1109/TPAMI.2007.1166

Hirschmuller, H., & Scharstein, D. 2009. Evaluation of stereo matching costs on images with radiometric differences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(9), 1582–1599. doi:10.1109/TPAMI.2008.221

Hirschmüller, H., Scholten, F., & Hirzinger, G. 2005. Stereo vision based reconstruction of huge urban areas from an airborne pushbroom camera (HRSC). In W. Kropatsch, R. Sablatnig, & A. Hanbury (Eds.), *Pattern Recognition*, Vol. 3663, pp. 58–66. Berlin: Springer.

IGN (National Geographic Institute). 2014. *MicMac*. Retrieved from http://logiciels.ign.fr/?Micmac

Institute for Photogrammetry. 2014. *Photogrammetric surface reconstruction from imagery*. Retrieved from http://www.ifp.uni-stuttgart.de/publications/software/sure/index.en.html

Ishikawa, H. 2003. Exact optimization for Markov random fields with convex priors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(10), 1333–1336. doi:10.1109/Tpami.2003.1233908

Jaboyedoff, M., Oppikofer, T., Abellán, A., Derron, M.-H., Loye, A., Metzger, R., & Pedrazzini, A. 2012. Use of LIDAR in landslide investigations: A review. *Natural Hazards*, 61(1), 5–28. doi:10.1007/s11069-010-9634-2

James, M. R., & Robson, S. 2012. Straightforward reconstruction of 3D surfaces and topography with a camera: Accuracy and geoscience application. *Journal of Geophysical Research: Earth Surface*, 117(F3), F03017. doi:10.1029/2011JF002289

Jeong, J., Yang, C., & Kim, T. 2015. Geo-positioning accuracy using multiple-satellite images: IKONOS, QuickBird, and KOMPSAT-2 stereo images. *Remote Sensing*, 7(4), 4549–4564. doi:10.3390/rs70404549

Jiang, W. 2004. Multiple aerial image matching and automatic building detection. *PhD thesis*, Wuhan University, Wuhan, China.

Kanade, T., & Okutomi, M. 1994. A stereo matching algorithm with an adaptive window: Theory and experiment. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(9), 920–932. doi:10.1109/34.310690

Kim, T. 2000. A study on the epipolarity of linear pushbroom images. *Photogrammetric Engineering and Remote Sensing*, 66(8), 961–966.

Li, Z., Wang, K., Jia, W., Chen, H.-C., Zuo, W., Meng, D., & Sun, M. 2015. Multiview stereo and silhouette fusion via minimizing generalized reprojection error. *Image and Vision Computing*, 33(0), 1–14. doi:10.1016/j.imavis.2014.10.008

Lillesand, T. M., Kiefer, R. W., & Chipman, J. W. 2008. *Remote Sensing and Image Interpretation*. Hoboken, NJ: John Wiley & Sons.

Lowe, D. G. 2004. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2), 91–110. doi:10.1023/B:VISI.0000029664.99615.94

Marr, D., & Poggio, T. 1979. A computational theory of human stereo vision. *Proceedings of the Royal Society B: Biological Sciences*, 204(1156), 301–328.

Matthies, L., Kanade, T., & Szeliski, R. 1989. Kalman filter-based algorithms for estimating depth from image sequences. *International Journal of Computer Vision*, 3(3), 209–238. doi:10.1007/BF00133032

Meer, P., & Georgescu, B. 2001. Edge detection with embedded confidence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(12), 1351–1365. doi:10.1109/34.977560

Moffitt, F. H., & Mikhail, E. M. 1980. *Photogrammetry*. New York: Harper & Row Publishers, Inc.

Moravec, H. P. 1979. Visual mapping by a robot rover. Paper presented *at the Proceedings of the 6th International Joint Conference on Artificial Intelligence, Vol. 1*, Tokyo, Japan.

Morgan, M., Kim, K. O., Jeong, S., & Habib, A. 2006. Epipolar resampling of space-borne linear array scanner scenes using parallel projection. *Photogrammetric Engineering and Remote Sensing*, 72(11), 1255-1263.

NASA. 2014. *The Ames Stereo Pipeline: NASA's open source automated stereogrammetry software (v2.4.2)*. Ames Research Center Intelligent Robotics Group. Retrieved from https://ti.arc.nasa.gov/tech/asr/groups/intelligent-robotics/ngt/stereo/

NeoGeographyToolkit. 2014. *MultipleViewPipeline*. Retrieved from https://github.com/NeoGeographyToolkit/StereoPipeline

Noh, M. J., Cho, W., & Bang, K. I. 2012. Highly dense 3D surface generation using multi-image matching. *ETRI Journal*, 34(1), 87–97. doi:10.4218/etrij.12.1611.0011

Oh, J., Lee, W. H., Toth, C. K., Grejner-Brzezinska, D. A., & Lee, C. 2010. A piecewise approach to epipolar resampling of pushbroom satellite images based on RPC. *Photogrammetric Engineering and Remote Sensing*, 76(12), 1353–1363.

Okutomi, M., & Kanade, T. 1993. A multiple-baseline stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(4), 353–363. doi:10.1109/34.206955

OTB Development Team. 2014. *The ORFEO Tool Box Software Guide Updated for OTB-4.2*. Retrieved from http://www.orfeo-toolbox.org

Paparoditis, N., Thom, C., & Jibrini, H. 2000. Surface reconstruction in urban areas from multiple views with aerial digital frame cameras. Paper presented *at the ISPRS Congress Amsterdam*, Amsterdam, Netherlands.

Pierrot-Deseilligny, M., & Paparoditis, N. 2006. A multiresolution and optimization-based image matching approach: An application to surface reconstruction from SPOT5-HRS stereo imagery. *Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36(1/W41), 1–5.

Pix4D. 2015. *Pix4Dmapper*. Retrieved from https://pix4d.com

Remondino, F., Spera, M. G., Nocerino, E., Menna, F., & Nex, F. 2014. State of the art in high density image matching. *Photogrammetric Record*, 29(146), 144–166. doi:10.1111/phor.12063

Rogmans, S., Lu, J., Bekaert, P., & Lafruit, G. 2009. Real-time stereo-based view synthesis algorithms: A unified framework and evaluation on commodity GPUs. *Signal Processing Image Communication*, 24(1-2), 49–64. doi:10.1016/j.image.2008.10.005

Roy, S., & Cox, I. J. 1998. A maximum-flow formulation of the N-camera stereo correspondence problem. Paper presented *at the Sixth International Conference on Computer Vision*.

Scharstein, D. 1994. Matching images by comparing their gradient fields. Paper presented *at the Proceedings of the 12th IAPR International Conference on Image Processing*.

Schmid, C., & Zisserman, A. 2000. The geometry and matching of lines and curves over multiple views. *International Journal of Computer Vision*, 40(3), 199–233. doi:10.1023/a:1008135310502

Sedaghat, A., Ebadi, H., & Mokhtarzade, M. 2012. Image matching of satellite data based on quadrilateral control networks. *Photogrammetric Record*, 27(140), 423–442. doi:10.1111/j.1477-9730.2012.00699.x

Seitz, S., Curless, B., Diebel, J., Scharstein, D., & Szeliski, R. 2015. *Multi-view stereo evaluation*. Retrieved from http://vision.middlebury.edu/mview/eval/

Seitz, S. M., Curless, B., Diebel, J., Scharstein, D., & Szeliski, R. 2006. A comparison and evaluation of multi-view stereo reconstruction algorithms. Paper presented *at the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*.

Stentoumis, C., Grammatikopoulos, L., Kalisperakis, I., & Karras, G. 2014. On accurate dense stereo-matching using a local adaptive multi-cost approach. *ISPRS Journal of Photogrammetry and Remote Sensing*, 91, 29–49. doi:10.1016/j.isprsjprs.2014.02.006

Stumpf, A., Malet, J. P., Allemand, P., Pierrot-Deseilligny, M., & Skupinski, G. 2015. Ground-based multi-view photogrammetry for the monitoring of landslide deformation and erosion. *Geomorphology*, 231(0), 130–145. doi:10.1016/j.geomorph.2014.10.039

Szeliski, R. 2010. *Computer Vision: Algorithms and Applications*. Heidelberg, Germany: Springer.

Tack, F., Buyuksalih, G., & Goossens, R. 2012a. 3D building reconstruction based on given ground plan information and surface models extracted from spaceborne imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 67(1), 52–64. doi:10.1016/j.isprsjprs.2011.10.003

Tack, F., Buyuksalih, G., & Goossens, R. 2012b. A mixed spaceborne sensor approach for surface modelling of an urban scene. *International Journal of Remote Sensing*, 33(19), 6035–6059. doi:10.1080/01431161.2012.676745

Tack, F., Goossens, R., & Buyuksalih, G. 2012c. Assessment of a photogrammetric approach for urban DSM extraction from tri-stereoscopic satellite imagery. *Photogrammetric Record*, 27(139), 293–310. doi:10.1111/j.1477-9730.2012.00691.x

Takeuchi, S. 2005. Effect of speckle reduction filter and multi-window image matching on stereo SAR. Paper presented *at the Asian Association on Remote Sensing—26th Asian Conference on Remote Sensing and 2nd Asian Space Conference*, *ACRS 2005.*

Tao, C. V., & Hu, Y. 2001. A comprehensive study of the rational function model for photogrammetric processing. *Photogrammetric Engineering and Remote Sensing*, 67(12), 1347–1357.

Tola, E., Lepetit, V., & Fua, P. 2010. DAISY: An efficient dense descriptor applied to wide-baseline stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(5), 815–830. doi:10.1109/TPAMI.2009.77

Toldo, R., Fantini, F., Giona, L., Fantoni, S., & Fusiello, A. 2013. Accurate multiview stereo reconstruction with fast visibility integration and tight disparity bounding. *ISPRS—International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XL-5/W1, 243–249.

Tong, X., Liu, X., Chen, P., Liu, S., Luan, K., Li, L., Liu, S., et al. 2015. Integration of UAV-based photogrammetry and terrestrial laser scanning for the three-dimensional mapping and monitoring of open-pit mine areas. *Remote Sensing*, 7(6), 6635–6662. doi:10.3390/rs70606635

Toutin, T., Schmitt, C. V., & Wang, H. 2012. Impact of no GCP on elevation extraction from WorldView stereo data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 72, 73–79. doi:10.1016/j.isprsjprs.2012.05.009

Tsai, R. Y. 1983. Multiframe image point matching and 3-D surface reconstruction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 5(2), 159–174. doi:10.1109/TPAMI.1983.4767368

Vineet, V., & Narayanan, P. J. 2008. CUDA cuts: Fast graph cuts on the GPU. Paper presented *at the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*.

Vogiatzis, G., Torr, P. H. S., & Cipolla, R. 2005. Multi-view stereo via volumetric graph-cuts. Paper presented *at the Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*.

Wang, M., Hu, F., & Li, J. 2011. Epipolar resampling of linear pushbroom satellite imagery by a new epipolarity model. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(3), 347–355. doi:10.1016/j.isprsjprs.2011.01.002

Wrobel, B. P. 1991. Least-squares methods for surface reconstruction from images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 46(2), 67–84. doi:10.1016/0924-2716(91)90017-P

Wu, B., Zhang, Y. S., & Zhu, Q. 2012. Integrated point and edge matching on poor textural images constrained by self-adaptive triangulations. *ISPRS Journal of Photogrammetry and Remote Sensing*, 68(1), 40–55. doi:10.1016/j.isprsjprs.2011.12.005

Xu, Z., Wu, L., Shen, Y., Li, F., Wang, Q., & Wang, R. 2014. Tridimensional reconstruction applied to cultural heritage with the use of camera-equipped UAV and terrestrial laser scanner. *Remote Sensing*, 6(11), 10413–10434. doi:10.3390/rs61110413

Yu, W., Chen, T. H., Franchetti, F., & Hoe, J. C. 2010. High performance stereo vision designed for massively data parallel platforms. *IEEE Transactions on Circuits and Systems for Video Technology*, 20(11), 1509–1519. doi:10.1109/Tcsvt.2010.2077771

Yuan, X., & Ming, Y. 2009. A novel method of multi-image matching using image and space synthesis information. *Acta Geodaetica et Cartographica Sinica*, 38(3), 216–222.

Zabih, R., & Woodfill, J. 1994. Non-parametric local transforms for computing visual correspondence. In J.-O. Eklundh (Ed.), *Computer Vision—ECCV '94*, Vol. 801, pp. 151–158. New York: Springer-Verlag.

Zeng, C., King, D. J., Richardson, M., & Shan, B. 2017a. Fusion of multispectral imagery and spectrometer data in UAV remote sensing. *Remote Sensing*, 9(7). doi:10.3390/rs9070696

Zeng, C., Richardson, M., & King, D. J. 2017b. The impacts of environmental variables on water reflectance measured using a lightweight unmanned aerial vehicle (UAV)-based spectrometer system. *ISPRS Journal of Photogrammetry and Remote Sensing*, 130, 217–230. doi:10.1016/j.isprsjprs.2017.06.004

Zeng, C. Q., Wang, J. F., & Shi, P. J. 2013. A stereo image matching method to improve the DSM accuracy inside building boundaries. *Canadian Journal of Remote Sensing*, 39(4), 308–317. doi:10.5589/m13-039

Zeng, C. Q., Zhao, T., & Wang, J. F. 2014. A multicriteria evaluation method for 3-D building reconstruction. *IEEE Geoscience and Remote Sensing Letters*, 11(9), 1619–1623. doi:10.1109/Lgrs.2014.2302586

Zhang, B., & Miller, S. 1997. Adaptive automatic terrain extraction. Paper presented *at the Integrating Photogrammetric Techniques with Scene Analysis and Machine Vision III, Orlando*, FL.

Zhang, L. 2005. *Automatic Digital Surface Model (DSM) Generation from Linear Array Images*. Ph.D. disseration, Swiss Federal Institute of Technology, Zurich.

Zhang, L., & Gruen, A. 2006. Multi-image matching for DSM generation from IKONOS imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 60(3), 195–211. doi:10.1016/j.isprsjprs.2006.01.001

Zhu, Z., Stamatopoulos, C., & Fraser, C. S. 2015. Accurate and occlusion-robust multi-view stereo. *ISPRS Journal of Photogrammetry and Remote Sensing*, 109, 47–61. doi:10.1016/j.isprsjprs.2015.08.008

Zitnick, C. L., Kang, S. B., Uyttendaele, M., Winder, S., & Szeliski, R. 2004. *High-quality video view interpolation using a layered representation. ACM SIGGRAPH*, Vol. 23, pp. 600–608. Los Angeles, CA: Association for Computing Machinery, Inc.

Zitnick, C. L., & Webb, J. A. 1996. *Multi-baseline Stereo Using Surface Extraction*. Retrieved from Computer Science Department, Carnegie Mellon University. Pittsburgh, PA.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

6 High-Resolution Radar Data Processing and Applications

Joseph R. Buckley

CONTENTS

6.1 FUNDAMENTALS OF SYNTHETIC APERTURE RADAR

There are many treatises in the literature that lay out in detail the principles of synthetic-aperture radar (SAR) (Curlander and McDonough 1991; Henderson and Lewis 1998). The reader is directed to them for a complete description of the theory of SAR. What follows here is a condensed, essentially nonmathematical presentation of the fundamentals of SAR that are relevant to the high-resolution imaging of natural environments.

6.1.1 Basic Terminology and Electromagnetics

Radar is an active remote sensing system in which a radio signal is sent out from an antenna, and the echo from a target is received. Invented in the 1930s and brought to maturity in the Second World War, this system is fundamentally different from all other remote sensing systems. It provides its own signal, allowing it to operate in darkness as well as sunlight (as does LiDAR), and operates at sufficiently low frequency to be reasonably immune to the effects of clouds and weather. The basic operating principle is that the radar emits pulses of electromagnetic energy and records the time and intensity of the echoes resulting from those pulses. The time from emission of a pulse to the receipt of its echo, when multiplied by the speed of light, determines the total path length the pulse has traveled, which is twice the distance between radar and target. An image is constructed from all the echoes received by the radar. Consequently, spaceborne radars must be side-looking, not nadir-looking like most optical sensors.

Most spaceborne radar systems are monostatic, meaning that the transmitter and receiver share the same antenna on the satellite. Bistatic systems are those where the transmit and receive antennae are separated, either on the same satellite or on different ones. The following discussion assumes a monostatic system.

The operation of all radar systems is governed by the following equation:

$$P_r = P_t G_t \times \frac{1}{4\pi R^2} \times \sigma \times \frac{1}{4\pi R^2} \times A_e \tag{6.1}$$

where P_r and P_t are the received and transmitted power respectively, G_t is the antenna gain, A_e is the effective antenna area, R is the distance between the radar and the target, and σ is the radar cross section (Skolnik 1990). The right side of this equation is divided into five parts: the first is the amount of power emitted by the radar in the direction of the target, the second is the inverse square dispersion of the transmitted signal, the third is the interaction of the signal with the target, the fourth is the inverse square dispersion of the reflected signal, and the last is the area of the antenna to receive the reflected signal.

When an electromagnetic wave is incident on a surface, the wave is absorbed by dielectric dipoles on the surface and then reradiated (Elachi and van Zyl 2006, p. 202). The radar cross section, σ, is a measure of this reradiation. By definition, it is the ratio of the actual reflected radar intensity to the intensity of reflection from an isotropic reflector. It contains all the information on how the radar signal is modified by the target. Since the value can range over many orders of magnitude, it is usually displayed logarithmically in decibels.

The radar backscatter cross section is affected by many bio- and geophysical parameters. It is sensitive to the surface morphology on scales similar to or greater than the wavelength of the radar, and to the dielectric constant of the surface. This dielectric constant is in turn dependent not only on the surface material but also on the water content of the material. The relatively long wavelength of radar (in comparison to optical wavelengths) means that there is some penetration of the surface by the radar signal, and hence that the reradiated response comes from the material below

the surface, as well as from the interface itself. This effect may penetrate up to a few wavelengths deep into the material.

Spaceborne radar operates in the microwave region of the electromagnetic spectrum, in a nominal range from about 3 to 70 cm. Much of this spectral region is allocated for other uses, so only a few narrow bands are available for SAR. These bands are shown in Table 6.1. The majority of SAR systems currently in space operate in C band, with a few systems in X and L bands.

6.1.2 Aperture Synthesis and Image Formation

The angular beam width of emissions from a radar antenna is defined by

$$\theta = \frac{\lambda}{L} \tag{6.2}$$

where λ is the radar wavelength, L is the length of the radar antenna, and θ is in radians. For a satellite at an altitude h from the ground, looking at an angle φ from the vertical (the incidence angle), Equation 6.2 gives a ground resolution of (Elachi and van Zyl 2006, Eq 6-73)

$$X_r = \frac{2\lambda h}{L\cos\varphi} \tag{6.3}$$

When populated with values appropriate for a modern spaceborne radar, in this case RADARSAT-2, the raw ground resolution is

$$X_r = \frac{2(0.052)(0.8\times10^6)}{15\cos 45^\circ} = 7.84 \text{ km} \tag{6.4}$$

This resolution is evidently far too coarse for useful imaging of the Earth's surface except in the broadest of senses. Since the radar operating wavelength is restricted to a relatively small range of values, as is the satellite's orbital altitude, the only parameter that can be varied in Equation 6.3 is the antenna length. To achieve a ground resolution of 10 m, an antenna length of $L = 5.88$ km is required. It is not

TABLE 6.1
Bands and Wavelengths Available for Space-Borne SAR

Band	Nominal Wavelength (cm)
X	3
C	5
S	10
L	25
P	70

currently practical to place a rigid structure of this length into Earth orbit, so such an antenna must be synthesized mathematically. The key observation here is that a target on the ground returns a reflection to the satellite for the entire time it is illuminated by the radar beam. Both the magnitude and phase of the echoes from each pulse are recorded, and then are recombined mathematically to simulate the response that would have been received by a physical antenna of that length. The actual process is rather computationally intense (Cumming and Wong 2005).

Noting that the size of the antenna footprint on the ground is

$$\mathscr{L} = \frac{2\pi h}{L} \tag{6.5}$$

a point P on the ground will be in this footprint during the time that the satellite moves this distance, therefore $\mathscr{L}$ is also the length of the synthetic aperture. This aperture will have an angular beamwidth of

$$\theta_s = \frac{\lambda}{\mathscr{L}} = \frac{L}{2h} \tag{6.6}$$

Leading to the result that

$$X_s = h\theta_s = \frac{L}{2} \tag{6.7}$$

where X_s is the size of the footprint of the synthetic aperture, and hence the smallest resolvable distance on the ground (Elachi and van Zyl 2006, pp. 250–251). This result is perhaps counterintuitive for two reasons. First, the synthetic-aperture radar resolution is independent of the altitude of the sensor, and second, the smaller the antenna, the finer the resolution. There are engineering constraints on antenna size that limit how small it can be made, but physical antenna lengths between 5 and 15 m are in use in the current generation of SAR satellites, so resolution of a few meters is achievable with spaceborne SAR.

There are two basic modes of image acquisition: stripmap, where the antenna beam is pointed perpendicularly to the satellite flight path, and spotlight, where the beam is electronically steered to stay pointed at a specific location on the ground for some period of time, up to the entire time that location is visible from the satellite. Stripmap mode has the advantage of allowing continuous acquisition parallel to the flight path of the satellite, while spotlight mode acquires a finer resolution image at the expense of only being able to acquire discontinuous snapshots along the flight path.

6.1.3 Coherent and Incoherent Scattering

Radar waves interact significantly only with structures that are similar to or larger than the wavelength of the transmitted signal. Structures smaller than the radar wavelength

are essentially invisible to the radar. Consequently, the area covered by the footprint of the radar may be considered to be composed of a large number of independent scatterers, each the size of the radar wavelength. For example, if a 5 cm radar wave is incident upon a 10 × 10 m footprint, there will be $(10.0/0.05)\times(10.0/0.05) = 4\times10^4$ individual scatterers. Since they are all illuminated by a coherent wave front from the radar, the reflected wave is the vector sum of the responses from each of these scatterers. If there is one strong reflecting object within the footprint, such as a building, pole, or other man-made object, then its echo can dominate the echoes from weaker scatterers like natural terrain, and the overall response from the footprint will be just that of the strong reflector (a discrete scatterer). If, on the other hand, there are no single strong reflectors, then most or all of the individual scatterers will contribute to the backscatter (a distributed scatterer). Natural surfaces tend to be distributed, while man-made objects tend to be discrete. Whether the response from a footprint or, alternatively, a pixel in the image, is discrete or distributed is significant to the processing and interpretation of the image.

6.1.4 Speckle

In the case of distributed scatterers within a pixel, it may be assumed that the magnitudes of the responses from individual scatterers are statistically distributed according to a Rayleigh distribution, and the phases according to a uniform distribution (Elachi and van Zyl 2006, pp. 244–246). A significant result of these distributions is that there is a statistical variability about the mean value of the magnitude that is proportional to the magnitude. This variability is distinct from the additive thermal and system noise common in optical imagery. Because of the appearance of this statistical noise in the image, it is commonly referred to as speckle. As with any electronic radio system, there will also be system noise present.

$$V_m = (1+\upsilon)V_t + \epsilon \tag{6.8}$$

Here, the measured backscatter voltage is V_m, the "true" voltage is V_t, the speckle is υ, and the system noise ϵ. Through good system design, system noise can be minimized, but speckle is unavoidable. The magnitude of speckle is significant, only 3.66 dB lower than the "true" backscatter itself. For useful image interpretation, the effects of speckle must be minimized. Techniques for this minimization will be discussed in "Basic SAR Image Processing."

6.1.5 Texture

Texture, a pixel-to-pixel variation in image intensity caused by subtle differences in the surface reflectivity, is a completely different phenomenon than speckle. Texture is a property of the reflecting surface, while speckle is a property of the image formation, independent of the surface. Texture does occur in SAR imagery, but is usually far less significant than speckle, so is not often used as a classifying parameter.

6.1.6 Polarimetry

Electromagnetic waves are oscillations of perpendicular electric and magnetic fields that propagate in a direction such that both the electric and magnetic field oscillations are transverse to the direction of propagation. When an electromagnetic wave is generated from an antenna, then the plane of polarization of the electric field is fixed by the antenna, and hence so is the polarization of the signal. Radar signals are therefore, by definition, polarized. The plane of polarization may be either fixed in time, giving linear polarization, or varying in time, giving elliptical or circular polarization. All Earth-observing satellite radar systems to date use linear polarization.

The nomenclature for radar polarization arises from the earliest land-based systems, where the polarization was either horizontal, that is, the plane of oscillation of the electromagnetic wave was parallel to the horizon, or vertical, where the plane of oscillation was perpendicular to the horizon. Spaceborne systems retain this terminology, where horizontal polarization is horizontal, but vertical polarization, being perpendicular to both the horizon and to the direction of propagation, is not vertical.

A detailed discussion of the theoretical basis for radar polarimetry is presented by Boerner et al. (1998) and Cloude (2010), among others. What follows here is a very basic summary.

Transmitting and receiving polarizations are, in principle, independent of one another. Given appropriate antenna design, it is possible to transmit one polarization and receive a different one. The possible combinations are horizontal transmit, horizontal receive (HH); vertical transmit, vertical receive (VV); horizontal transmit, vertical receive (HV); and vertical transmit, horizontal receive (VH).

When multiple polarizations are considered, the radar cross section may be redefined as a four-element matrix

$$\sigma = \begin{bmatrix} \sigma_{\mathrm{HH}} & \sigma_{\mathrm{HV}} \\ \sigma_{\mathrm{VH}} & \sigma_{\mathrm{VV}} \end{bmatrix} \tag{6.9}$$

showing that the reflecting surface may react differently to each of the polarization combinations. The signal $\mathbf{s}$ that the radar receives therefore also has four complex terms

$$\mathbf{s} = \begin{bmatrix} s_{\mathrm{HH}} & s_{\mathrm{HV}} \\ s_{\mathrm{VH}} & s_{\mathrm{VV}} \end{bmatrix} \tag{6.10}$$

These four terms are all complex numbers, representing both amplitude and phase of the backscatter. The phase measured here is the total number of SAR wavelengths between the satellite and the radar footprint on the Earth's surface. This number is large and not meaningful in any useful sense. However, the relative phase between terms is meaningful, so the phase of one term is subtracted from all four phases, thereby setting one to zero and the other three relative to it. Therefore, there are a total of seven independent values (four magnitudes, three relative phases)

potentially measurable from an appropriately designed SAR. However, under almost all circumstances in a natural or man-made environment the two cross-polarized terms are equal, that is, $s_{HV} = s_{VH}$, removing two of these seven parameters. Due to this redundancy, it is usual to rewrite the scattering matrix as a feature vector, either lexicographic

$$\mathbf{f}_{3L} = [s_{HH} \quad \sqrt{2}s_{HV} \quad s_{VV}]^T \tag{6.11}$$

or Pauli

$$\mathbf{f}_{3P} = [s_{HH} + s_{VV} \quad s_{HH} - s_{VV} \quad 2s_{HV}]^T \tag{6.12}$$

It is from either of these feature vectors that most polarimetric SAR imagery analysis ensues.

Operational spaceborne SAR systems vary in the amount of polarimetric information they acquire, and will return from one to five polarimetric parameters, depending on the specific system design and imaging mode. The first generation of SAR satellites (Seasat, ERS-1, ERS-2, Almaz) were capable of transmitting and receiving in only one polarization, either HH or VV, and so only returned a single magnitude for each image pixel (synthetic radar footprint). Many subsequent satellite SARs have been capable of transmitting and receiving multiple polarizations, giving rise to cross-polarized (HV or VH) modes returning one magnitude value per pixel, dual-polarization (HH/HV or VV/VH) modes that return two magnitudes and possibly the interchannel phase, and quad-polarization (HH, HV, VH, VV) modes that usually return all five available parameters. While it might seem that getting more information through quad-polarization imagery is better than getting less through other modes of imaging, there are trade-offs of image extent and resolution that make the choice of imaging mode application specific.

6.2 BASIC SAR IMAGE PROCESSING

6.2.1 Multilooking and Speckle Reduction

Although the pixel size of a SAR image may be small enough to be considered high resolution, the influence of speckle in the image makes the image at maximum resolution practically unusable. In order to reduce the effect of speckle, some averaging of the image is necessary. A simple technique used in the early days of SAR processing was multilooking. This involved splitting the synthetic aperture into a number of segments (looks), processing each one separately, and averaging the resulting images. Since each look was computed using a fraction of the original synthetic aperture, the resolution of the image was reduced, that is, if the aperture was split into four segments, the resolution of the image was reduced by a factor of four. Reducing the aperture size also had the beneficial effect of reducing the computational time required to produce the image. In the early years of SAR processing, computing power was a limiting factor, but, with the advances in computer hardware and software

design, this is no longer the case. It is standard now to process SAR imagery to the single look complex (SLC) form. This form retains the magnitude and phase of the SAR signal in each pixel, but also retains all the speckle. Speckle reduction is now carried out in the data processing after image formation from the raw SAR data.

Although modern processing techniques no longer rely on multilook processing, the terminology has been maintained. The number of looks corresponds roughly to the number of pixels averaged into each pixel of the speckle filtered image. As the number of looks, or the number of pixels averaged, increases, the severity of speckle decreases as does its influence on polarimetric parameters (Lee et al. 2008).

There are two basic classes of spatial speckle filters: nonadaptive and adaptive. Nonadaptive filters work on the image as a whole and do not consider local image statistics. The simplest of these is a block, or boxcar, average of the image

$$s'(i,j) = \frac{1}{MN} \sum_{k=i-(N/2)}^{i+(N/2)} \sum_{l=j-(M/2)}^{j+(M/2)} s(k,l) \tag{6.13}$$

where M and N are odd integers and usually $M = N$. Each pixel is replaced by the average of the pixels around it. This type of filter, however, reduces the resolution of the image by a ratio of $1 : MN$, a result that is not useful for high-resolution imagery. Adaptive filters use the image statistics in the local area around a pixel to determine how to best estimate a speckle-reduced value for the pixel. Examples are Lee (1981) or Frost et al. (1982). These filters smooth homogeneous areas while preserving the sharpness of boundaries between them.

Speckle reduction in polarimetric images is more complicated due to the requirement to maintain the phase between channels. Several adaptive filters have been developed for polarimetric including the Lopez filter (López-Martínez and Fabregas 2003), the refined Lee filter (Lee et al. 1999b), and the Lee sigma filter (Lee et al. 2009). This last-mentioned filter preserves discrete scatterers as well as edges between regions of distributed scatterers. Because it preserves the fine structure in an image while reducing the overall effect of speckle, filtering of this sort is a requirement for high-resolution use of SAR imagery. An example of this filtering is shown in Figure 6.1. This figure shows a RADARSAT-2 image of a solar power facility in northern China. The unfiltered image in Figure 6.1a shows a field of solar panels surrounded by a fence. The white dots near the bottom of the image are small buildings regularly spaced throughout the solar array. The larger bright area near the center of the image is a group of buildings. The panels themselves do not show up well in the image because they are smooth and are oriented facing south, and therefore reflect the radar energy away from the satellite. A road is visible on the right side of the image. In spite of the relatively homogeneous area of solar panels, the image appears very noisy. In Figure 6.1b, the image has been filtered through a 7×7 boxcar (block) filter. The speckle has been significantly reduced, but the entire image is blurred. If only the statistics of the image in the larger homogeneous regions are important, then this filtering technique is excellent. In Figure 6.1c, the image has been filtered with a Lee sigma filter. The resolution of the small-scale features has been maintained, and the speckle in areas away from these features has been significantly

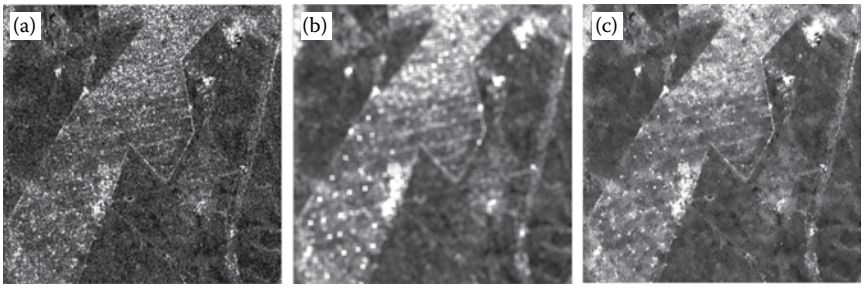


FIGURE 6.1 Speckle filtering applied to a RADARSAT-2 image of a solar power facility: (a) unfiltered image; (b) boxcar filtered image; (c) Lee sigma filtered image.

reduced, thus preserving the image statistics on both small and large scales. The only issue with this adaptive filtering is that different amounts of averaging are applied to different parts of the image and therefore the number of looks varies throughout the image, creating some theoretical difficulty in interpreting some statistical measures in the image. For most uses, however, this issue is not important, and this adaptive filtering produces a very usable product.

Under specific circumstances it is possible to use temporal filtering instead of spatial filtering. When the objects of interest in the scene are unlikely to change over time, then speckle filtering using a time series of imagery may be performed to maintain the original resolution of the imagery. The process usually used is the simple averaging of each pixel through the stack of imagery once the images have been coregistered. Thus, a stack of N images will yield N looks at each pixel location. This technique is often used to reduce speckle in man-made objects such as buildings and roads when an appropriate time series of imagery is available.

6.2.2 Polarimetric Parameters and Target Decomposition Theorems

Fully polarimetric radar contains five times as much information as do single channel systems. However, examination of the three channel amplitudes and two interchannel phases by themselves does not usually give much insight into the properties of the underlying surface. Some basic analysis of these parameters is necessary to provide this insight. The analysis is normally based on derived statistics from these parameters. A thorough review of these derived quantities and their interpretation is given in Touzi et al. (1992).

These analyses have now been largely supplanted by efforts to model the surface scattering properties, and to apply these models to the polarimetric data. In the general case of distributed scatterers, appropriate for natural surfaces and many larger structures, the surface backscatter is modeled as a combination of:

1. Single bounce, or surface scattering: This sort of scattering comes from fields, water surfaces, pavement, and similar sorts of surfaces that are relatively smooth. In color imagery, this type of scattering is traditionally displayed in the blue channel.

2. Double bounce, or dihedral scattering: This scattering arises when the incoming radar signal is reflected off two surfaces perpendicular to each other. Examples of this sort of scattering include reflections off the ground then the side of a building or a tree trunk, or off a water surface then off stalks of emergent vegetation. In color imagery, this type of scattering is traditionally displayed in the red channel.
3. Volume scattering: This scattering comes principally from vegetation, such as tree branches or bushy crops. In color imagery, this type of scattering is traditionally displayed in the green channel.

The Pauli feature vector Equation 6.12 contains some basic information about these types of scattering. The first term $s_{HH}+s_{VV}$, is associated with single-bounce backscatter, the second term, $s_{HH}-s_{VV}$, with double-bounce backscatter, and the third, s_{HV}, with volume scattering. A resulting Pauli image, in which $(s_{HH}+s_{VV}, s_{HH}-s_{VV}, s_{HV})$are displayed as (blue, red, green) can give a good general impression of the features of the landscape, but does not convey any sense of the actual amount of backscatter from different sources. To do that, use of a target decomposition theorem is required. These theorems may be either model based, or eigenvector/eigenvalue based. A thorough review of the basis of these models is presented by Cloude and Pottier (1996).

The starting point for most polarimetric modeling is either the polarimetric covariance matrix

$$\mathbf{C}=\left\langle \mathbf{f}_{3L}\cdot\mathbf{f}_{3L}{}^{\dagger}\right\rangle=\begin{bmatrix}\left\langle\left|s_{HH}\right|^{2}\right\rangle & \sqrt{2}\left\langle s_{HH}s_{HV}{}^{*}\right\rangle & \left\langle s_{HH}s_{VV}{}^{*}\right\rangle\\ \sqrt{2}\left\langle s_{HV}s_{HH}{}^{*}\right\rangle & \left\langle\left|s_{HV}\right|^{2}\right\rangle & \sqrt{2}\left\langle s_{HV}s_{VV}{}^{*}\right\rangle\\ \left\langle s_{VV}s_{HH}{}^{*}\right\rangle & \sqrt{2}\left\langle s_{VV}s_{HV}{}^{*}\right\rangle & \left\langle\left|s_{VV}\right|^{2}\right\rangle\end{bmatrix} \tag{6.14}$$

or the polarimetric coherency matrix

$$\mathbf{T}=\left\langle \mathbf{f}_{3P}\cdot\mathbf{f}_{3P}{}^{\dagger}\right\rangle=$$
$$\frac{1}{2}\begin{bmatrix}\left\langle\left|s_{HH}+s_{VV}\right|^{2}\right\rangle & \left\langle\left(s_{HH}+s_{VV}\right)\left(s_{HH}-s_{VV}\right)^{*}\right\rangle & 2\left\langle\left(s_{HH}+s_{VV}\right)s_{HV}{}^{*}\right\rangle\\ \left\langle\left(s_{HH}-s_{VV}\right)\left(s_{HH}+s_{VV}\right)^{*}\right\rangle & \left\langle\left|s_{HH}-s_{VV}\right|^{2}\right\rangle & 2\left\langle\left(s_{HH}-s_{VV}\right)s_{HV}{}^{*}\right\rangle\\ 2\left\langle s_{HV}\left(s_{HH}+s_{VV}\right)^{*}\right\rangle & 2\left\langle s_{HV}\left(s_{HH}-s_{VV}\right)^{*}\right\rangle & 4\left\langle\left|s_{HV}\right|^{2}\right\rangle\end{bmatrix} \tag{6.15}$$

The averaging implied by $\langle . \rangle$ may be performed as part of the creation of the matrix, but is more often performed at the speckle filtering stage of image processing.

The matrix itself is computed on a pixel-by-pixel basis with no averaging, and then the matrix is processed through a speckle filter (Lee and Pottier 2009). Both the coherence and the coherency matrices contain nine independent parameters: six magnitudes and three off-diagonal phases. Due to the HV-VH symmetry, the elements below the diagonal are just the complex conjugates of those above the diagonal, and therefore do not contain any independent information. Note that the diagonal of the **T** matrix is just an averaged squared Pauli feature vector.

Model-based decompositions separate the total backscattered energy into amounts of surface, double-bounce, and volume energy. One of the first of these models was derived by Freeman and Durden (1998). This relatively simple model uses only five of the nine parameters in the coherency matrix. Omitting some of the information present in the full coherency matrix means that not all the information present in the polarimetric response of the surface is used. Although it has been shown to overestimate volume scattering, even to the point of exceeding the total backscatter (van Zyl et al. 2011), and can exhibit some instabilities, it is still in common use. Seeking improvements on this model, Yamaguchi et al. (2005), and subsequently Singh et al. (2013), added a fourth scattering type, helix scattering, that may arise from man-made structures and urban areas, and used more parameters in the coherency matrix to provide better and more stable estimates of the amounts of the different scattering types. van Zyl et al. (2011) proposed a revised three-component model in which the volume scatter was estimated stably.

The most commonly used eigenvalue/eigenvector decomposition is that of Cloude and Pottier (1997). In it, the eigenvalues and eigenvectors of the coherency matrix are determined as follows:

$$\langle \mathbf{T} \rangle = \sum_{i=1}^{3} \lambda_i [\mathbf{T}_i] = \sum_{i=1}^{3} \lambda_i \mathbf{u}_i \cdot \mathbf{u}_i^{*\mathrm{T}} \tag{6.16}$$

where λ_i are the eigenvalues, and $\mathbf{u}_\mathrm{i}$ are the eigenvectors of the **T** matrix. The eigenvectors represent the three basic scattering mechanisms, and the eigenvalues their relative importance. If one eigenvector is much larger than the other two, then the pixel contains a single scattering type, and if all three are close to the same size, then there is no dominant scatterer. This behavior is presented in the three parameters entropy (H), average alpha angle (α), and anisotropy (A). Entropy is a measure of the disorder of the polarization, and ranges from 0 to 1, where 0 represents a single pure scattering type, and 1 represents complete disorder, with no dominant type. The average alpha angle defines the dominant scattering type. It has a range from 0° to 90°, where 0° represents surface scattering, 45° volume scattering, and 90° double-bounce scattering. Anisotropy represents the difference in magnitude of the second and third eigenvalues where $0 \leq A \leq 1$, 0 implying equality of the two lesser eigenvalues, and 1 implying the dominance of the second over the third. Images displaying these parameters, either singly or in combination, are often helpful in interpreting the polarimetric data.

Decomposition algorithms have recently been developed that combine model-based with eigenvector approaches (van Zyl et al. 2011).

6.2.3 Image Classification and Segmentation

Before performing an image classification, it is useful to estimate the statistical likelihood of the separability of the classes. A metric that has been used is the Bhattacharyya distance (Morio et al. 2007). This metric is defined as

$$D_{\text{bhat}} = \frac{1}{8}(M_j - M_i)^T \left[\frac{\Sigma_i + \Sigma_j}{2}\right]^{-1} (M_j - M_i) + \frac{1}{2}\ln \frac{\left|(\Sigma_i + \Sigma_j)/2\right|}{\sqrt{\left|\Sigma_i\right|\left|\Sigma_j\right|}} \tag{6.17}$$

where M_i and M_j are the centroids of two training groups of pixels and Σ_i and Σ_j are their respective covariance matrices. The first term estimates the separability due to the distance between the class means, and the second due to the class covariance matrices. From this equation, a Bayesian upper bound on the probability of correctly classifying two classes separated by a distance D_{bhat} may be defined as

$$\varepsilon_{\text{bhat}} = 0.5\exp(-D_{\text{bhat}}) \tag{6.18}$$

Table 6.2 shows the Bhattacharyya distance associated with the maximum classification error computed from Equation 6.18. To ascertain the probability of successful classification of more than two classes, an agglomerative procedure was used. First, the Bhattacharyya distance was computed between each pairing of classes. The two classes that are closest are combined into a single larger class, and then the process was repeated until only two classes remain.

The underlying statistical distribution of radar backscatter values for distributed targets is different from that of optical and thermal imaging sensors. While the raw backscatter values in a multipolarized system follow a multivariate complex Gaussian distribution, once they have been processed into a covariance or coherency matrix, they follow a complex Wishart distribution (Lee et al. 1999a). Consequently, it is

TABLE 6.2
Maximum Classification Error and Bhattacharyya Distance

Maximum Classification Error	Bhattacharyya Distance
0.05	2.30
0.10	1.61
0.15	1.20
0.20	0.92
0.25	0.69
0.30	0.51
0.35	0.36
0.40	0.22
0.45	0.11

important when performing any classification operations on this imagery to use tools either based on these distributions or that are nonprobabilistic.

Lee et al. (1994) developed a supervised classification scheme based on minimizing Wishart distances from pixel cluster centers in the polarimetric covariance matrix that were defined by the training data set. This methodology was refined by Anfinsen et al. (2007).

In recent studies for imagery decomposed by any of the model-based techniques, supervised classification of images has shown good results from either a support vector machine (SVM) (Smith and Buckley 2011), or a random forest (Deschamps et al. 2012) methodology. Both methods do not rely on the particular statistical model underlying the data.

All these methods are, to some extent, arbitrary, in the sense that there is no absolute truth against which they can be tested. It has been shown that the classification accuracies of any of the model-based decompositions are similar, and are all reasonably high (Buckley and Smith 2015). It was also found that classification based on any multiple polarization imagery, whether quad polarization, compact polarization, or dual polarization, yielded reasonably high classification accuracy. However, the $H/A/\alpha$ decomposition showed a much lower accuracy, apparently because this decomposition does not use span (total power) information.

There have been extensive studies of the unsupervised classification of imagery decomposed by the $H/A/\alpha$ scheme. Cloude and Pottier (1997) classified imagery by dividing the H/α plane into nine arbitrary regions based on low, medium, or high entropy, and surface, volume, or double-bounce scattering. Since high entropy surface scattering was shown to be not possible theoretically, their classification resulted in eight distinct classes. Lee et al. (1999a) adapted their 1994 supervised classification protocol and used pixels in the eight classes defined by Cloude and Pottier as the initial data set, and iteratively computed class centroids that minimized Wishart distance, yielding eight unsupervised classes. Each of these classes were then split into two parts based on whether their anisotropy value was greater or less than 0.5. The iterative minimization procedure was repeated, yielding 16 classes.

One oddity of this procedure was that, as class centroids were moved in $H/A/\alpha$ space in response to the minimization procedure, they sometimes moved from one dominant scattering type to another. Lee et al. (2004) revised this procedure to constrain class centroids from changing the dominant scattering type.

One property common to most of these classification procedures is that, in general, they perform as well as classifications based on optical imagery (Smith and Buckley 2011). Given that the radar sensors function under all weather conditions and are independent of solar illumination, these polarimetric SAR systems can provide classified imagery more frequently and reliably than optical systems, and therefore are often the system of choice for regular monitoring programs.

6.3 SATELLITE SYSTEMS

The most recent generation of SAR satellites are quite versatile, and can operate in many different imaging modes. Four systems, representative of the current state of the art, are discussed here. They share some common properties in that their phased-array

antennas allow electronic beam steering to make image acquisition possible over a wide range of incidence angles and they all have some selectable polarimetric capability. They are all in polar dawn/dusk orbits to maximize their coverage of the Earth and the amount of time their solar panels are exposed to sunlight. Two of these systems operate in C band, and two in X band. Complete descriptions of these systems are not presented here since many of their capabilities lie in medium-resolution, wide-area imaging and fall outside this volume's focus on high-resolution imagery. No modes with range resolution greater than 15 m are considered. Only imaging capabilities with resolutions higher than this are discussed.

6.3.1 RADARSAT-2

RADARSAT-2 is a Canadian polarimetric SAR satellite launched in December 2007. This satellite operates in C band at 5.406 GHz. It is capable of a wide range of imaging modes, polarizations, and incidence angles (MDA 2016). Table 6.3 gives some orbital parameters for the satellite, and Table 6.4 lists characteristics of the beam modes most relevant to high-resolution imaging. This satellite is capable of routine high-resolution quad-polarization imaging over moderate-sized areas, as well as dual-polarization imaging over larger swaths. Although much of the satellite's capacity is used for operational and commercial purposes, a significant amount of capacity is used for research and application development.

6.3.2 TerraSAR-X/TanDEM-X

TerraSAR-X is an X-band (9.65 GHz) satellite designed, built, and operated by the German Aerospace Center (DLR) in conjunction with Astrium (now Airbus Defence and Space). It was launched on June 15, 2007. A second satellite, TanDEM-X, was launched on June 21, 2010.

The two identical satellites orbit in close formation to allow collection of imagery for the purpose of creating a digital elevation model (DEM) through SAR interferometry. Orbital characteristics of these satellites are given in Table 6.5, and the available high-resolution beam modes in Table 6.6 (Airbus Defence and Space 2014). DLR makes some of the satellites' capacity available for research and application development. The satellites are also commercially available for customer-specified imaging in any imaging mode.

TABLE 6.3
RADARSAT-2 Orbital Characteristics

Characteristic	Value
Orbital Altitude	798 km
Repeat Period	24 days
Orbit Inclination	98.6°
Orbital Period	100.7 min
Ascending Node	18 h

TABLE 6.4
RADARSAT-2 High-Resolution Modes

Mode	Resolution Range × Azimuth (m)	Swath Width (km)	Polarization(s)	Incidence Angle Range
Standard	13.8 × 7.7	100	HH, HH + HV, VV, VV + VH, VH, HV	20°–50°
Wide	13.8 × 7.7	150	HH, HH + HV, VV, VV + VH, VH, HV	20°–45°
Extended High	13.5 × 7.7	75	HH	49°–60°
Extended Low	9.0 × 7.7	170	HH	10°–23°
Fine	5.2 × 7.7	50	HH, HH + HV, VV, VV + VH, VH, HV	30°–50°
Wide Fine	5.2 × 7.7	150	HH, HH + HV, VV, VV + VH, VH, HV	20°–45°
Extra Fine	3.1 × 4.6	125	HH, HV, VH, VV	22°–49°
Multilook Fine	3.1 × 4.6	50	HH, HV, VH, VV	30°–50°
Wide Multilook Fine	3.1 × 4.6	90	HH, HV, VH, VV	29°–50°
Ultrafine	1.6 × 2.8	20	HH, HV, VH, VV	20°–54°
Wide Ultrafine	1.6 × 2.8	50	HH, HV, VH, VV	29°–50°
Spotlight	1.3 × 0.4	18	HH, VV, HV, VH	20°–54°
Standard Quad Polarization	13.5 × 7.6	25	HH + HV + VH + VV	18°–49°
Wide Standard Quad Polarization	13.5 × 7.6	50	HH + HV + VH + VV	18°–42°
Fine Quad Polarization	5.2 × 7.6	25	HH + HV + VH + VV	18°–49°
Wide Fine Quad Polarization	5.2 × 7.6	50	HH + HV + VH + VV	18°–42°

6.3.3 SENTINEL-1

The Sentinel-1 mission is the European Radar Observatory for the Copernicus joint initiative of the European Commission (EC) and the European Space Agency (ESA) (ESA 2013). The two satellites of the constellation were launched April 3, 2014, and

TABLE 6.5
TerraSAR-X/TanDEM-X Orbital Characteristics

Characteristic	Value
Orbital Altitude	514 km
Repeat Period	11 days
Orbit Inclination	97.44°
Orbital Period	94.79 min
Ascending Node	18 h

TABLE 6.6
TerraSAR-X/TanDEM-X High-Resolution Modes

Mode	Resolution Range × Azimuth (m)	Swath Width (km)	Polarization(s)	Incidence Angle Range
Staring Spotlight	0.6 × 0.24	4	VV, HH	20°–45°
High-Resolution Spotlight (300 MHz)	0.6 × 1.1	10	VV, HH	20°–55°
High-Resolution Spotlight	1.2 × 1.1 (single polarization) 1.2 × 2.2 (dual polarization)	10	VV, HH, HH + HV, VV + VH, VV + HH	20°–55°
Spotlight	1.2 × 1.7 1 (single polarization) 1.2 × 3.4 (dual polarization)	10	VV, HH, HH + HV, VV + VH, VV + HH	20°–55°
Stripmap	1.70–3.49 × 3.3 (7.6 dual polarization)	30 (single polarization), 15 (dual polarization)	VV, HH, HH + HV, VV + VH, HH + VV	20°–45°
ScanSAR	1.70–3.49 × 18.5	100	HH, VV	20°–45°

April 25, 2016. They operate in C band (5.405 GHz). The satellites share a common orbit and are spaced 180° apart. The orbital parameters of these satellites are given in Table 6.7, and descriptions of the higher resolution modes in Table 6.8. They acquire imagery in the interferometric wide (IW) mode (VV/VV + VH) over land in a fixed operational scenario, and only change to other modes under exceptional circumstances. Note that they do not support any higher resolution spotlight modes. All imagery acquired is freely available.

6.3.4 COSMO-SkyMed

COSMO-SkyMed is an Italian constellation of four X-band (9.8 GHz) satellites launched between June 2007 and November 2010. The constellation was conceived by

TABLE 6.7
Sentinel-1 Orbital Characteristics

Characteristic	Value
Orbital Altitude	693 km
Repeat Period	12 days
Orbit Inclination	98.18°
Orbital Period	98.74 min
Ascending Node	18 h

TABLE 6.8
Sentinel-1 High-Resolution Imaging Modes

Mode	Resolution Range × Azimuth (m)	Swath Width (km)	Polarization(s)	Incidence Angle Range
Stripmap	5 × 5	80	HH + HV, VV + VH	20°–47°
Interferometric Wide Swath	5 × 20	250	HH + HV, VV + VH	31°–46°

TABLE 6.9
COSMO-SkyMed Orbital Characteristics

Characteristic	Value
Orbital Altitude	619.6 km
Repeat Period	16 days
Orbit Inclination	97.86°
Orbital Period	97.1 min
Ascending Node	06 h

the Italian Space Agency (ASI), funded by the Italian Ministries of Research (MUR) and Defence (MoD), and operated by ASI and MoD. It is designed to serve the needs of both the military and civilian communities (ASI 2007). The four satellites share a common orbit, and are equispaced in it with 90° separation between them. Their orbital characteristics are shown in Table 6.9.

The satellites support both spotlight and stripmap modes with single-channel imagery in any combination of H or V transmitting and H or V receiving. There is also a dual-polarization stripmap mode in which any two polarimetric combinations may be acquired, but without phase information between the two channels. Imaging modes are listed in Table 6.10. While some of the constellation's capacity is used by the Italian military, imagery may be acquired commercially to the user's specification.

TABLE 6.10
COSMOS-SkyMed High-Resolution Imaging Modes

Mode	Resolution Range (m)	Swath Width (km)	Polarization(s)	Incidence Angle Range
Spotlight 2	1	10	HH, VV, HV, VH	25°–60°
Stripmap (HIMAGE)	3	40	HH, VV, HV, VH	25°–60°
Stripmap (PingPong)	15	30	Any two of HH, VV, HV, VH	25°–60°

6.4 CASE STUDIES

6.4.1 Invasive Species in Grasslands

The invasive weed species leafy spurge (*Euphorbia esula*) causes significant problems in the prairie grasslands of Western Canada. This plant, which is toxic to cattle, grows in dense patches and propagates robustly to the exclusion of native grassland species (Belcher and Wilson 1989). Eradication of this weed is very difficult due to its ability to propagate from root tips and from seeds that are propelled explosively up to 15 m from the plant (Thunhorst and Swearingen 2005). It is very difficult to map with multispectral optical imagery because it is only different in color from the surrounding grasses when its brilliant yellow flowers are in bloom. However, the plant is structurally different from grasses, and therefore ought to be visible in polarimetric SAR imagery.

Agriculture and Agri-Food Canada, through their Lethbridge Research and Development Centre under the direction of A. M. Smith, conducted a study in the grasslands of southern Alberta, Canada, to investigate the potential of using RADARSAT-2 quad-polarization imagery to detect and delineate leafy spurge. This study has been described in detail by Coleman (2010). The bulk of the following description is based on this document.

The study site was chosen in a coulee, a river valley deeply eroded into the relatively flat prairie. Twenty-five images were acquired between April and October 2009 at three different incidence angles (beam modes): 23.17° (FQ4), 26.30° (FQ6),

FIGURE 6.2 RADARSAT-2 FQ17 image over southern Alberta, Canada. The study area is shown by the white box. The total backscattered power of the image is displayed.

and 37.22° (FQ17). Ground truth measurements were made in June, July, and August by mapping the edges of patches of leafy spurge with handheld global positioning system (GPS) units. Areas of other types of vegetation that could possibly be confused with leafy spurge in SAR imagery were also mapped. Further validation information was collected through ground-based and aerial photography.

A typical image is shown in Figure 6.2. The SAR imagery was processed through a Lee sigma filter, georeferenced using the image tie points provided in the image header, and then decomposed using both the Freeman-Durden and the Cloude-Pottier methods. These two decompositions, along with the coherency matrix, were used in supervised classification of the imagery. Some results of the experiment are shown in Figure 6.3.

The most detailed field work and ground truth sample collection was done in the two small rectangular areas shown in Figure 6.3a.

Eight classes were chosen to represent the range of natural and human-influenced conditions in the region. They are listed in Table 6.11. They are grouped into flat surfaces, shrubs, tall vegetation, and infrastructure. The total size of training sets for these classes shows that the individual regions of each are very small. These small sizes pose a problem for SAR classification since it is usual for the effects of speckle, even though reduced through filtering, to be mitigated by the use of large training

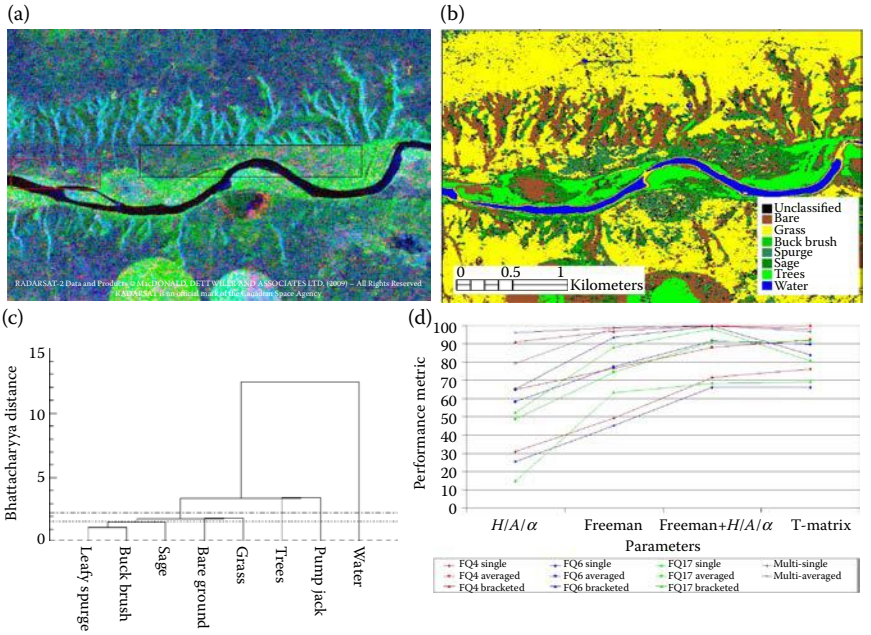


FIGURE 6.3 (a) The region considered by the study, showing the two areas where ground truth measurements were made. (b) SVM classification of the layer stack of the average coherency matrices at the three incidence angles. (c) Bhattacharyya distance tree for the averaged coherency matrices for the FQ4 beam mode. (d) Classification accuracy results for all beam modes and combinations of imagery and classification basis.

TABLE 6.11
Classes Used for Invasive Weed Study

Class Number	Class Content	Class Category	Training Pixels
1	Water		54
2	Bare Ground	Flat Surfaces	74
3	Grass		87
4	Trees	Tall Vegetation	60
5	Pump Jack	Human Infrastructure	125
6	Leafy Spurge	Shrub	76
7	Buckbrush		66
8	Sage		39

data sets. One of the aims of this study, then, was to see if the SAR imagery could be pushed into yielding a high-resolution classification result.

Since the number of pixels in each class was small, tests were performed on various combinations of imagery to determine the best way of identifying leafy spurge. The following combinations of imagery were used:

- An individual image: One image from each of the three beam modes (incidence angles) was chosen.
- Averages of all images at a specific incidence angle: In this procedure, there were three (or nine) parameters per pixel $(\overline{D}_{FD}, \overline{V}_{FD}, \overline{S}_{FD})$,$(\overline{H}, \overline{A}, \overline{\alpha})$ or $(\overline{T}_{ij}, i = 1,\ldots,3, j = i,\ldots,3)$. Three averages were computed, one for each incidence angle.
- Layer stacked at each incidence angle: In this procedure, each image was considered to be independent of the others, so, for a three image set, classifications used nine parameters for the decompositions and 27 for the coherency matrix. Three-layer stacks were created, one for each incidence angle.
- A layer stack of all single images: Since there were three images used, classifications used nine parameters for the decompositions and 27 for the coherency matrix.
- A layer stack of the three averaged images: Classifications of this stack used nine parameters for the decompositions and 27 for the coherency matrix.

As a preliminary check on the probability of meaningful classification, Bhattacharyya distances were computed for the averaged coherency matrices at the three incidence angles. Results of this procedure for the FQ4 beam mode are shown in Figure 6.3c.

In this figure, the dash-dot line represents 95% confidence in classification accuracy, the dotted line 90%, and the dashed line 55%. It can be seen that the water class, the infrastructure (pump jack) class, and the tree class should all be recognized as having better than 95% accuracy and can be separated from the other classes. Grass and bare ground should be correctly classified as having better than 90% accuracy, and the three shrub classes—leafy spurge, buckbrush, and sagebrush—as having

better than 80% accuracy. Results for the other two incidence angles were similar in form, but with marginally lower accuracy estimates.

Classification of the imagery was performed using the support vector machine methodology. An example of the output of the classifier is shown in Figure 6.3b.

The classification appears to be subjectively correct, with trees and shrubs near the river banks, bare soil on the steep slopes of the valley, and grass on the surrounding prairie. A quantitative assessment of the quality of the classification is more difficult to make. The training areas were very small, and there was no independent validation data set. It was possible to compute a producer's accuracy, that is, the percentage of training area pixels classified correctly. This was expected to be overly optimistic (Lillesand et al. 2008, p. 578), but seemed to be the only quantifiable measure that could be extracted from the classification.

The classification was run for all the combinations of imagery listed above, for the Cloude-Pottier $H/A/\alpha$ eigenvalue decomposition, the Freeman-Durden decomposition, a combination of both the Cloude-Pottier and Freeman-Durden decompositions, and the full coherency ($\mathbf{T}$) matrix. A summary of the results is shown in Figure 6.3d.

Several trends are evident from this figure. The Cloude-Pottier decomposition performed more poorly than the others. This result was not unexpected since its algorithm does not use the pixel intensity (span). This quantity is important in both the Freeman-Durden and coherency matrix formulations. The more information input to the classifier, the better the result, with the nine-term coherency matrix performing better than the six-term combination, which in turn performed better than the Freeman-Durden. Incidence angle is important. As was noted in the discussion of Bhattacharyya distances, the steepest incidence angle (FQ4) performed better than the shallower angles (FQ5 and FQ17). Finally, the more images averaged together or otherwise included in the classification, the better the result.

From the perspective of high-resolution remote sensing, this study showed that polarimetric synthetic-aperture radar can be a useful tool as long as the data processing focuses on maintaining the high-resolution aspect of the imagery. Any SAR imagery must be speckle reduced to be statistically useful. Choice of an appropriate speckle filter algorithm is critical, and must be one that is adaptive and edge preserving. The improved Lee sigma filter used here proved appropriate. The effect of speckle was further reduced by multitemporal and multiangle averaging. Through these processes, this study showed that relatively small patches of unique vegetation can be extracted from polarimetric SAR imagery.

6.4.2 Infrastructure

Human-built features in a SAR image are often very difficult to identify. Sharp edges and features smaller than one SAR pixel in size are often blurred or filtered out in speckle-reduction processing. However, by using the higher resolution modes of SAR systems, processed appropriately, human-built infrastructure may be located and identified.

There are two distinct types of features that are discernable in SAR imagery: discrete scatterers such as power poles, fence posts, and other objects that are good radar reflectors, and hard-edged or linear features such as large buildings or roads.

6.4.2.1 Discrete Scatterers: Power Lines, Fence Lines, and Wind Turbines

Objects smaller than the size of a SAR pixel may be imaged successfully if their backscatter is much stronger than that of the distributed scatter from the natural environment surrounding them. RADARSAT-2 fine quad-polarization imagery over grasslands in the Inner Mongolia Autonomous Region, China, illustrates this concept.

Figure 6.4 shows a small area of grassland near the town of Siziwang, Inner Mongolia, China, imaged twice on October 24, 2016. The ascending (satellite moving NNW) and descending (satellite moving SSW) images show many features in common, but also illustrate a significant difficulty with imaging linear features. The circles numbered 1 locate wind turbines. As tall cylinders, they appear equally well both ascending and descending, in double-bounce and volume scattering. They

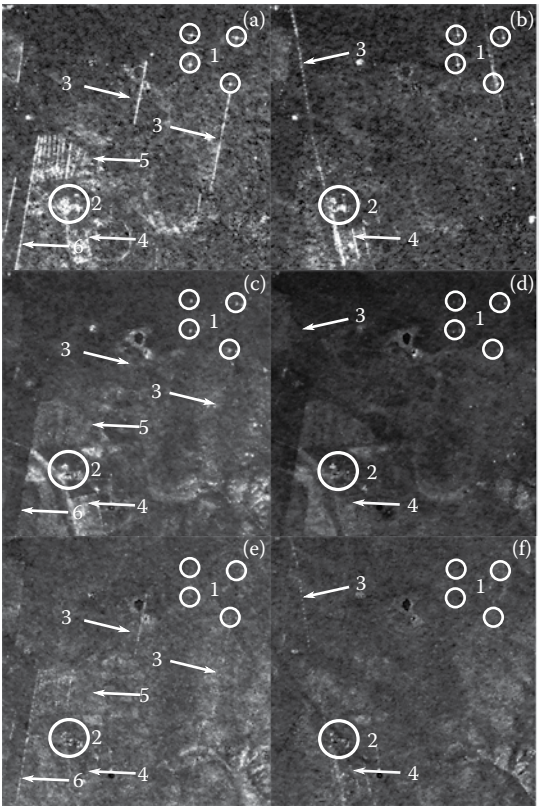


FIGURE 6.4 RADARSAT-2 fine quad-polarization imagery near Siziwang, China, October 24, 2016, shown in a Yamaguchi decomposition. (a), (c), and (e) are descending images; (b), (d), and (f) are ascending images. (a) and (b) show double-bounce scattering, (c) and (d) show volume scattering, (e) and (f) surface scattering. The numbers refer to image features described in the text.

do not appear in the surface scattering images. The circle labeled 2 encloses a group of buildings. Their walls are oriented parallel or perpendicular to the ascending direction of the satellite, and hence appear as double-bounce scatterers. However, they are at an oblique angle to the descending direction, and therefore the double bounce causes a change in polarization from vertical to horizontal and vice versa, and that change is interpreted as volume scattering. The arrows labeled 3 point to power lines. As linear features, they appear only when they are parallel to the satellite's path. They are imaged as double-bounce scatterers. The most likely mechanism is that there is a bounce off the relatively flat ground that hits the power line and is reflected back to the satellite. The reverse path is of course possible too, from the wire to the ground to the satellite. In the ascending images, the power line is imaged more by its supporting poles than by the wire itself. These poles are also weakly visible in the descending image. Arrow 4 points to a field enclosed by a wire fence. The field was prepared for crops and was tilled in the direction of its longer side. Accordingly, in the ascending image, it appears to be a distributed double-bounce scatterer, but in the descending image it is a volume scatterer.

The arrow labeled 6 indicates a wire fence line, imaged in the same way as the power lines. The arrow labeled 5 points to an area of animal pens that are enclosed by wire fencing. As with the fence indicated by arrow 6, these fences are only visible as double-bounce objects where they are parallel to the flight direction of the satellite.

6.4.2.2 Distributed Scatterers: Roads and Well Pads

Some man-made features in SAR imagery are not characterized by discrete or dominant scatterers, but still contain fine detail and sharp edges. These sorts of features may be seen in RADARSAT-2 imagery of an oilfield near Cold Lake, Alberta, Canada (Brown et al. 2015). These images show some well pads (graveled areas surrounding producing oil wells) and the roads connecting them.

In this test, a stack of ultrafine (2 m resolution) HH polarization images acquired from 2011 to 2013 was used. A RapidEye image of the area is shown in Figure 6.5a. Different ways of image preparation were tested to assess which produced the most accurate random forest classification. Treating the image stack as 32 independent channels with no speckle reduction, or performing a minimum noise fraction transform on the image stack, performed better than temporally averaged imagery. Results from the unfiltered image stack are shown in Figure 6.5b.

Roads and well pads are clearly visible in the classified SAR image. A road, indicated by the red arrow, is clearly delineated, but is misidentified as a water feature. An unused well pad, identified by the black arrow, is correctly identified as a grassy area, while a well pad in use (magenta arrow), while clearly separated from the surrounding vegetation, is classified into a mixture of exposed land, urban, and grass. The overall classification accuracy of this imagery was only 37%, showing that, although this high-resolution imagery preserves the boundaries of man-made features, its single HH polarization is not very effective in classifying different types of vegetation.

In a similar region near Slave Lake, Alberta, Canada, imagery was acquired in ultrafine HH and HV and fine quad-polarization modes (Henley et al. 2016). A RapidEye image of the area is shown in Figure 6.5c and an orthophoto in Figure 6.5d. The ultrafine imagery has high spatial resolution (∼2 m), but only one real-valued SAR

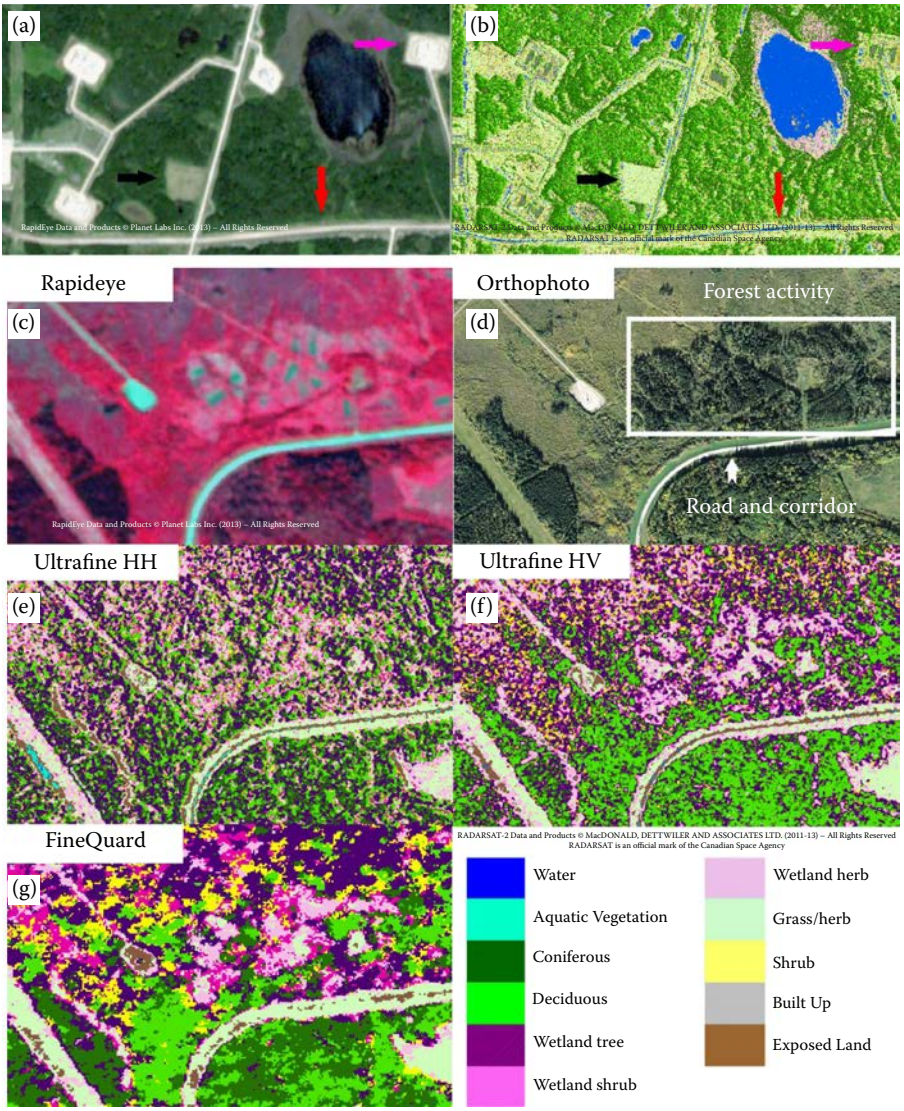


FIGURE 6.5 (a) RapidEye true color image of well pads and roads near Cold Lake, Alberta, Canada. (b) RADARSAT-2 ultrafine classified averaged image of the same area. (From Brown et al. 2015. MARA: Multi-Sensor Assessment of Reclaimed Areas. Milestone 2 Report. ASL Environmental Sciences Inc. for the Canadian Space Agency under Contract No. 9F043-130644-008. Sidney, BC, Canada, Fig 5-21). (c) Roads and a well pad near Slave Lake, Alberta, Canada, as seen by RapidEye, (d) aerial photography, and (e)–(g) three different beam modes of RADARSAT-2. (From Henley et al. 2016. MARA: Multi-Sensor Assessment of Reclaimed Areas. Final Report. ASL Environmental Sciences Inc. for the Canadian Space Agency under Contract No. 9F043-130644-008. Sidney, BC, Canada, Fig 5-3).

channel, while the quad-polarization imagery has lower resolution (~7 m), but with four complex channels of SAR information, allowing an assessment of the trade-offs between the two types of SAR imagery. The ultrafine imagery was speckle filtered using an enhanced Frost filter, and the quad-polarization imagery with a Lee sigma filter. A comparison of classification results from these modes is shown in Figure 6.5e through g. The orthophoto was acquired some time before the other images, and the trees in the area identified as "Forest Activity" were cut down and removed before the other images were acquired. In all three SAR images, the road and its corridor are delineated, although in the coarser resolution of the quad-polarization image the road itself appears to be discontinuous. As was seen in the HH image of Figure 6.5b, the HH image here classifies the road surface as water. This misclassification is not present in either the HV or the quad-polarization image. The well pad in the left center of the image is classified in the quad-polarization image as being exposed land with a grassy border and a wetland tree or built-up structure in the center. With the exception of the tree class, this classification is correct. Neither of the single-polarization images classify the well pad correctly. In general, the quad-polarization classification is smoother and less speckled than the single-polarization results, and does a much better job of separating coniferous from deciduous forest. In the HV image, however, all three SAR classifications overestimate the amount of wetland vegetation.

In summary, the ultrafine imagery is better at delineating man-made features, but the quad-polarization imagery is better at identifying the content of features, both man-made and natural.

6.5 CONCLUDING OBSERVATIONS

Synthetic-aperture radar imagery does have the capability of useful high-resolution monitoring of the Earth's surface. Its all-weather day/night imaging ability is a significant advantage over visible imagery in conducting consistent environmental monitoring programs. However, the presence of speckle in the imagery requires careful processing in order not to decrease the resolution of the imagery.

There are two significant and separate issues that are addressed by SAR imagery: the delineation of features and the identification of the content of the features. For persistent features in imagery, such as man-made structures or alterations to the landscape, high-resolution delineation is relatively easy and reliable and may usually be done reliably with single-channel SAR imagery. For natural landscape features, acquiring imagery in more than one polarimetric channel significantly improves the accuracy of the identification of the content of these features. However, acquiring on multiple channels implies a reduction in spatial resolution. Smaller features may not be imaged at all, but larger features will be imaged and identified more reliably. With suitable speckle filtering, edges between features will be maintained.

REFERENCES

Airbus Defence and Space. 2014. TerraSAR-X Image Product Guide.

Anfinsen, SN, R Jenssen, and T Elthoft. 2007. Spectral Clustering of Polarimetric SAR Data with Wishart-Derived Distance Measures. In *Proceedings of the 3rd International*

Workshop on Science and Applications of SAR Polarimetry and Polarimetric Interferometry, pp. 22–26. Frascati, Italy: European Space Agency.

ASI (Agenzia Spaziale Italiana). 2007. *COSMO-SkyMed System Description & User Guide*. ASI-CSM-ENG-RS-093-A. http://www.cosmo-skymed.it/docs/ASI-CSM-ENG-RS-093-A-CSKSysDescriptionAndUserGuide.pdf.

Belcher, JW, and SD Wilson. 1989. Leafy Spurge and the Species Composition of a Mixed-Grass Prairie. *Journal of Range Management*, 42(2):172–175.

Boerner, WM, H Mott, E Luneburg, C Livingstone, B Brisco, RJ Brown, JS Paterson et al. 1998. Polarimetry in Radar Remote Sensing: Basic and Applied Concepts. In *Manual of Remote Sensing: Principles and Applications of Imaging Radar*, Chapter 5, edited by Henderson FM and Lewis AJ, pp. 271–356. John Wiley and Sons, Inc, Hoboken, NJ.

Brown, L, M Henley, N Rochdi, K Ersahin, and E Loos. 2015. "MARA: Multi-Sensor Assessment of Reclaimed Areas. Milestone 2 Report." ASL Environmental Sciences Inc. for the Canadian Space Agency under Contract No. 9F043-130644-008. Sidney, BC, Canada.

Buckley, JR, and AM Smith. 2015. "Identifying Native Rangelands: How Well Will the RCM SAR Perform?" In *Proceedings of the 10th Advanced SAR Workshop*. Saint-Hubert, QC, Canada.

Cloude, SR. 2010. *Polarisation: Applications in Remote Sensing*. Oxford University Press, New York, NY.

Cloude, SR, and E Pottier. 1996. "A Review of Target Decomposition Theorems in Radar Polarimetry." *IEEE Trans. Geosci. Rem. Sens.* 34 (2): 498–518.

Cloude, SR, and E Pottier. 1997. "An Entropy Based Classification Scheme for Land Applications of Polarimetric SARs." *IEEE Trans. Geosci. Rem. Sens.* 35 (2): 68–78.

Coleman, AJ. 2010. "Polarimetric Radar Identification of Invasive Plant Species in a Prairie Landscape." *MSc thesis*, Royal Military College of Canada.

Cumming, IG, and FH Wong. 2005. *Digital Processing of Synthetic Aperture Radar Data: Algorithms and Implementation*. Artech House, Norwood, MA.

Curlander, JC, and RN McDonough. 1991. *Synthetic Aperture Radar: System and Signal Processing*. New York: John Wiley and Sons, Inc.

Deschamps, B, H McNairn, J Shang, and X Jiao. 2012. "Towards Operational Radar-Only Crop Type Classification: Comparison of a Traditional Decision Tree with a Random Forest Classifier." *Can. J. Rem. Sens.* 38 (1): 60–68. doi:10.5589/m12-012.

Elachi, C, and JJ van Zyl. 2006. *Introduction to the Physics and Techniques of Remote Sensing*, 2nd ed. John Wiley and Sons Inc, Hoboken, NJ.

ESA (European Space Agency). 2013. "*Sentinel-1 User Handbook*." https://www.scribd.com/doc/259520850/Sentinel-1-User-Handbook

Freeman, A, and SL Durden. 1998. "A Three-Component Scattering Model for Polarimetric SAR Data." *IEEE Trans. Geosci. Rem. Sens.* 36 (3): 963–973.

Frost, VS, JA Stiles, KS Shanmugan, and JC Holtzmann. 1982. "A Model for Radar Images and Its Application to Adaptive Digital Filtering of Multiplicative Noise." *IEEE Trans. Patt. Anal. Mach. Intellig. PAMI-4* (2): 157–165.

Henderson, FM, and AJ Lewis. 1998. "Principles and Applications of Imaging Radar." In *Manual of Remote Sensing*, edited by RJ Ryerson, Volume 2, 3rd ed. John Wiley and Sons, Inc, Hoboken, NJ.

Henley M, L Brown, K Ersahin, E Loos, and N Rochdi. 2016. *"MARA: Multi-Sensor Assessment of Reclaimed Areas.* Final Report." ASL Environmental Sciences Inc. for the Canadian Space Agency under Contract No. 9F043-130644-008. Sidney, BC, Canada.

Lee, J-S. 1981. "Refined Filtering of Image Noise Using Local Statistics." *Comp. Graphic. Image Process.* 15: 380–389.

Lee, J-S, TL Ainsworth, J Kelly, and C Lopez-Martinez. 2008. "Statistical Evaluation and Bias Removal of Multi-Look Effect on Entropy/α/Anisotropy in Polarimetric Target Decomposition." In *7th European Conference on Synthetic Aperture Radar.*

Lee, J-S, MR Grunes, TL Ainsworth, L-J Du, DL Schuler, and SR Cloude. 1999a. "Unsupervised Classification Using Polarimetric Decomposition and the Complex Wishart Classifier." *IEEE Trans. Geosci. Rem. Sens.* 37 (5): 2249–2258.

Lee, J-S, MR Grunes, and G de Grandi. 1999b. "Polarimetric SAR Speckle Filtering and Its Implication for Classification." *IEEE Trans. Geosci. Rem. Sens.* 37: 2363–2373.

Lee, JS, MR Grunes, and R Kwok. 1994. "Classification of Multi-Look Polarimetric SAR Imagery Based on Complex Wishart Distribution." *Int. J. Rem. Sens.* 15: 2299–2311.

Lee, J-S, MR Grunes, E Pottier, and L Ferro-Famil. 2004. "Unsupervised Terrain Classification Preserving Scattering Characteristics." *IEEE Trans. Geosci. Rem. Sens.* 42 (4): 722–731.

Lee, J-S, and E Pottier. 2009. *Polarimetric Radar Imaging: From Basics to Applications.* Boca Raton, FL: CRC Press.

Lee, J-S, J-H Wen, TL Ainsworth, K-S Chen, and AJ Chen. 2009. "Improved Sigma Filter for Speckle Filtering of SAR Imagery." *IEEE Trans. Geosci. Rem. Sens.* 47 (1): 202–213.

Lillesand, TM, RW Kiefer, and JW Chipman. 2008. *Remote Sensing and Image Interpretation*, 6th ed. John Wiley and Sons Inc.

López-Martínez, C, and X Fabregas. 2003. "Polarimetric SAR Speckle Noise Model." *IEEE Trans. Geosci. Rem. Sens.* 41 (10): 2232–2242.

MDA. 2016. "*RADARSAT-2 Product Description*." RN-SP-52-1238, Richmond, BC, Canada.

Morio, J, PH Réfrégier, F Goudail, P Dubois-Fernandez, and X Dupuis. 2007. "Application of Information Theory Measures to Polarimetric and Interferometric SAR Images." In *Physics in Signal and Image Processing, 2007.* Mulhouse, France.

Singh, G, Y Yamaguchi, and SE Park. 2013. "General Four-Component Scattering Power Decomposition with Unitary Transformation of Coherency Matrix." *IEEE Trans. Geosci. Rem. Sens.* 51 (5): 3014–3022. doi:10.1109/TGRS.2012.2212446.

Skolnik, MI. 1990. "An Introduction to Radar." In *Radar Handbook*, edited by MI Skolnik, 2nd ed., 1.1–1.21. New York, NY: McGraw-Hill.

Smith, AM, and JR Buckley. 2011. "Investigating RADARSAT-2 as a Tool for Monitoring Grassland in Western Canada." *Can. J. Rem. Sens.* 37 (1): 93–102. doi:10.5589/m11-027.

Thunhorst, G, and JM Swearingen. 2005. "Fact Sheet: Leafy Spurge." Plant Conservation Alliance's Alien Plant Working Group. www.nps.gov/plants/alien/fact/eues1.htm.

Touzi, R, S Goze, T Le Toan, A Lopes, and E Mougin. 1992. "Polarimetric Discriminators for SAR Images." *IEEE Trans. Geosci. Rem. Sens.* 30 (5): 973–980.

van Zyl, JJ, M Arii, and Y Kim. 2011. "Model-Based Decomposition of Polarimetric SAR Covariance Matrices Constrained for Nonnegative Eigenvalues." *IEEE Trans. Geosci. Rem. Sens.* 49 (9): 3452–3459. doi:10.1109/TGRS.2011.2128325.

Yamaguchi, Y, T Moriyama, M Ishido, and H Yamada. 2005. "Four-Component Scattering Model for Polarimetric SAR Image Decomposition." *IEEE Trans. Geosci. Rem. Sens.* 43 (8): 1699–1706.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Section II

Algorithms and Techniques

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

7 Structure from Motion Techniques for Estimating the Volume of Wood Chips

Travis L. Howell, Kunwar K. Singh, and Lindsey Smart

CONTENTS

7.1 INTRODUCTION

Various versions of unmanned aerial systems (UASs) have existed, but until recently, they have been too expensive or incapable of acquiring useful remotely sensed data without requiring extensive prior knowledge (Zhou and Zang 2007). Commercial and research applications of UASs have recently gained momentum due to technological advancements that have reduced costs and increased efficiency. UASs' incredible potential is now being realized due to technological advances like battery-integrated power management, lightweight brushless rotors, and miniature high-resolution digital cameras. As a result of these advancements, common UAS applications have emerged that include coverage analysis for cellular networks, timber harvest auditing, natural disaster inspection, hydrology management, archaeological surveys, wetland mapping, and vegetation monitoring (Berni et al. 2009; Breckenridge et al. 2011; Chiabrando et al. 2009; Goddemeier et al. 2010; Themistocleous et al. 2015; Wallace et al. 2012). These studies use image matching and photogrammetric techniques that allow for the generation of high-density point

clouds from the very-high-resolution photos collected by UASs (Wallace et al. 2012). Initially, data are acquired at a suitable altitude and angle (oblique and/or nadir), with an appropriate percent of overlap to seamlessly stitch together to create a high-resolution 3D point cloud (Dandois et al. 2015). Then, structure from motion (SfM) programs are used to generate, view, and manipulate high-resolution point clouds and orthophotos that can be spatially referenced to centimeter-level accuracy, if ground control points (GCPs) are available. SfM programs initially create a sparse point cloud using the manual marking of GCPs to generate orthomosaics and a rectified dense point cloud.

SfM is a combination of photography and LiDAR techniques (Higgins 2016), which is also referred to as photo detection and ranging (PhoDAR). Compared to costly data collection and processing of LiDAR data, the SfM technique is capable of producing data with similar accuracies but at much lower costs. LiDAR remote sensing uses differences in laser return times to create a 3D point cloud representation of the Earth's surface, while the SfM follows principles of photogrammetry and uses software that creates 3D point clouds by processing high-resolution imagery. Natural resource applications of SfM include, but are not limited to, site inspection, plant species identification, vegetation height models, timber harvest auditing, hydrology management, and prescribed or wild fire reconnaissance (Gates 2015). The key to the accuracy and success of SfM are images with a high percentage of overlapping images. Software applications stitch overlapping images by matching known objects to generate 3D point clouds, orthophotos, and mosaics. SfM generates a point derived by each pixel within the photos, and therefore the point density directly reflects the spatial resolution of the collected photos. SfM has adopted and improved upon traditional stereophotogrammetry principles.

The quality of SfM data depends on the type of equipment, for example, the use of a consumer- versus professional-grade camera. Consumer UASs include platforms similar to DJI Phantoms, wherein the platform acquisition cost is below $1,600 but the camera cannot be interchanged with another sensor. On the other hand, professional-grade UAS platforms cost on average more than $4,000, but come with interchangeable camera/sensor capabilities, including better global positioning system (GPS) equipment and hovering accuracy. Ultimately, all variables depend on the objectives of a project and the required accuracy for the intended outcomes. High-definition sensors (i.e., digital cameras), capable of collecting data in visible, infrared, and thermal portions of the electromagnetic spectrum, can be attached beneath a UAS platform to gather real-time data. For example, Wallace et al. (2012) developed a low-cost UAS-LiDAR system and an accompanying workflow to produce 3D point clouds. They used a modified processing workflow that included a trajectory-determining algorithm for fusing observations from a GPS receiver, an inertial measurement unit, and a high-definition video camera. In order to mount a sensor to a UAS, the device would need to be assisted with an appropriate gimbal and a suitable battery source for seamless data acquisition. von Bueren et al. (2015) deployed and compared data from four different cameras, including a high-resolution spectrometer using two different UASs to monitor ryegrass pastures. While UAS hardware is becoming both easy to use and cost efficient, pragmatic applications of UASs, such as volume of wood chips, need to be explored.

Wood chips are used in the production of medium-density fiberboard (MDF), which includes a wide range of sustainable forest products such as plywood, lumber, and panels. The volume of chips that enter a facility at the MDF plant is correlated with current and future production of MDF products. This requires an accurate estimate of the chip volume at the start of each month that allows a facility to calculate the existing inventory and make accurate manufacturing decisions, including those related to personnel, resources, and profitability. Current survey methods rely upon third-party ground surveyors. Surveyors gather data on the chip pile at defined intervals, using a GPS and volumetric surveying equipment to recreate the pile in computer software for estimating the pile volume. Typically, it takes ground surveyors more than an hour to collect the needed data, and another hour to process the measurements digitally to estimate a final volume. Application of UASs may help overcome these constraints in estimating the volume of wood chips.

The objective of this study is to examine the use of a consumer-grade DJI Phantom 3 Professional quadcopter for estimating wood chip volume. We collected UAS photos from multiple positions and then analyzed them using various SfM programs. From this, we outline a workflow for generating reliable estimates of wood chip volume. Since the ground surveyors are contract employees and prohibited from revealing specifics, the exact time and cost for the ground survey and estimation of wood chip volume is unknown for this study.

7.2 METHODS

7.2.1 Study Area

The wood chip pile examined in this study is located on an ARAUCO MDF plant (a private enterprise) in Moncure, Chatham County, North Carolina (Figure 7.1). Transfer trucks arrive at the plant daily to empty their more than 25 tons of cargo (both hardwood and softwood chips) for refinement into MDF board. Chips are unloaded via a hydraulic lift, and transported into an appropriate storage bay for categorization using an overhead conveyor belt. At the time of this study, there were four bays for the separation of hardwood and softwood chips.

7.2.2 Unmanned Aerial Systems Logistics

This study examined the use of a consumer-grade DJI Phantom 3 Professional quadcopter for photo acquisition of chip volume. The consumer UAS was equipped with a Sony Exmor RGB 12.4 megapixel camera with a 94° field of view. The focal length of this camera is 2.8 mm with an ISO of 100–1600. The quadcopter was flown at a speed of 4.7 mph. Shutter speeds can range from 8 seconds to 1/8000 of a second. In this study, we used a shutter speed of 1/8000 of a second. Each imagery contained 4000 × 3000 pixels and was acquired in JPEG format. The DJI Phantom 3 has an integrated camera and records the spatial coordinates of each photo automatically in the exchangeable image file format. The average viable flying time per battery is around 18 minutes, with a horizontal range limited by Federal Aviation Administration (FAA) line of sight regulations and a maximum legal altitude of 400

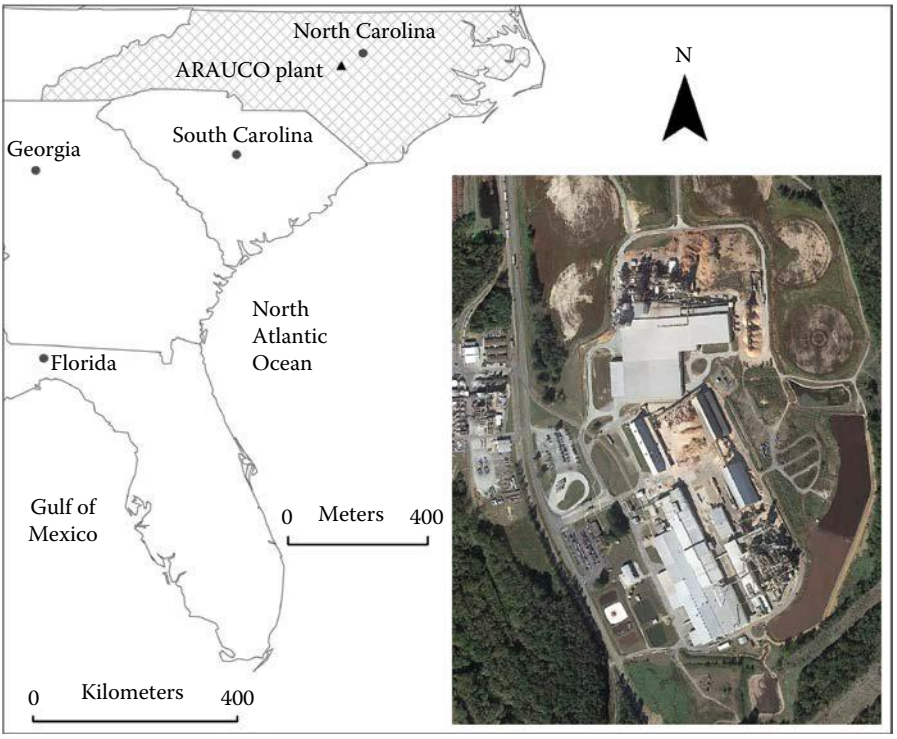


FIGURE 7.1 Study site with wood chip piles and hydraulic lift station at ARAUCO medium-density fiberboard plant in Moncure, North Carolina.

feet above ground level (AGL). The autonomous flight Map Pilot application by Maps Made Easy was used to capture nadir images. The Map Pilot application is one of the few that will allow for specific user customization regarding image overlap, battery management, photo capture, flight speed, and multibattery missions. The Phantom 3 can fly over approximately 100 acres at maximum legal altitude using approximately 80% forward and 70% side overlap.

7.2.3 Photo Acquisition

We collected UAS data once in August and once in September 2016. We collected aerial photos immediately after the ground surveyors finished their data collection (~8:30 a.m.), to ensure consistency in the amount of wood chips (Figure 7.2). We did not acquire UAS data at the same time as the ground surveys, in order to eliminate any possibility of bias in the data. The UAS was controlled from two main positions that kept the pilot and visual observer out of the way of transfer trucks and moving equipment (Figure 7.3). Because the sun was still rising at this time in the morning, the western side of the wood chip pile experienced shadow effects during the second survey. Figure 7.4 shows the area of the chip pile that

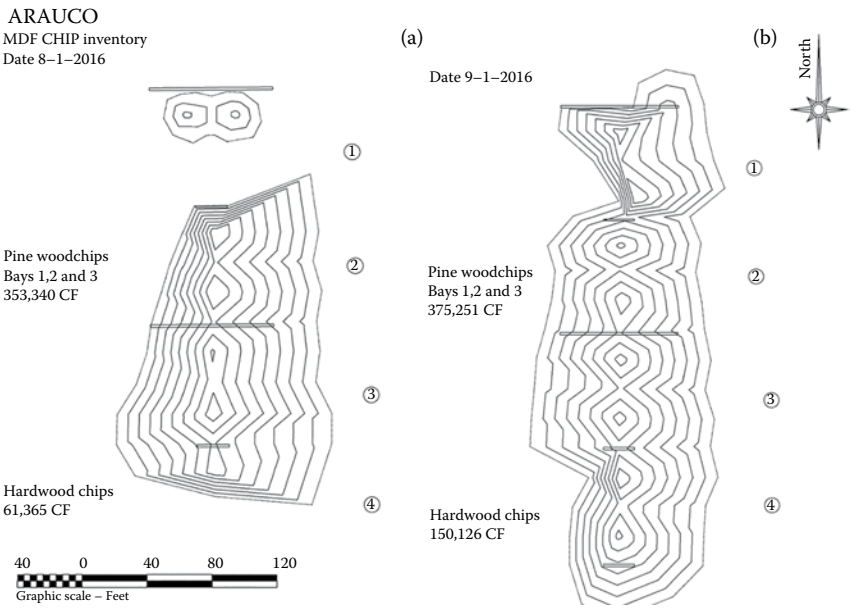


FIGURE 7.2 Wood chip pile topographic model produced by the ground survey crew for (a) August, and (b) September 2016.

was affected by shadows from the low sun angle in September. The inventory day in August was slightly overcast, and the sun was not visible, resulting in imagery without shadowing effects.

Prior to each survey, we required an average setup time of less than 5 minutes to visually inspect the aircraft, perform a hardware check on the DJI system, calibrate the compass, and designate the GPS return to home point location (within the DJI GO application). After initial setup, the next step was to acquire photos using the Map Pilot application by Maps Made Easy. This autonomous flight software was used to generate a specific flight plan for the August and September surveys in order to have unique nadir capture locations. The percentage of overlap and the flying altitude were consistently 80% and 130 feet AGL, respectively. At this altitude, there were no physical obstructions, and the resulting ground surface resolution was approximately 0.70 inches. Users can customize the Map Pilot application on an iOS tablet or smart phone, with or without the aid of cellular service or Wi-Fi. At present, the Map Pilot app is only available through Apple Inc. products; therefore, we used the iPad mini tablet for visualizing real-time video during UAS operations. In this case, we used a smart phone with a personal hot spot, which enabled background satellite imaging to aid in flight line creation on the iPad mini. The percent overlap in nadir photography is critical to matching pixel values, and greater accuracy is possible from a higher percentage of overlap.

We used the autonomous flight software by Maps Made Easy to capture photos in August and September from the nadir position. The DJI GO application was operated to acquire photos of approximately 45-degree oblique in September (Figure 7.5). The

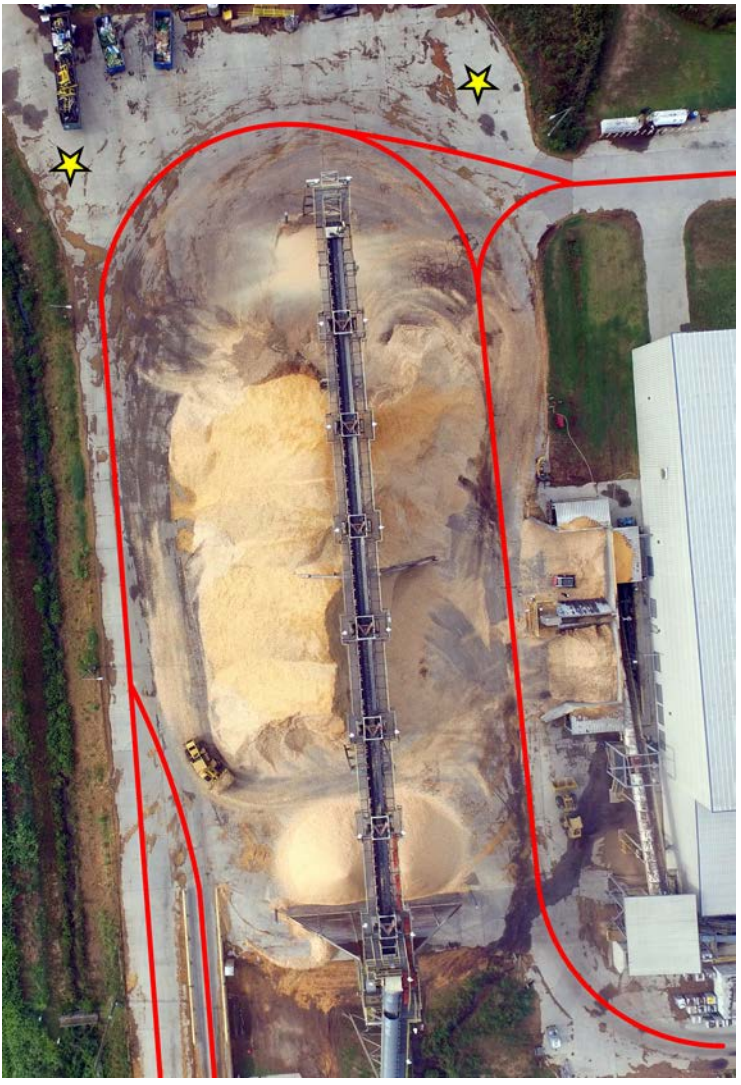


FIGURE 7.3 Mosaic generated using Agisoft PhotoScan of the study site, with an overlay of UAS pilot positions (star) and equipment traffic areas (thick line).

DJI Go flight application was operated in real time, with the UAS pilot controlling photo capture, platform altitude, and sensor angle. Obliques were acquired from all sides of the chip pile, at various percentages of overlap, since there is no way to control overlap in the DJI GO application. A 60% side-lap in oblique photos was desired because there were additional photos to add detail in areas the nadir sensor could not capture. Flying altitudes ranged from 0 to 100 feet AGL for capturing the oblique imagery. Total photo acquisition time from setup to breakdown of the DJI Phantom was approximately 20 minutes for each survey.

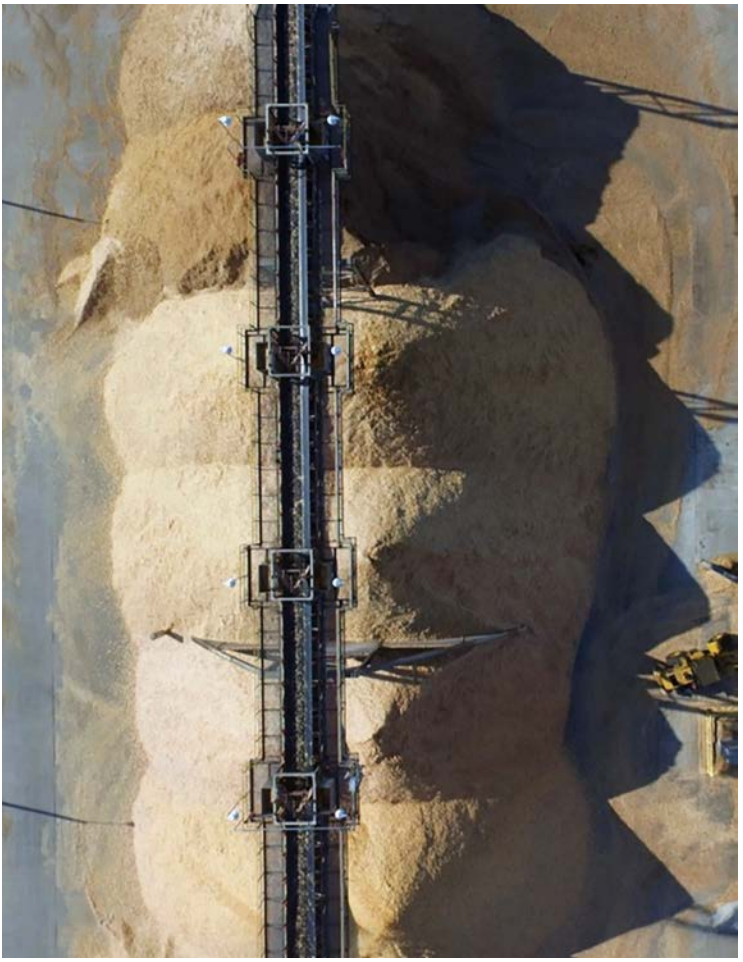


FIGURE 7.4 Westward shadow effect on the wood chip pile in the storage bay area of the study site.

7.2.4 UAS Imagery Processing

We analyzed each data set from the surveys using three different SfM programs: (1) Agisoft PhotoScan Pro, (2) Pix4Dmapper Pro, and (3) Maps Made Easy (Table 7.1). Agisoft PhotoScan Pro and Pix4Dmapper Pro are licensed desktop programs. They require users to upload, process, and edit the 3D point cloud on their own computer. Pix4D also permits license holders to use online image processing, but we did not investigate that option in this study. The Maps Made Easy online image processor is a pay as you go service that allows the user to specify which images to upload, and then cloud processing begins.

We preprocessed data prior to volume estimation using Agisoft PhotoScan Pro and Pix4Dmapper Pro to analyze the correct points within each dense point cloud.

Area:	2.14 acres
Distance:	0.85 km
Max speed:	9.0 mph
Duration:	4 minutes 7 seconds
Batteries:	1
Imagery:	57
Points:	Free
Storage	0.28 GB
Altitude:	131 feet
Resolution:	0.7 inches/pixel

FIGURE 7.5 Map Pilot flight plan and a screenshot of the real-time unmanned aerial system flight lines while using the application.

TABLE 7.1
SfM Program Costs

	Cost Per Unit License		
	Agisoft PhotoScan Pro	**Pix4Dmapper Pro**	**Maps Made Easy**
Professional	$3499.00	$3500.00[a] to $8700.00	$0.022 to $0.067/acre (pay-as-you-go-system)[b]
Student	$549.00	$1990.00	—

[a] Desktop yearly license.
[b] https://www.mapsmadeeasy.com/pricing.

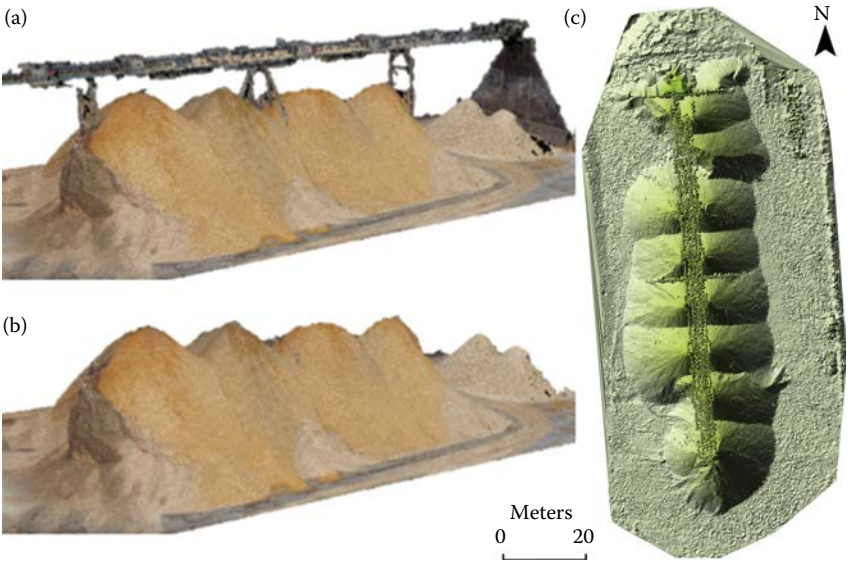


FIGURE 7.6 3D point cloud of wood chip pile 1.1 (a) with the conveyor belt, (b) without the conveyor belt, and (c) a surface model of wood chip pile with the conveyor belt.

For example, we masked the conveyor belt out of the 3D point clouds for accurately estimating the chip volume using nadir and oblique photos (Figure 7.6). We estimated wood chip volumes with the conveyor belt, using nadir photos for August and September surveys, and used nadir and oblique photos for the September survey. While estimating wood chip volume with the conveyor belt does not allow for the comparison of volumes between the UAS and ground surveyor, it helps to compare the performance of each SfM program. We also estimated wood chip volume without the conveyor belt to compare estimates generated from SfM programs (i.e., Agisoft PhotoScan Pro and Pix4Dmapper Pro) and ground-surveyed measurements. Maps Made Easy does not allow customization of the point cloud, and therefore we did not estimate volume. We repeated the entire process five times to eliminate biased results from each calculation. The average volume and standard error of the measurements were used for each method. Images of the woodchip pile that contained shadow changed the color of the woodchip pile but not the surface texture because such shadow did not affect the quality of surface models. Figure 7.7 represents a general SfM workflow.

Agisoft PhotoScan Pro has a simple graphical user interface (GUI) that allows users to select each step and the associated parameters. The workflow is quite simple,

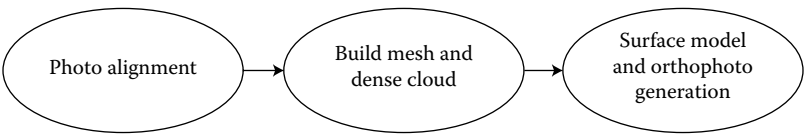


FIGURE 7.7 General structure for motion workflow.

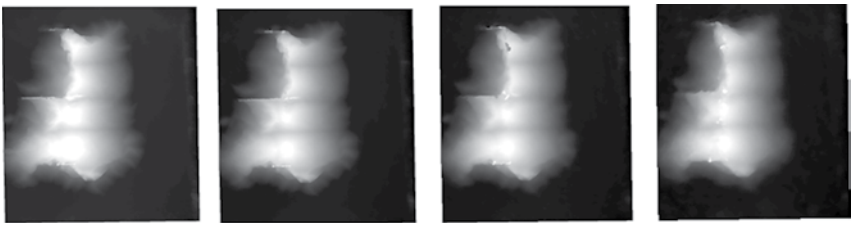


FIGURE 7.8 Agisoft PhotoScan Pro dense point cloud quality setting comparison for DSM generation.

and can be batch processed after the initial setup. The final step involves exporting the desired results to the correct file directory. Within Agisoft PhotoScan Pro, we used high, medium, low, and the lowest quality parameters to determine the best setting for creating the digital surface model (DSM). The medium-quality DSM produced the most accurate representation of the chip pile surfaces in Agisoft PhotoScan Pro (Figure 7.8). The high-quality processing required a substantial time to stitch points that were considered noisy or above the reference plane. The low and lowest settings appeared to provide incorrect estimates. We also selected medium-quality parameters for creating the DSM within Pix4Dmapper Pro. The Pix4Dmapper Pro comes with a GUI and requires three main processing steps. Steps can be generated individually or batch processed together if selected. The initial phase includes the image capture locations overlaid on satellite imagery, which is helpful to display the data set and to check for photo inconsistencies. Pix4Dmapper Pro and Agisoft PhotoScan Pro allow users to customize the SfM parameters to generate a point cloud that contains the highest accuracy possible. The Maps Made Easy application provides systematic instructions on its webpage that require no prior image analysis experience because there is little customization of processing parameters. Once photos are uploaded and export parameters are selected, an email is sent to the user with the suggested time of completion, and the user is again notified via email when the process is completed.

7.3 RESULTS

Volume estimates varied across the applied SfM programs. The Pix4Dmapper Pro program produced the smallest standard errors in estimating wood chip volume, both with and without the conveyor belt, using nadir photos for August and September UAS surveys (128 and 26 m^3; 83 and 14 m^3, respectively) (Table 7.2). Maps Made Easy produced slightly higher volume estimates with the conveyor belt for August and September (3.5% and 2%, respectively), compared to the Pix4Dmapper Pro program, but at the cost of the highest standard errors (380 and 523 m^3), whereas Agisoft PhotoScan Pro produced volume estimates approximately 19% higher (Table 7.2). The three SfM programs produced the highest volume estimates and standard errors using a combination of nadir and oblique photos for September (Table 7.3). These volume estimates were far from those of ground-based survey methods. When we removed the conveyor belt from the 3D point cloud, we observed volume estimates similar to

TABLE 7.2
Estimates of Wood Chip Pile Volume with Conveyor Belt from Nadir Photos, Using Three Structure for Motion Programs for August and September

	August		September	
	Volume (m^3)	Standard Error	Volume (m^3)	Standard Error
Ground Survey	11,743.1	—	14,876.5	—
Agisoft PhotoScan Pro	15,720.7	147.8	20,128.7	57.6
Pix4Dmapper Pro	13,252.8	128.1	19,468.8	25.7
Maps Made Easy	13,708.1	380.0	19,859.5	522.5

TABLE 7.3
Estimates of Wood Chip Pile Volume with Conveyor Belt from Nadir Photos for September, Using Three Structure for Motion Programs

	September	
	Volume (m^3)	Standard Error
Ground Survey	14,876.5	—
Agisoft PhotoScan Pro	20,362.6	113.4
Pix4Dmapper Pro	19,459.4	192.6
Maps Made Easy	20,750.4	668.3

those of the ground survey, while standard errors were the lowest (Table 7.4). The Pix4Dmapper Pro program produced the lowest standard error (14 m^3), followed by the Agisoft PhotoScan Pro programs (27 m^3 without the conveyor belt). Volume estimates from these two SfM programs were 8% to 12% lower compared to the ground survey (Table 7.4). We found that the standard error of volume estimates decreased as the number of photos used in the image processing increased. The SfM programs that produced the lowest average standard error used the highest number of images, such as the Pix4Dmapper Pro program.

7.4 DISCUSSION

The UAS allowed for seamless data collection amid heavy equipment operating around the chip pile. Since there is minimal site disturbance through the UAS data collection procedure, UAS can be used on a regular basis for monitoring wood chip pile volumes in order to update inventory. The Pix4Dmapper Pro program outperformed the other two SfM programs, with the lowest standard error and volume estimates, similar to those of the ground survey. The Maps Made Easy application was much simpler and more user friendly. It is important to keep in mind that the data collected via ground

TABLE 7.4
Estimates of Wood Chip Pile Volume without Conveyor Belt from Nadir and Oblique Photos for August and September, Using the Agisoft PhotoScan Pro and Pix4Dmapper Pro

	August		September	
	Volume (m^3)	Standard Error	Volume (m^3)	Standard Error
Ground Survey	11,743.10	—	14,876.50	—
Agisoft PhotoScan Pro	10,358.70	193.80	16,626.80	27.40
Pix4Dmapper Pro	10,769.00	83.00	16,231.50	13.80
Maps Made Easy[a]	—	—	—	—

[a] Preprocessing, such as removal of unwanted 3D points, is not available in this program.

survey is only an estimation and could be biased. Besides generating and saving multiple volumes at once, SfM programs provide high-resolution orthophotos and 3D interactive models that could be used for additional purposes. By-products, such as the mosaic generated by SfM techniques when estimating volumes, may help in inspecting conveyor belts and auditing infrastructure.

Photo acquisition should follow the procedure to minimize the shadow effect. While shadows in wood chip pile photos did not affect the quality of the DSM, they did affect the red, green, and blue (RGB) values of photos. UAS pilots should account for the clarity of the atmosphere, sun angle, wind speed, and changes in camera exposure to acquire usable photos in a timely manner. If these parameters are not addressed, the photos may be blurry or out of focus and unable to be stitched together. We successfully imported imagery collected by the DJI Phantom 3 Pro into multiple SfM programs to generate and view high-resolution 3D point clouds that were spatially referenced. SfM allows for the mapping of areas at resolutions that were previously not possible, and is a valuable aid to natural resource management. However, each program has its benefits and drawbacks, and should be used accordingly to maximize the objectives of a project. The Agisoft PhotoScan Pro application performs photogrammetric processing of digital images and generates 3D spatial data. This provides an easy and rapid calculation of volumes, but it is difficult to get exact results for large areas with specific boundaries. Tools that are used for volume delineation do not allow the user to select exact pixel locations, and can cause confusion if irregular or vast areas are to be calculated. The general workflow is quite simple, and photos can be batch processed after a quick setup. Generated mosaics are of high quality, with uniform color across the entire project area. The Pix4Dmapper Pro application uses three main processing steps that generate a simple workflow. It provides accurate clipping and volume extraction methods that allow the user to save and view multiple volume measurements at once. Another appealing feature that Pix4Dmapper Pro offers is the ability to save the entire workflow as a template; therefore, it can be easily loaded for a new project. This decreases the possibility of

error when selecting the SfM parameters, and allows for time efficiency with repeated image processing. A GCP identification and placement tool is constructed through an accurate and easy-to-follow interface.

The Maps Made Easy application operates through a system that requires the user to pay for the image processing, using a point scale associated with the amount and resolution quality of the photos. There is a direct correlation between resolution and suggested computing power; therefore, Maps Made Easy accounts for this input variable. The current version permits projects that are less than 250 points to be processed free of cost, to allow users to test the system and run sample areas before processing large data sets. Each chip pile data set did not exceed 250 points, and was able to be processed without cost. Maps Made Easy is the easiest method of processing because it requires very little user input and knowledge. Volume areas are delineated through a 2D interactive window, where the user selects volume boundaries via the generated mosaic. However, users do not have the ability to edit the point cloud, making volume calculations include all space below overhead features. Project deliverables are easily accessible to view and download in 2D and 3D formats.

SfM programs are able to account for minimal changes in topography. Desktop-based programs, such as Agisoft PhotoScan Pro and Pix4Dmapper Pro, are optimal for acquiring and preprocessing data that include removal of 3D points of unwanted overhead features that may obstruct the view of the UAS camera. If the area surveyed does not have any overhead features, a program such as Maps Made Easy is another outlet at a low cost of operation. The front-end loader that is visible in the eastern side of the mosaic was able to operate continually when the UAS survey was being conducted (Figure 7.2), while this was not the case with the ground survey. Therefore, the UAS pilot was able to conduct inventory from a safe distance away from moving equipment compared to ground surveys, decreasing the possibility of an accident occurring during the UAS survey. The ability to provide a standard error, specifically tailored to each data set, with UAS measurements can offer additional data regarding the volume estimation in comparison with a typical ground survey method.

7.5 CONCLUSION

This study demonstrates that postprocessing of UAS imagery using multiple software approaches for estimating wood chip pile volume is repeatable and comparable to estimates produced by ground surveys. The point cloud generated through only nadir photos had a larger standard error when compared to those images gathered in nadir and oblique positions. Oblique photos captured the vertical texture of the pile in a way that the nadir sensor was unable to capture. Photos captured with a tilted sensor acquired at various altitudes and angles mapped different aspects and surfaces that were unseen in nadir photography (especially underneath the conveyor belt). We found that the standard error of the chip pile volume decreases as the number of photos used in the image processing increases. The data set with the highest number of images had the lowest average standard error. A direct comparison of known wood chip volumes to those estimated by SfM programs is needed to improve the accuracy of volume estimates. This will inform us of the bounds and errors of each used program, and their ability to successfully generate an accurate 3D point cloud.

Adding survey GCPs to tie the models to true geographic coordinates may increase the accuracy of estimated volumes and should be taken into consideration.

7.6 LIMITATIONS AND FUTURE RESEARCH

As per the FAA regulations, each UAS pilot has to have an FAA UAS license. They are typically bound to fly at 400 feet AGL or below, which limits the mappable area. If the UAS is piloted within 5 nautical miles of an airport, the pilot is required to notify the appropriate air traffic control tower to allow for increased situational awareness of local manned aircraft. The pilot is also limited by visual line of sight regulations that instruct the pilot or associated observers to physically keep visual contact with a UAS platform at all times. This can restrict the area flown when operated in forested areas or between tall buildings, where the crew is unable to keep unobstructed view of a UAS. While these regulations are set for safety reasons, they limit the capabilities of UASs because current platforms have the ability to map at greater altitudes due to high-resolution integrated cameras. UAS pilots also have to keep a very close eye on the geographic location of the flights because any slight change in weather can affect the quality of data. Most platforms are not waterproof, but can fly in considerable amounts of wind. High winds do require more rotor thrust and decrease the overall time in the air due to battery consumption. While these limitations restrict usability of UASs in large area applications, cost effectiveness is one of the key drivers that would open doors for UASs in both research and commercial applications (Canis 2015).

Civil and commercial applications include aerial reconnaissance, search and rescue, survey of forest, timber production, crop production, disaster damage estimation, disaster management, agricultural activities, telecommunications, oil and gas exploration, and geophysical surveys (Zhang et al. 2016). The ease of UAS deployment will allow repeated mapping at a scale that was previously time and cost consuming; for example, frequent forest survey for the detection of windblown and nonproductive areas (Cunliffe et al. 2016; Nasi et al. 2015). Use of SfM techniques will drastically change how timber can be inventoried for modeling timber height. Since repeated surveys and rapid response using UASs are user friendly, it can be used to quickly access disaster areas for damage assessments and management using 3D mapping. With the aid of high-resolution cameras and brushless rotors, visualizations of potentially hazardous sites will increase site awareness without having to place humans in contact with harmful substances (Yuan et al. 2015). RGB and infrared-equipped UAS platforms are suitable for real-time data collection on prescribed or wildfire situations. Lightweight and portable UASs can be carried by ground crews and quickly deployed if aerial views are needed to monitor fire movement and direction.

The applicability of UASs due to low operational and material costs, control of spatial and temporal resolution, and the absence of risk to crews will also open doors for both applied and developmental research (Tang and Shao 2015). For example, because of UASs' flexibility in the temporal dimensions of data acquisition, UASs can be used to study phenological variability due to climate change at broad scales. While data collection from aircraft is costly, it overcomes the limitations of traditional satellite sensors where data acquisition is bounded by the orbital characteristics of

the platform (Weiss and Baret 2017). Therefore, the capabilities of UASs can help in planning flights very close to phenological stages important for detection of plant stress, impacts of climate change, or preharvest characterization (Burkart et al. 2017). These UAS flights can be conducted using a variety of sensors, such as RGB, multispectral, near-infrared, hyperspectral, or LiDAR. In addition, such portable sensors may open doors to new research directions, such as system calibration, sensor development, data registration, and data fusion with onboard cameras.

REFERENCES

Berni, J.A.J., Zarco-Tejada, P.J., Suarez, L., & Fereres, E. 2009. Thermal and Narrowband Multispectral Remote Sensing for Vegetation Monitoring from an Unmanned Aerial Vehicle. *IEEE Transactions on Geoscience and Remote Sensing*, 47, 722–738.

Breckenridge, R.P., Dakins, M., Bunting, S., Harbour, J.L., & White, S. 2011. Comparison of Unmanned Aerial Vehicle Platforms for Assessing Vegetation Cover in Sagebrush Steppe Ecosystems. *Rangeland Ecology & Management*, 64, 521–532.

Burkart, A., Hecht, V.L., Kraska, T., & Rascher, U. 2017. Phenological Analysis of Unmanned Aerial Vehicle Based Time Series of Barley Imagery with High Temporal Resolution. *Precision Agriculture*, 19, 1–13.

Canis, B. 2015. *Unmanned aircraft systems (UAS): Commercial outlook for a new industry.* Congressional Research Service, Washington, DC.

Chiabrando, F., Marenchino, D., Nex, F., Piatti, D., & Rinaudo, F. 2009. UAV Systems for Photogrammetric Recording in Archaeological Areas: Problems, Suggestions, and Products. *Remote Sensing for Environmental Monitoring, GIS Applications, and Geology IX, 7478.*

Cunliffe, A.M., Brazier, R.E., & Anderson, K. 2016. Ultra-Fine Grain Landscape-Scale Quantification of Dryland Vegetation Structure with Drone-Acquired Structure-from-Motion Photogrammetry. *Remote Sensing of Environment*, 183, 129–143.

Dandois, J.P., Olano, M., & Ellis, E.C. 2015. Optimal Altitude, Overlap, and Weather Conditions for Computer Vision UAV Estimates of Forest Structure. *Remote Sensing*, 7, 13895–13920.

Gates, D. 2015. Drone Tracks Fire Hot Spots in Successful Olympic Forest Test. *Seattle Times*, http://www.seattletimes.com/business/boeing-aerospace/drone-tracks-fire-hotspots-in-successful-national-park-test/.

Goddemeier, N., Daniel, K., & Wietfeld, C. 2010. *Coverage Evaluation of Wireless Networks for Unmanned Aerial Systems*. 2010 IEEE Globecom Workshops, 1760–1765.

Higgins, S. 2016. PhoDAR vs. LiDAR—Is This New Technology a Giant Killer? *SPAR3D*, http://www.spar3d.com/news/software/phodar-vs-lidar-new-technology-giant-killer/.

Nasi, R., Honkavaara, E., Lyytikainen-Saarenmaa, P., Blomqvist, M., Litkey, P., Hakala, T., Viljanen, N., Kantola, T., Tanhuanpaa, T., & Holopainen, M. 2015. Using UAV-Based Photogrammetry and Hyperspectral Imaging for Mapping Bark Beetle Damage at Tree-Level. *Remote Sensing*, 7, 15467–15493.

Tang, L.N., & Shao, G.F. 2015. Drone Remote Sensing for Forestry Research and Practices. *Journal of Forestry Research*, 26, 791–797.

Themistocleous, K., Ioannides, M., Agapiou, A., & Hadjimitsis, D.G. 2015. The Methodology of Documenting Cultural Heritage Sites Using Photogrammetry, UAV and 3D Printing Techniques: The Case Study Of Asinou Church In Cyprus. *Third International Conference on Remote Sensing and Geoinformation of the Environment (Rscy2015), 9535.*

von Bueren, S.K., Burkart, A., Hueni, A., Rascher, U., Tuohy, M.P., & Yule, I.J. 2015. Deploying Four Optical UAV-Based Sensors over Grassland: Challenges and Limitations. *Biogeosciences*, 12, 163–175.

Wallace, L., Lucieer, A., Watson, C., & Turner, D. 2012. Development of a UAV-LiDAR System with Application to Forest Inventory. *Remote Sensing*, 4, 1519–1543.

Weiss, M., & Baret, F. 2017. Using 3D Point Clouds Derived from UAV RGB Imagery to Describe Vineyard 3D Macro-Structure. *Remote Sensing*, 9, 111.

Yuan, C., Zhang, Y.M., & Liu, Z.X. 2015. A Survey on Technologies for Automatic Forest Fire Monitoring, Detection, and Fighting Using Unmanned Aerial Vehicles and Remote Sensing Techniques. *Canadian Journal of Forest Research*, 45, 783–792.

Zhang, J., Hu, J.B., Lian, J.Y., Fan, Z.J., Ouyang, X.J., & Ye, W.H. 2016. Seeing the Forest from Drones: Testing the Potential of Lightweight Drones as a Tool for Long-Term Forest Monitoring. *Biological Conservation*, 198, 60–69.

Zhou, G., and Zang, D. 2007. Civil UAV system for earth observation, *IEEE International Geoscience and Remote Sensing Symposium*, Barcelona, pp. 5319–5322. doi: 10.1109/IGARSS.2007.4424063.

8 A Workflow to Quantify the Carbon Storage in Urban Trees Using Multispectral ALS Data

Xinqu Chen and Jonathan Li

CONTENTS

8.1 INTRODUCTION

With the continuing growth of the global population, urbanization has become an inevitable trend. Urban expansion accelerates with the advancements of socioeconomic activities and population growth. Today, more than 50% of the global population resides in cities, and by 2050 urban areas will hold up to another 2.5 billion people, equivalent to 66% of the global population (United Nations, 2014). Intensive urban developments and economic activities increase energy consumption and result in greenhouse gas emissions (Corfee-Morlot et al., 2009; Satterthwaite, 2009; Dodman et al., 2012). As urban land covers expand, direct losses of vegetation cover also occur. Without vegetation covers acting as the largest carbon sink, deforestation becomes the second largest contributor of greenhouse gases (Johnson, 2009).

Increased awareness of high energy use and urban heat islands has led urban municipalities to put conservation efforts into implementing strategies that can protect the urban climate and energy use at the city scale. Urban vegetation has drawn the direct attention of city planners and policy makers, considering the importance of trees in in urban climate modification and energy conservation. In the context of Canada's climate, annual cooling energy use can be reduced 10%–19% by planting vegetation proximate to the houses and increasing the albedo of urban surfaces (Xu et al., 2012; Sawka et al., 2013). City wide, urban trees reduce air pollution through direct dry deposition, and also influence the cooling of the ambient temperature, which slows smog formation. Vegetation contributes the largest proportion of carbon storage, which in turn reduces the rate of climate warming and urban heat islands (Davies et al., 2011). Urban trees both sequester CO_2 and store excess carbon in biomass (71% of total urban carbon storage), which significantly influences environmental quality and human health (Donovan & Butry, 2010; Roy et al., 2012). Preserving carbon storage and improving green space infrastructure in urban areas has significant environmental benefits (van den Berg et al., 2015). Consequently, estimation and monitoring of urban carbon stocks and green space becomes important indeed. In the literature, the carbon content stored in individual trees can be assessed through aboveground dry-weight biomass calculation using allometric equations. Dendrometric parameters, such as individual tree height or crown diameters, are generally used in the allometric equations to derive diameter at breast height (DBH). Dry biomass is then calculated using the allometric model, with DBH as the input, and further transformed to carbon storage with a conversion rate around 0.5 (Lieth, 1963; Whittaker & Likens, 1973).

With the development of LiDAR technology, airborne laser scanning (ALS) and very-high-resolution multispectral imagery have proven to be promising tools for the derivation of dendrometric parameters. Previous studies have successfully extracted detailed vegetation covers from very-high-resolution multispectral data, such as QuickBird, and applied it to the ALS-derived canopy height model (CHM) for dendrometric measurements and biomass estimation in both forestry and urban studies (Huang et al., 2013; Schreyer et al., 2014; Pasher et al., 2014; Raciti et al., 2014; Sousa et al., 2015). Raciti et al. (2014) estimated carbon storage in urban trees for the City of Boston by developing an ALS-height-only regression model to estimate carbon storage across the city. A total of 404 accurately segmented tree crowns from the normalized digital surface model (nDSM) were split into 284 samples for model fitting and 120 samples for validation. The reason for using a simple linear regression of tree biomass and height was to avoid the influence of crown segmentation results. An R^2 of 0.79 was found between field-estimated biomass and model-predicted biomass. Schreyer et al. (2014) estimated the carbon storage in urban trees, and its distribution was extrapolated to the entire city of Berlin in terms of land use types. This study did not propose a region-specific allometric model for the study area, but applied the ALS-DBH model developed by Zhao et al. (2009) and a carbon allometric model using DBH as the only independent variable. A total of 87% of the modeled DBH showed an underestimation, which was further calibrated by a weighted arithmetic average DBH. The carbon storage in urban trees was calculated as half of the model-based biomass, regardless of the genus. Meanwhile, the crown base height was assumed to be half of the ALS-derived height, and the crown width was calculated in 16 directions

with a series of criteria, instead of the conventional estimation of crown width in four directions as proposed by Popescu et al. (2003) or as used by Xu et al. (2014).

With the emerging multispectral ALS technology, it became possible to obtain both the range and multiple reflectance data from a single data source. The first commercial multispectral ALS system, Optech Titan, released by Teledyne Optech Inc., has integrated three laser channels at wavelengths of 532, 1064, and 1550 nm, respectively. These three channels produce independent scan lines by sending pulses with separate forward angles (the near-infrared [NIR] channel has a 0° forward angle; the short-wave infrared [SWIR] channel has a 3.5° forward angle; the green channel has a 7.0° forward angle). This emerging ALS system showed great potential in land cover mapping without the aid of passive multispectral images (Bakuła, 2015; Wichmann et al., 2015). This active laser system can largely avoid those factors commonly associated with passive optical sensors, such as weather conditions and shadow effects. However, the potential of ALS intensity currently remains undervalued, and the applications of this newly released system are at an early stage of development.

Given the current state of development of tree inventory with typical ALS data, this chapter aims to explore the feasibility of multispectral ALS range and intensity data in carbon storage estimation. To achieve this, vegetation covers are first classified based on multispectral ALS range and intensity data by applying a support vector machine (SVM) classifier. Second, dendrometric parameters such as tree height and crown diameter are derived in order to establish an allometric relationship between ALS-derived measurements (tree height and crown diameter) and the field-measured parameter (DBH) through regression modeling. Last, this study quantifies the carbon storage in urban trees for the Town of Whitchurch-Stouffville, Ontario, Canada.

8.2 STUDY AREA AND DATA SOURCES

The study area is located in the Town of Whitchurch-Stouffville, Ontario, Canada, which is a municipality in the Greater Toronto Area (Figure 8.1). The area is characterized by a typical residential landscape that contains two water bodies (Musselman Lake and Windsor Lake) and three land cover types: residential area, open area (grassland and woody area), and park and recreation (DMTI Spatial, 2015). The residential area consists of single detached dwellings with mature street and backyard trees planted at least 10 years ago (Google Earth, 2004). The tree species in the study area include deciduous trees such as maple, ash, oak, elm, black cherry, basswood, and conifers (Brook McIlroy Inc., 2002).

Two multispectral ALS images were acquired by the Teledyne Optech Titan multispectral ALS system on July 2, 2015. The multispectral ALS data sets had two flight lines that covered and intersected at the study area. The flight altitudes were above 1000 m with a pulse frequency of 100 kHz for each channel, yielding an average point spacing of 0.8 m per point and an average point density of 7.7 point/m^2 (Table 8.1).

Field data were collected on February 9, 2016 (Figure 8.2). A total of 40 trees were sampled in the field that contained four attributes (i.e. height, DBH, crown diameter, and biomass) for each single tree. Tree heights were measured in meters using a hypsometer. DBH was measured in centimeters with a diameter tape. Since the field measurement was conducted during the leaf-off season, the third attribute,

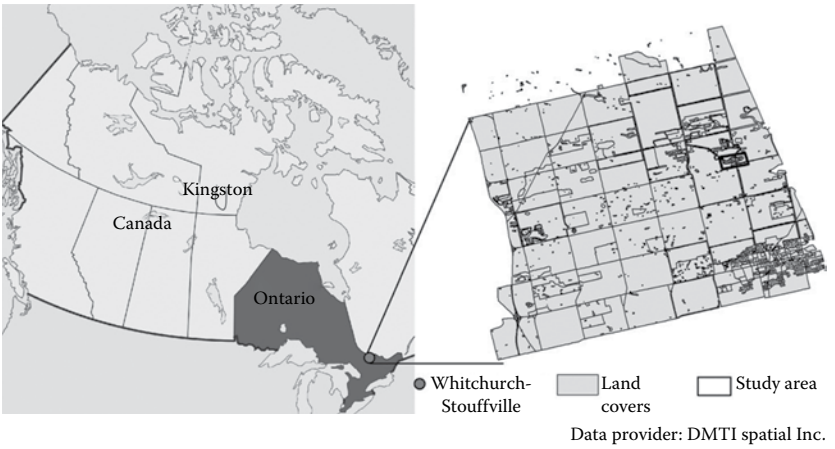


FIGURE 8.1 Map of study area.

TABLE 8.1
Summary of the Multispectral Titan ALS Data Set

Characteristic	Value
Laser channels	532 nm; 1064 nm; 1550 nm
Field of view (FOV)	30°
Pulse repetition frequency (PRF)	100 kHz per channel; 300 kHz in total
Flight height	Strip 1: 1030 m
	Strip 2: 1043 m
Flight speed	69–70 m/s
Flight date	July 2, 2015
Number of ALS strips	2
Number of returns	4
Number of points	Channel (532 nm): 3,665,956
	Channel (1064 nm): 3,692,211
	Channel (1550 nm): 3,072,895
Average point density	7.7 points/m^2 in total

crown diameter, was measured using aerial photos from Google Earth. The crown diameter, in meters, is defined as the mean of the maximum crown diameter and the diameter measured at the direction perpendicular to the maximum, using the Ruler tool in Google Earth. The fourth attribute, single-tree carbon storage, is estimated by plugging the field-measured DBH and tree height into the Canadian national aboveground all-species biomass equations (Lambert et al., 2005). The equations calculate the dry aboveground biomass by relating tree height and DBH to each biomass component, such as wood, bark, and foliage, with a uniform relationship: Biomass $= \beta_1 \mathrm{DBH}^{\beta_2} H^{\beta_3}$; where β_1, β_2, and β_3 are the parameters generated for the all-species group with different values according to the tree

FIGURE 8.2 Location of 40 sampled trees in the field.

compartments. The carbon storage in sampled trees was then defined as half of the sum of the dry aboveground biomass in each compartment.

8.3 METHODOLOGY

The methodology of this study consisted of four parts: ALS data processing, vegetation isolation, dendrometric parameter extraction, and carbon storage modeling (Figure 8.3).

8.3.1 ALS Data Processing

After removing outliers and rectifying the ALS intensity values, point clouds from two flight strips were merged together; in total, three ALS point clouds were acquired by the Titan laser channels. Then each point cloud was rasterized into an intensity image with a ground resolution of 1 m. The pixel size was selected according to the point spacing of the data set. By selecting the pixel size close to the point spacing, most of the pixels can contain at least one point, and the vertical distribution of the points can be largely highlighted. In this way, points were grouped into 1 m grids and the pixel values were assigned by the mean intensity of the points within the grid. For the grids that had no point filled in, the grid values were interpolated linearly by searching the neighbors. Here, three mean intensity raster data were generated for channels 532, 1064, and 1550 nm, respectively.

Besides the generation of ALS intensity images, a digital surface model (DSM) and a digital terrain model (DTM) were created from the raw ALS by a ground-point filtering and rasterization process. The whole ALS data set was then classified into ground and nonground classes. DTM raster data were generated by rasterizing all the ground points into 1 m grids, based on the linear interpolation method. The DSM was generated in a similar way by using the nonground class, and the maximum

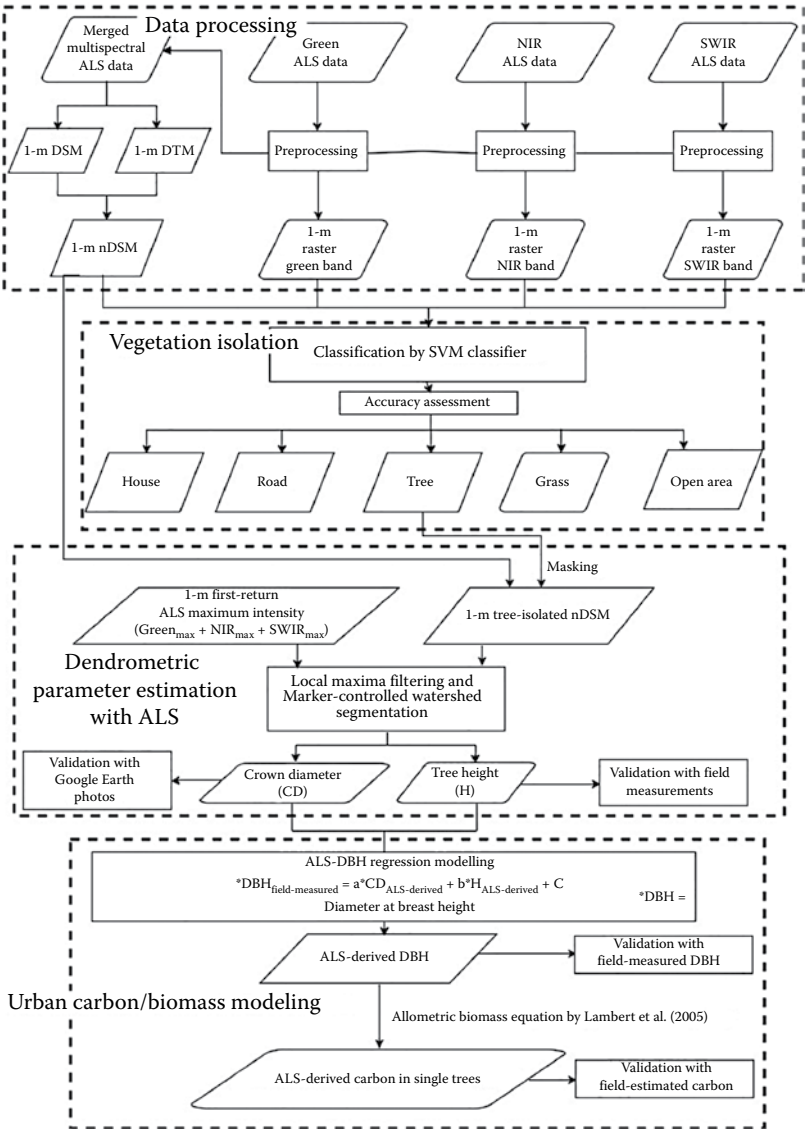


FIGURE 8.3 Workflow of the proposed methodology.

height within the grid was assigned to the pixel values. In this way, the points that represented the treetops could be largely reserved. Finally, DSM raster data were acquired by subtracting the DTM from the DSM.

8.3.2 Vegetation Isolation

Besides the multispectral intensity and nDSM data described in Section 8.3.1, two additional indices were derived as follows:

$$\text{pNDWI} = \frac{C_{\text{Green}} - C_{\text{NIR}}}{C_{\text{Green}} + C_{\text{NIR}}} \tag{8.1}$$

$$\text{pNDVI} = \frac{C_{\text{NIR}} - C_{\text{SWIR}}}{C_{\text{NIR}} + C_{\text{SWIR}}} \tag{8.2}$$

where C_{Green}, C_{NIR}, and C_{SWIR} refer to the laser channels at 532, 1064, and 1550 nm, respectively.

By visually examining the pseudo normalized difference water index (pNDWI) and pseudo normalized difference vegetation index (pNDVI) (Figure 8.4), both the pNDVI and pNDWI showed good discrimination for artificial objects. These two indices can facilitate the manual selection of training samples and work as the ancillary data in the classification process. The contribution of these two indices to the overall classification accuracy was analyzed. A total of six input data, including (1) green channel intensity, (2) NIR channel intensity, (3) SWIR channel intensity, (4) nDSM, (5) pNDWI, and (6) pNDVI, were generated and input into the classification scheme.

Because the study area was characterized by a simple residential landscape, it could be grouped into six land cover types, including water, house, road, grass, tree, and open area. However, due to the bathymetry capability of the channel at 532 nm, water points in the study area tended to have irregular intensity that was induced by the interaction of laser points with both the water surface and the organic matter underneath. Hence, water bodies including Musselman Lake, Windsor Lake, and one small water region were masked out of the data set, resulting in only five land cover types being trained and classified in this study. Interpreting land cover types from ALS intensity was not as easy as from passive optical images. Certain land cover types, such as grass and open area, could be distinguished from only one or two intensity data and could hardly be identified from the rest.

An SVM classifier was selected to perform land cover classification with multispectral ALS-derived data due to the popularity of SVM in typical ALS-related classification studies, making the classification result of this study comparable to previous studies.

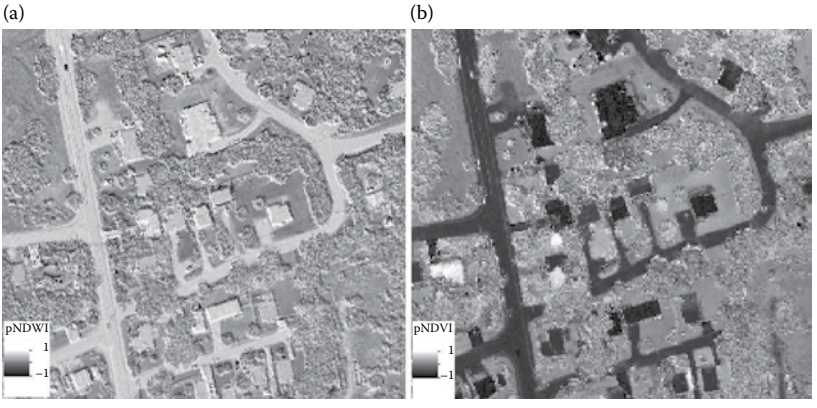


FIGURE 8.4 Examples of ALS-intensity-derived indices.

TABLE 8.2
Combinations of Input Data for SVM Classification

	Combination
Selection 1	Green + NIR + SWIR + nDSM + pNDWI+ pNDVI
Selection 2	Green + NIR + SWIR + nDSM
Selection 3	NIR + nDSM

The SVM classification was performed on three combinations of the input data, shown in Table 8.2. The selection of input data was mainly designed for showing the benefits of multispectral ALS in land cover classification (Selection 2) compared with typical ALS data (Selection 3). Meanwhile, Selection 1 was designed to examine the contribution of multispectral-intensity-derived indices in overall classification accuracy.

8.3.3 Dendrometric Parameter Estimation

The tree-isolated nDSM was also referred to as the CHM, a model displaying tree positions by tree crowns in a top-down view and storing the height values in pixels. A 3×3 local maxima filter was first employed on the CHM to detect treetops. A pixel with the highest value among its eight neighbors was defined as the treetop. To eliminate the commission errors associated with the local maxima filtering and detect the true treetop pixels, the local maxima in the CHM were further filtered by the ALS intensity data. The ALS intensity values were dependent not only on the reflectivity of the object, but also on the range between the sensor and the object. A true treetop pixel would have both high intensity and height values. Under this assumption, the sum of maximum intensity of first return is generated at 1 m for each channel. The clusters of local maxima in the CHM, which had more than 15 pixels together, were further extracted, and only the pixels that were also the local maxima in the maximum intensity layer were retained in the final treetop results. Previous studies relied on changing the window size and shape of the filter to refine the treetops (Chen et al., 2006; Zhao et al., 2009). These approaches were not suitable here because the CHM resolution (1 m) generated in this study was relatively coarse, so increasing the window size of the local maxima filter would result in excluding small tree crowns.

The marker-controlled watershed segmentation was applied to segment the CHM into individual tree crowns by defining the pre-detected treetops as the markers. In this way, every pre-detected treetop would have one closed segment. The performance of the segmentation in isolating tree crowns was evaluated by the absolute accuracy, calculated as

$$\text{Absolute accuracy}_{\text{tree isolation}} = \frac{n_{1,1}}{n_{\text{total}}} \tag{8.3}$$

where $n_{1,1}$ is the number of detected crown segments that have a one-to-one relationship to the ground truth; and n_{total} is the number of tree crowns in the ground truth. Tree

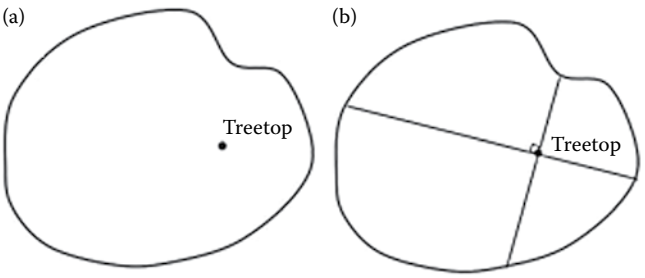


FIGURE 8.5 Measurement of tree height and crown diameter: (a) location of the pixel with the maximum value within a tree crown; (b) illustration of the crown diameter measurement.

height is defined as the average of the local maxima within each segment, and crown diameter is defined as the average of the maximum crown diameter passing through the center of the local maxima and the one measured at the perpendicular direction (Figure 8.5).

To evaluate the accuracy of the ALS-derived dendrometric parameters, the ALS-derived tree height and crown diameter were compared with the field measurements. The crown segments generated from the ALS data were matched with the 40 field-sampled trees, and the root-mean-square error (RMSE) and bias were calculated to compare the ALS-derived dendrometric parameters with field samples:

$$\text{Bias} = \frac{\sum_{i=1}^{n} X_{\text{ALS},i} - X_{\text{field},i}}{n} \tag{8.4}$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (X_{\text{ALS},i} - X_{\text{field},i})^2}{n}} \tag{8.5}$$

$$\text{RMSE\%} = \frac{\text{RMSE}}{\bar{X}_{\text{ALS}}} \tag{8.6}$$

where n is the number of field samples, which equaled 40 trees in this study; X refers to the values of dendrometric parameters (height or crown diameter) measured either in the field or from the ALS data; and $\bar{X}_{\text{ALS}}$ is the arithmetic mean of the ALS-derived measurements. Moreover, a linear regression model was fit to the ALS-derived tree height and crown diameter to determine if there was a strong correlation between these two variables.

8.3.4 Carbon Storage Modeling

In order to predict the carbon storage in trees, a multiple linear regression model was developed empirically to fit the data, with the ALS-derived dendrometric parameters as the independent predictors and the field-measured DBH as the predicted variable.

The empirical equation derived from the ALS-DBH linear regression model has the form

$$\mathrm{DBH}_{\mathrm{field}} = a \cdot CD_{\mathrm{ALS}} + b \cdot H_{\mathrm{ALS}} + c \tag{8.7}$$

The 40 field-sampled trees were split into two data sets, with 20 trees used for model parameterization and the remaining 20 trees reserved for validation. To eliminate the influence of tree locations in model fitting, each 10 adjacent tree samples were grouped together under one sampling location, resulting in a total of four sampling groups. Six combinations of training and validation data sets were chosen by selecting two sampling locations out of four for model development and using the remaining two locations for validation. All six models were developed at a 0.05 significance level and were fitted through a cross-validation process. The parameters generated for each model were collected. The predictive power of the regression models and the performance were inspected by the coefficient of determination (R^2), and the accuracy of the prediction was examined by the RMSE of the predicted parameters. Comparing the R^2 and RMSE of the six ALS-DBH regression models, the one with a high R^2 in the model fitting and a low RMSE in the validation was selected to predict DBH.

After selecting the ALS-DBH regression model, the ALS-estimated DBH and height were plugged into the Canadian national aboveground biomass equations proposed by Lambert et al. (2005) to estimate the carbon storage in trees. From Lambert et al. (2005), the set of equations based on DBH and height for all species was selected to calculate the biomass since no genus or species information was available in this study. The aboveground biomass was estimated as the sum of biomass in tree compartments (foliage, branch, wood, and bark). The carbon stored in trees was estimated as half of the total biomass. The carbon storage predicted by the ALS-derived parameters was compared with that estimated by the field-measured DBH and height and evaluated by the RMSE and R^2.

To show whether carbon storage in trees varied with land cover types, the ALS-derived tree crown segments were first converted into vector data in a geographic information system (GIS), with the amount of carbon storage stored in the attributes. The carbon storage within each land cover type was calculated by adding up all the carbon storage in trees and dividing by the area of the land cover type. For the Town of Whitchurch-Stouffville, the carbon stocks were extrapolated by multiplying the specific carbon amount per unit area with the total area of each land cover type. For those land cover types that were excluded in the study area (namely, government and institutional areas and industrial sites), the carbon stored in government and institutional areas was given the same amount per unit area as the residential area, but the carbon stored in industrial sites was given zero. Then a citywide carbon storage map was created.

8.4 RESULTS AND DISCUSSION

8.4.1 Analysis of Multispectral ALS Data for Land Cover Classification

Because the classification at the tree genus or species level is important, especially for precise biomass estimation, the spectral patterns of the tree class were examined

further to find if current multispectral ALS data sets could distinguish the tree class into conifer and deciduous trees. After confirming the tree types in Google Earth, there were no obvious distinctions between the tree genus by visual observations. However, through close visual observation of the pNDVI data set, the trees with dark-color leaves, such as the Crimson King Maple tree, could stand out from the tree class (Figure 8.6). The multispectral ALS data sets generated in the chapter may not be sensitive enough to provide separate classification for conifers and deciduous trees. However, the analysis presented here shows that the multispectral ALS intensity may be influenced by factors such as the color or reflectivity of the objects, which in turn will be beneficial to studies of tree mortality, rooftop solar energy, and so on.

Accuracy assessment was conducted on the classification results in order to examine the classification performance of multispectral ALS data. Figure 8.7 shows the final classification map. The combination of all six input data achieved the highest overall accuracy among the three selections (Table 8.3). The 89% overall accuracy was achieved using the ALS-derived raster data, which indicated that the contributions of the two calculated indices to the overall accuracy were not significant. The comparison between typical ALS data and multispectral ALS data on land cover classification was also conducted; the same classification process was applied to the nDSM and NIR bands only. The overall accuracy of classification in use of typical ALS data was around 79%.

8.4.2 Results for the Local Maxima Selection

The second step of the methodology was to find the local maxima in the tree-isolated nDSM (the so-called CHM) as the treetop candidates. In this chapter, a new approach is proposed to reduce the number of pixels generated by the original 3 × 3 local maxima

FIGURE 8.6 Trees with dark-color leaves in Google Earth and the pNDVI. ((a), (b), (d), and (e) reprinted from Google Earth (2015).)

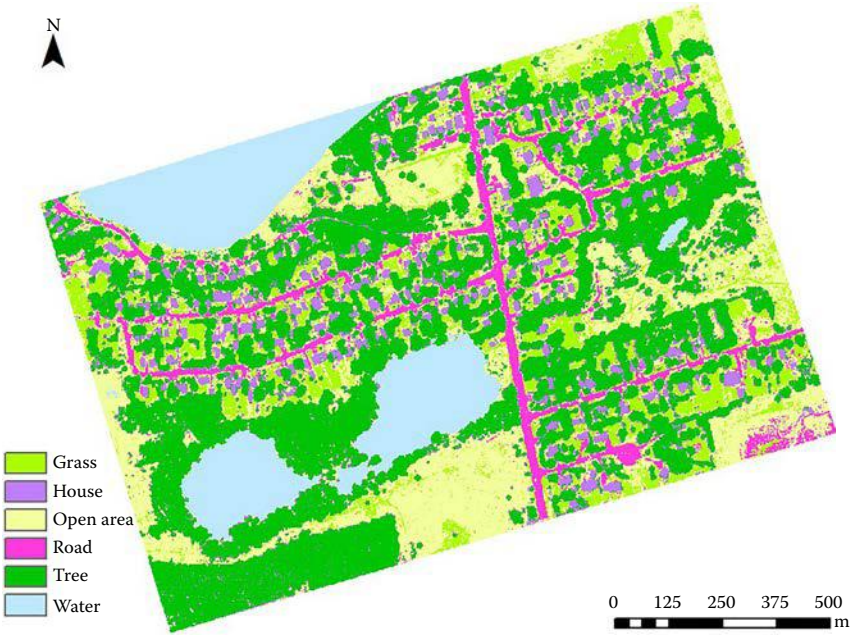


FIGURE 8.7 Classification map.

TABLE 8.3
Accuracy Report for the Classification Results

	Green	NIR	SWIR	nDSM	pNDWI	pNDVI	Overall Accuracy (%)
Selection 1	√	√	√	√	√	√	90.23
Selection 2	√	√	√	√			89.12
Selection 3		√		√			79.04

filter and to find the true treetop pixels. Figure 8.8 illustrates the comparison between the original local maxima result performed on the CHM and the refined result, which only retained the pixels that had a local maximum on both the maximum intensity data and the CHM. The clusters of local maxima in the original result were eliminated by the proposed method. Then the watershed segmentation algorithm was applied to generate crown segments, using the refined local maxima as the markers (Figure 8.8c).

8.4.3 Validations for the ALS-Derived Dendrometric Parameters

The accuracies of ALS-derived tree height and crown width were assessed by the samples measured in the field. For the ALS-derived tree height, an RMSE of 1.21 m (relative RMSE = 6.8%) and a negative bias of 0.2 m (relative bias = −0.1%) are given in Table 8.4. For the ALS-derived crown width, an RMSE of 1.47 m (relative

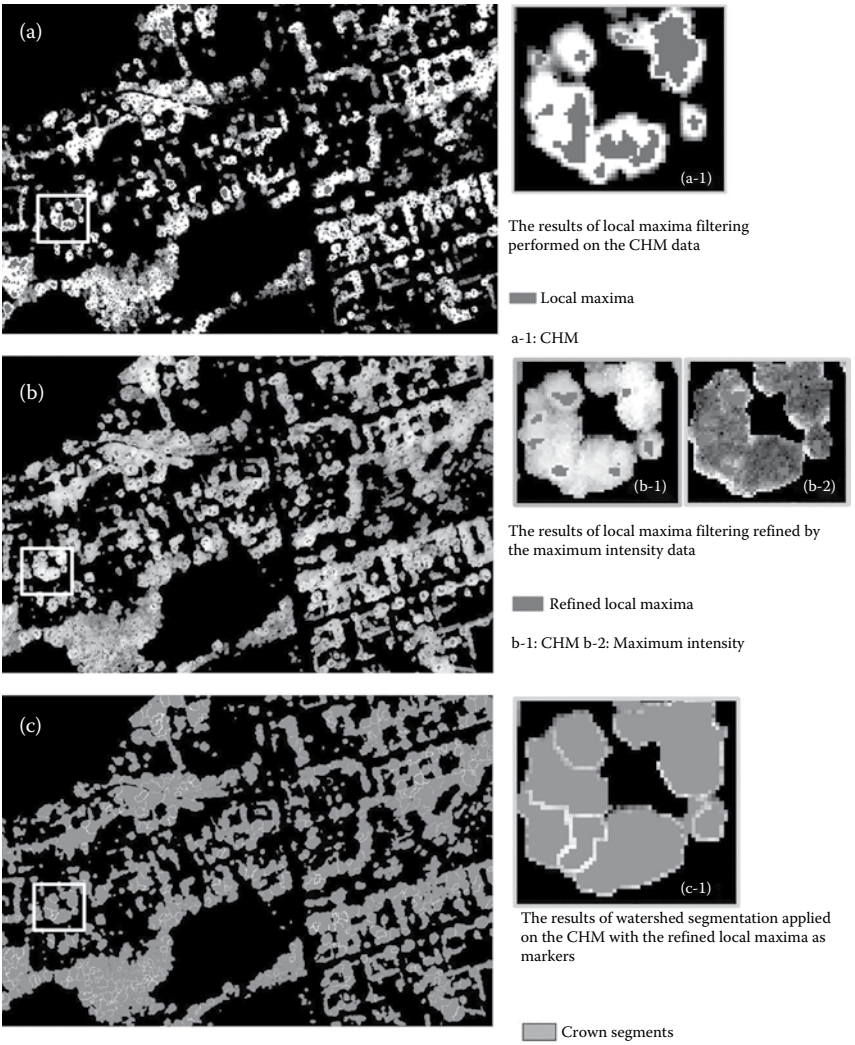


FIGURE 8.8 Results obtained by local maxima filtering and the marker-controlled watershed segmentation.

TABLE 8.4
Validation Statistics for the ALS-Derived Dendrometric Parameters

Parameter	RMSE	RMSE%	Bias	Bias%
Height (m)	1.21	6.8	−0.20	−0.1
Crown width (m)	1.47	16.4	−0.18	−2

TABLE 8.5
Results of Model Fitting and Model Validation

Model	Model Fit R^2	Model Fit RMSE (cm)	Validation R^2	Validation RMSE (cm)
1	0.83	5.35	0.80	6.82
2	0.86	6.60	0.76	5.60
3	0.86	3.86	0.71	8.25
4	0.75	7.89	0.77	4.82
5	0.78	5.59	0.85	6.55
6	0.81	7.00	0.83	5.20

RMSE = 16.4%) and a negative bias of 0.18 m (relative bias = −2%) were observed. The tree height was underestimated because of the undergrowth and factors related to the flight height (1000 m) and the point density (7 to 8 points/m^2). The relative RMSE of the crown width was mainly caused by the resolution of the CHM and the results of crown segmentation.

8.4.4 Validations for the ALS-Predicted DBH and Tree Carbon

The results of model fitting and validation of the six ALS-DBH regression models are listed in Table 8.5. The 40 field-measured samples were previously split into four groups, with 10 trees in each group. The models were iteratively fitted by 20 trees selected from two groups out of the four and were validated by the remaining 20 trees. Model 2 was selected as the overall best model to predict DBH in this study because it has a relatively high coefficient of determination ($R^2 = 86\%$) from model fitting and a relatively low RMSE (5.6 cm) from the validation. The regression equation is

$$\mathrm{DBH} = -11.2792 + (-0.2958)CD + 3.2637H \tag{8.8}$$

To determine the number of decimal places for the coefficients in the regression equation, the residuals between the DBH values predicted by the coefficients with eight decimal places and those predicted by coefficients rounded to two, three, and four decimal places, respectively, are compared in Figure 8.9. To keep high prediction accuracy, coefficients in the regression model are rounded to four decimal places. Moreover, the R^2 between field-measured tree height and crown width was calculated as 0.34, indicating insignificant correlation between these two variables. Though DBH cannot be directly measured on the CHM, all the generated ALS-DBH models showed that the DBH correlated well with ALS measurements. The accuracies of the ALS-modeled DBH and ALS-derived carbon storage are given in Table 8.6. The predicted DBH using ALS-derived parameters corresponded to an RMSE of 6.4 cm (relative RMSE = 13.1%) and a bias of 0.4 cm. The relationship between field-measured DBH and ALS-modeled DBH is plotted in Figure 8.10. The results are compared with the reference carbon storage estimated by field-measured DBH and plotted in Figure 8.11. The R^2 values of both DBH and carbon storage were both above

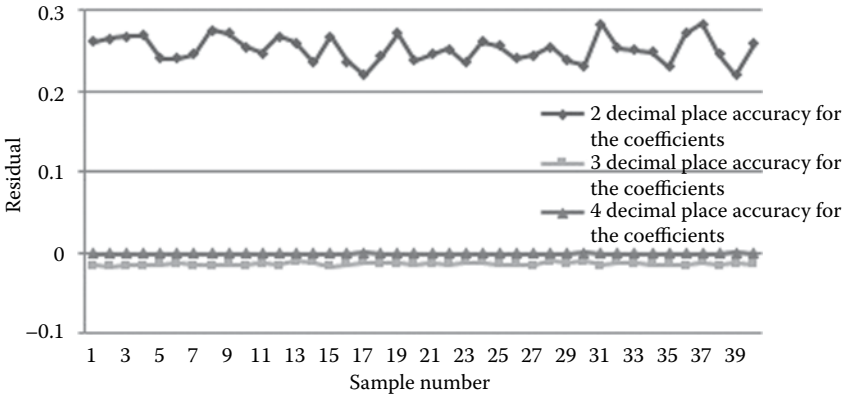


FIGURE 8.9 Residual plots for model-predicted DBH generated by coefficients with two, three, and four decimal places.

TABLE 8.6
Accuracy of ALS-Predicted versus Field-Measured Results

Parameter	RMSE	RMSE%	Bias	Bias%
DBH (cm)	6.39	13.1	0.44	0.1
Carbon (kg)	142.0	28.6	14.4	2.9

0.80. The predicted carbon storage using the ALS-modeled DBH corresponded to an RMSE of 142 kg (28.6%) and a bias of 14.4 kg.

8.4.5 Analysis of the ALS-Estimated Results

The accuracy of tree height measurements using ALS data has previously been studied by Yu et al. (2004), Kaartinen et al. (2012), and Hadaś & Estornell (2016). Yu et al. (2004) found that as flight altitude increased from 400 to 1500 m, the accuracy of tree heights lowered from 0.76 to 1.16 m for single tree species. Kaartinen et al. (2012) reported that the best methods that utilized the local maxima finding with a point density of 8 points/m^2 could obtain an RMSE of 60–80 cm for tree heights. Hadaś & Estornell (2016) showed that the bias of tree height measurements could decrease from −1.48 to −0.72 m if the point density increased from 3.5 to 9 points/m^2. The RMSE achieved in the present study is in line with these studies and is potentially affected by the errors generated during the field measurements. Both overestimations of the crown size and underestimations of the tree heights are likely a result of the overlaying crown covers of the dominant tree and the suppressed trees, but could be mitigated if the resolution of CHM is at the submeter level. The results of ALS-modeled DBH are in line with the findings in Hauglin et al. (2014) and Popescu (2007) regarding the tree height and crown diameter as good predictors to predict DBH using

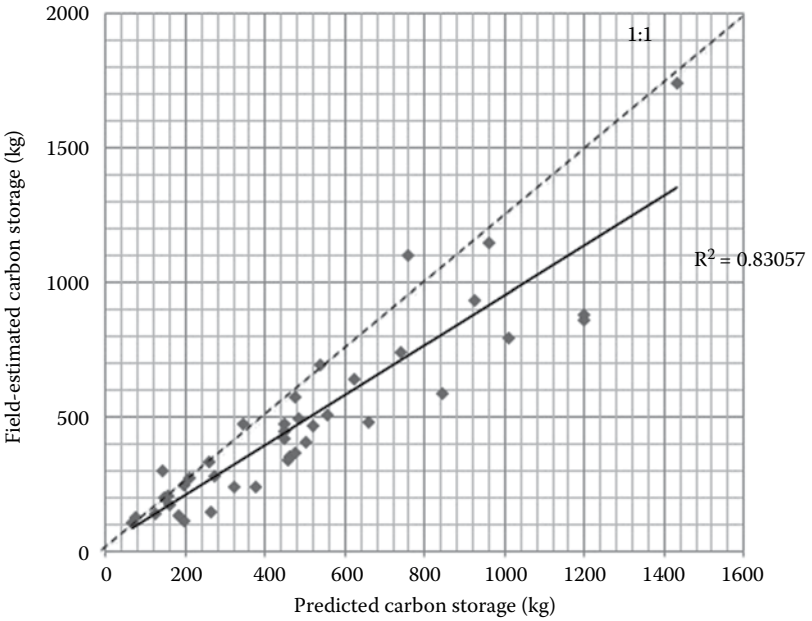


FIGURE 8.10 Scatterplot of the ALS-modeled DBH versus field-measured DBH.

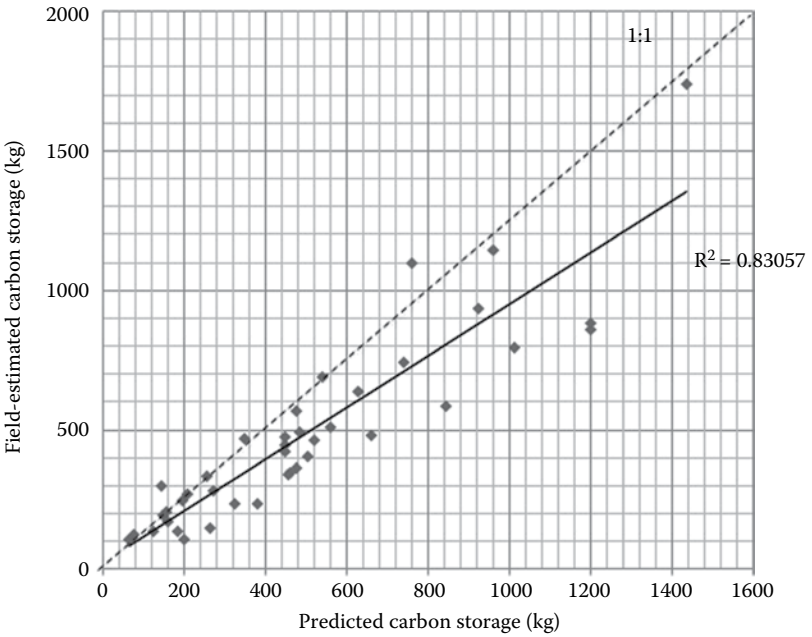


FIGURE 8.11 Scatterplot of the ALS-predicted carbon storage versus field-estimated carbon storage.

linear regression. The accuracy found in the present study is higher than in these two studies. Hauglin et al. (2014) reported an RMSE of 35% for ALS-estimated DBH of Norway spruce. Popescu (2007) reported a lower RMSE (4.9 cm) and higher R^2 compared to the present study. However, considering that the accuracies in Popescu (2007) were for a single tree species, which had an average DBH of 29.55 cm, and that they used all 43 sampled trees to construct the model and validate the model using the same data set, the RMSE% and R^2 in their study would understandably be higher than that in this study. For aboveground biomass, Popescu (2007) reported an RMSE of 47% for a single tree species. Kankare et al. (2013) reported RMSEs of 26.3% and 36.8% for pines and spruce, respectively. Hauglin et al. (2014) achieved an RMSE of 35.1% for biomass estimation. Though the achieved accuracy of carbon estimation is higher in the present study, these studies are not entirely comparable because some studies used field-destructive measurements as reference data, which were not available in this study. Also, because genus information is not available, the estimation of carbon stocks was done by the allometry equations for all species in Lambert et al. (2005), so that the derived carbon mainly depended on the ALS-derived DBH and height, with little consideration given to the differences in species.

8.4.6 Analysis of the Carbon Storage

There were a total of 2555 dominant trees in the study area. The trees were located along the roadsides, in backyards, and around the lakes. The average tree carbon was 484.3 kg and resulted in a total of 1.24 kt C. The study area could be divided into four land use types: residential, park and recreational, open area, and water. The open area occupied the largest portion of the study area, 47.0 ha. The residential area occupied 25.8 ha. The park had a small area of 2.5 ha. The extracted carbons in trees were therefore grouped based on the land use type, and the amount of carbon stored in each land cover type was calculated. Within the study area, the open area contained the largest tree carbon stocks (682.7 t), followed by the residential area (362.6 t), and finally the parks and recreational area (29.2 t). The tree carbon storage for the open area, residential area, and parks on a per unit area basis were 14.54, 14.08, and 11.57 t C/ha, respectively. City wide, open area occupied 83.7% (191.5 km^2) of the total city area and contained the largest carbon storage, 278.4 kt C. Residential area covered 8.3% (19.4 km^2) of the total city area and contained 27.3 kt C tree carbons. Parks covered 12.8 km^2 with a total carbon storage estimated at 14.8 kt C. The carbon storage map shows that large tree carbon stocks are accumulated in urban environments and are distributed heterogeneously among land use types (Figure 8.12).

The estimated tree carbons in the study area were in line with the estimation of carbon storage in Canadian urban trees conducted by researchers at Environment Canada (Pasher et al., 2014). Pasher et al. (2014) estimated the carbon stocks in trees by applying the crown cover area of urban trees and a Canadian-specific area-based growth rate for urban trees. Pasher et al. (2014) reported a total urban area of 5317 km^2 in the Ontario Mixedwood Plains, with an estimation of carbon storage at 9177.6 kt C, resulting in a carbon storage per unit urban area of 17 kt C/ha. This is slightly more than what has been predicted in this study (around 14 kt C/ha). The

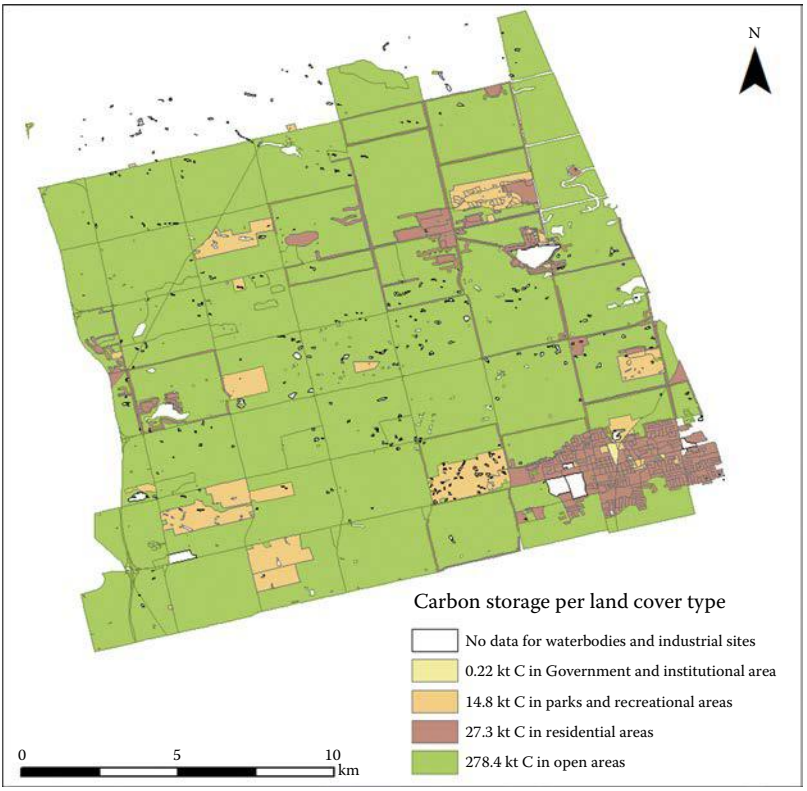


FIGURE 8.12 Carbon storage map for the Town of Whitchurch-Stouffville, Ontario, Canada.

main source of the difference comes from the approaches to estimating the biomass amounts between the present study and that used in Pasher et al. (2014).

8.5 CONCLUSIONS

This chapter proposed a workflow to map land covers and estimate aboveground carbon storage in trees at a spatial resolution of 1 m using multispectral ALS data. This chapter shows that good classification results can be obtained solely from multispectral ALS data sets. This study achieved an overall accuracy of 90%, which was 11% higher than that obtained using single-wavelength ALS data. Spectral patterns for impervious surfaces (road, rooftops) and single-return vegetation (grass) are observed to have similar patterns in the optical imagery. The dendrometric parameters at the single-tree level can be derived directly from the multispectral ALS data. It also shows how the use of both spectral and geometric properties of multispectral ALS data can improve the detection of treetops. The improvement could be more significant if the resolution of the CHM were finer.

This chapter presented the feasibility of applying forest-based allometric methods to assess carbon stocks in urban environments. Dominant trees with fewer underneath

or nearby trees were better detected and analyzed in the study. Though DBH cannot be directly measured from ALS data, the ALS-predicted DBH remains a powerful predictor for estimating tree carbon at the individual tree level. More accurate tree carbon measurements could be obtained if genus information and crown base heights were further investigated. An improvement of derivation of the crown width would also help in better prediction of tree carbon stocks. This chapter derived similar carbon amounts per unit area in both residential areas and open areas within the study area, because the open area had twice the size of the residential area but the density of the canopy cover was less than it was in the residential area. Citywide carbon storage estimation was derived in this chapter by extrapolating the values within the study area to the entire city based on the specific proportion of each land cover type in the city. This approach is applicable here because the Town of Whitchurch-Stouffville has a relatively simple city structure, the study area has included the major components of the city, and those land covers excluded in the study area only occupied a small proportion of the entire city.

Urban ecosystems are an important component in the global carbon cycle. In the context of urban sprawl, quantifying the carbon storage for urban areas is very important in terms of getting reliable estimations of carbon sequestration rate and magnitude, but it is a difficult and complex task that requires advanced analysis techniques and data sources to achieve fine-scale estimation. The methods developed here provide an accurate and detailed estimate of how urban trees in a Canadian city play the role of a carbon sink. The presented approach of estimating carbon stocks in urban trees takes advantage of the available Canada-wide allometry relationship between biomass and tree DBH and height, as well as the power of the ALS system in providing the estimation of dendrometric parameters. The methodology proposed does not require destructive sampling or large-scale field works. It is applicable to other urban areas and is beneficial to better understanding of urban carbon budgets and urban heat island effects. It also provides valuable information on the impact of climate change to city planners.

In conclusion, this chapter has developed a detailed workflow to estimate tree carbon stocks from multispectral ALS data by using a series of techniques including SVM classification, watershed segmentation, and allometry-based linear regression modeling. This chapter also demonstrated the strong capability of multispectral ALS data in land cover mapping and tree-level inventory in urban environments. The Titan scanner showed a trend for the development of airborne LiDAR technology. This scanner is able to realize a series of environmental and topographic applications that have been performed previously by photogrammetric data.

REFERENCES

Bakuła, K. 2015. Multispectral airborne laser scanning—A new trend in the development of LiDAR technology. *Archiwum Fotogrametrii, Kartografii i Teledetekcji* 27: 25–44.

Brook McIlroy Inc. 2002. Community of Stouffville urban design guidelines. http://www.townofws.ca/en/residents/resources/Documents/Planning/Stouffville_UrbanDesignGuidelines.pdf (Accessed February 20, 2016).

Chen, Q., D. Baldocchi, P. Gong, and M. Kelly. 2006. Isolating individual trees in a savanna woodland using small footprint LiDAR data. *Photogrammetric Engineering & Remote Sensing* 72(8): 923–932.

Corfee-Morlot, J., L. Kamal-Chaoui, M.G. Donovan, I. Cochran, A. Robert, and P.J. Teasdale. 2009. Cities, climate change and multilevel governance. *OECD Environment Working Papers, No. 14*, doi:10.1787/220062444715.

Davies, Z.G., J.L. Edmondson, A. Heinemeyer, J.R. Leake, and K.J. Gaston. 2011. Mapping an urban ecosystem service: Quantifying above-ground carbon storage at a city-wide scale. *Journal of Applied Ecology* 48(5): 1125–1134.

Dodman, D., J. Bicknell, and D. Satterthwaite, Eds., 2012. *Adapting Cities to Climate Change: Understanding and addressing the development challenges*. New York: Routledge.

Donovan, G.H., and D.T. Butry. 2010. Trees in the city: Valuing street trees in Portland, Oregon. *Landscape and Urban Planning* 94(2): 77–83.

DMTI Spatial Inc. 2015. *Land cover region. Geospatial Center at University of Waterloo.* http://geo1.scholarsportal.info.proxy.lib.uwaterloo.ca/proxy.html?URL=http:__maps.scholarsportal.info/files/zips/DMTI/DMTI_2015_CMCS_LandCoverRegion.zip (Accessed January 10, 2016).

Google Earth 7.1.2.2041. 2004. *Town of Whitchurch-Stouffville. 44°01'29.7"N 79°16'03.6"W, Eye alt 582 m. First Base Solutions 2016.* http://www.earth.google.com (Accessed February 26, 2016).

Google Earth 7.1.2.2041. 2015. *Town of Whitchurch-Stouffville.44°01'29.7"N 79°16'03.6"W, Eye alt 1.41 km. First Base Solutions 2016.* http://www.earth.google.com (Accessed February 26, 2016).

Hadaś, E., and J. Estornell. 2016. Accuracy of tree geometric parameters depending on the LiDAR data density. *European Journal of Remote Sensing* 49: 73–92.

Hauglin, M., T. Gobakken, R. Astrup, L. Ene, and E. Næsset. 2014. Estimating single-tree crown biomass of Norway spruce by airborne laser scanning: A comparison of methods with and without the use of terrestrial laser scanning to obtain the ground reference data. *Forests* 5(3): 384–403.

Huang, Y., B. Yu, J. Zhou, C. Hu, W. Tan, Z. Hu, and J. Wu. 2013. Toward automatic estimation of urban green volume using airborne LiDAR data and high resolution remote sensing images. *Frontiers of Earth Science* 7(1): 43–54.

Johnson, T. 2009. Deforestation and greenhouse-gas emissions. *Council on Foreign Relations.* http://www.cfr.org/forests-and-land-management/deforestation-greenhouse-gas-emissions/p-14919 (Accessed February 1, 2016).

Kaartinen, H., J. Hyyppä, X. Yu, M. Vastaranta, H. Hyyppä, A. Kukko, M. Holopainen, C. Heipke, M. Hirschmugl, F. Morsdorf, and E. Næsset. 2012. An international comparison of individual tree detection and extraction using airborne laser scanning. *Remote Sensing* 4(4): 950–974.

Kankare, V., M. Räty, X. Yu, M. Holopainen, M. Vastaranta, T. Kantola, J. Hyyppä, H. Hyyppä, P. Alho, and R. Viitala. 2013. Single tree biomass modelling using airborne laser scanning. *ISPRS Journal of Photogrammetry and Remote Sensing* 85: 66–73.

Lambert, M.C., C.H. Ung, and F. Raulier. 2005. Canadian national tree aboveground biomass equations. *Canadian Journal of Forest Research* 35(8): 1996–2018.

Lieth, H. 1963. The role of vegetation in the carbon dioxide content of the atmosphere. *Journal of Geophysical Research* 68(13): 3887–3898.

Pasher, J., M. McGovern, M. Khoury, and J. Duffe. 2014. Assessing carbon storage and sequestration by Canada's urban forests using high resolution Earth observation data. *Urban Forestry & Urban Greening* 13(3): 484–494.

Popescu, S.C. 2007. Estimating biomass of individual pine trees using airborne LiDAR. *Biomass and Bioenergy* 31(9): 646–655.

Popescu, S.C., R.H. Wynne, and R.F. Nelson. 2003. Measuring individual tree crown diameter with LiDAR and assessing its influence on estimating forest volume and biomass. *Canadian Journal of Remote Sensing* 29(5): 564–577.

Raciti, S.M., L.R. Hutyra, and J.D. Newell. 2014. Mapping carbon storage in urban trees with multi-source remote sensing data: Relationships between biomass, land use, and demographics in Boston neighborhoods. *Science of the Total Environment* 500: 72–83.

Roy, S., J. Byrne, and C. Pickering. 2012. A systematic quantitative review of urban tree benefits, costs, and assessment methods across cities in different climatic zones. *Urban Forestry & Urban Greening* 11(4): 351–363.

Satterthwaite, D. 2009. The implications of population growth and urbanization for climate change. *Environment and Urbanization* 21(2): 545–567.

Sawka, M., A.A. Millward, J. Mckay, and M. Sarkovich. 2013. Growing summer energy conservation through residential tree planting. *Landscape and Urban Planning* 113: 1–9.

Schreyer, J., J. Tigges, T. Lakes, and G. Churkina. 2014. Using airborne LiDAR and QuickBird data for modelling urban tree carbon storage and its distribution—A case study of Berlin. *Remote Sensing* 6(11): 10636–10655.

Sousa, A.M., A.C. Gonçalves, P. Mesquita, and J.R.M. da Silva. 2015. Biomass estimation with high resolution satellite images: A case study of Quercus rotundifolia. *ISPRS Journal of Photogrammetry and Remote Sensing* 101: 69–79.

United Nations, 2014. *World urbanization prospects 2014: highlights*. http://esa.un.org/unpd/wup/highlights/wup2014-highlights.pdf (Accessed February 01, 2016).

van den Berg, M., W. Wendel-Vos, M. van Poppel, H. Kemper, W. van Mechelen, and J. Maas. 2015. Health benefits of green spaces in the living environment: A systematic review of epidemiological studies. *Urban Forestry & Urban Greening* 14(4): 806–816.

Wichmann, V., M. Bremer, J. Lindenberger, M. Rutzinger, C. Georges, and F. Petrini-Monteferri. 2015. Evaluating the potential of multispectral airborne LIDAR for topographic mapping and land cover classification. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences* 1: 113–119.

Whittaker, R.H., and G.E. Likens. 1973. Carbon in the biota. In *Carbon and the Biosphere*, G. M. Woodell and E. V. Pecans, Eds., 281–302. Springfield, VA: U.S. National Technical Information Service.

Xu, Q., Z. Hou, M. Maltamo, and T. Tokola. 2014. Calibration of area based diameter distribution with individual tree based diameter estimates using airborne laser scanning. *ISPRS Journal of Photogrammetry and Remote Sensing* 93: 65–75.

Xu, T., J. Sathaye, H. Akbari, V. Garg, and S. Tetali. 2012. Quantifying the direct benefits of cool roofs in an urban setting: Reduced cooling energy use and lowered greenhouse gas emissions. *Building and Environment* 48: 1–6.

Yu, X., J. Hyyppä, H. Kaartinen, and M. Maltamo. 2004. Automatic detection of harvested trees and determination of forest growth using airborne laser scanning. *Remote Sensing of Environment* 90(4): 451–462.

Zhao, K., S. Popescu, and R. Nelson. 2009. Lidar remote sensing of forest biomass: A scale-invariant estimation approach using airborne lasers. *Remote Sensing of Environment* 113(1): 182–196.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

9 Suitable Spectral Mixing Space Selection for Linear Spectral Unmixing of Fine-Scale Urban Imagery

Jian Yang

CONTENTS

9.1 INTRODUCTION

To date, subpixel methods have most frequently been used to extract urban attributes (e.g., sealed surface, green space). As a prevalent component of subpixel methods, linear spectral unmixing has captured increasing attention for its use in sealed surface or green space mapping in urban areas. Since a spectrum within a pixel collected by a remote sensor is always mixed by different surface materials (Shimabukuro and Smith 1991), linear spectral unmixing is well known as a procedure to decompose a mixed spectrum into a group of fractions on the basis of pure surface materials (Keshava 2003). In the conventional sense, coarse-medium spatial resolution imagery is the most appropriate for implementing linear spectral unmixing, and thus is the most widely applied for mapping urban sealed surfaces (Phinn et al. 2002; Wu and Murray 2003; Lu and Weng 2006; Weng et al. 2008; Van de Voorde et al. 2009; Weng et al. 2009) or green space (Small 2001; Song 2005; Small and Lu 2006; Liu and Yang 2013). With the advances in remote sensing techniques, more endeavors have been

made to spectrally unmix high spatial resolution imagery to extract sealed surfaces (Lu and Weng 2009; Wu 2009; Yang and Li 2015; Yang and He 2017) and green space (Nichol and Wong 2007; Nichol et al. 2010) in urban areas.

In high spatial resolution imagery, an urban sealed surface is spectrally differentiated into high- and low-albedo materials, while bare soil becomes more spectrally similar to high-albedo materials (Yang et al. 2014; Yang and He 2017). From this point of view, the vegetation-high albedo-low albedo model (Small 2003) is more appropriate than the classic vegetation-impervious surface-soil model (Ridd 1995) for spectrally unmixing high spatial resolution urban imagery. As demonstrated by Yang et al. (2014), multispectral bands of high spatial resolution imagery can construct various spectral mixing spaces for the description of urban surface materials at fine scales. Specifically, a two-dimensional spectral mixing space constructed by a non-near-infrared (IR) band and a near-IR band is able, and is the most widely used, to implement the vegetation-high albedo-low albedo model for linear spectral unmixing (Yang et al. 2015; Yang and He 2017). Within this type of two-dimensional spectral mixing space, all of the unmixed pixels can form a scatterplot of a triangle, three vertices of which represent the pure pixels (hereafter referred to as end-members) of vegetation, high albedo, and low albedo, respectively. To identify these three end-members, Yang et al. (2014) proposed an automated method based on end-member spatial distribution in the two-dimensional spectral mixing space, which was further developed by Yang et al. (2015).

In addition to end-member identification, the accuracy of linear spectral unmixing (ALSU) is also strongly dependent upon the selection of spectral mixing space. As suggested by Small (2003), a good unmixing result is largely attributed to a good scatterplot triangle that is well spanned, with straight or convex edges within the given two-dimensional spectral mixing space. Based on this assumption, the space constructed by the red and near-IR1 bands has proven to be the best for spectrally unmixing high spatial resolution urban imagery (Yang et al. 2015; Yang and He 2017). In these studies, however, the selection of spectral mixing space was determined by visual interpretation rather than quantitative assessment. In the current study, we therefore propose a new indicator to quantify the structure of scatterplot triangles in different two-dimensional spectral mixing spaces, and then to select the best one for spectrally unmixing high spatial resolution urban imagery. To ensure a good scatterplot triangle with straight or convex edges, we hypothesized that a few pixels should be located outside the triangle, though not overly far away from the triangle, in case there were some outliers. The rest of this chapter is organized as follows. Section 9.2 will describe the workflow of linear spectral unmixing, as well as our proposed indicator for quantifying the structure of the scatterplot triangle. Section 9.3 will introduce the selected study site, experimental data set, and results. Section 9.4 will discuss our experimental results. Finally, the main conclusions are reiterated in Section 9.5.

9.2 METHODS

9.2.1 Linear Spectral Unmixing

Due to the impacts of low sun elevation, off-nadir viewing angle, and high-rise buildings, shadows are widely distributed across fine-scale urban imagery (Yang

et al. 2015). Because the spectral features of surface materials in unshadowed areas are significantly different from those in shadowed areas, it is necessary to implement linear spectral unmixing in unshadowed and shadowed areas separately (Yang and He 2017). In the current study, the object-based method was utilized to separate the unshadowed and shadowed areas. The multiresolution segmentation algorithm (Benz et al. 2004), which is implemented in eCognition Developer (Trimble Munich, Munich, Germany), was first used to produce meaningful segments, with shape, compactness, and scale parameter assigned as 0.1, 0.5, and 50, respectively. For each spectral band, the Digital Number (DN) values of all pixels within a segment were thereafter averaged to represent that segment. Furthermore, we averaged eight spectral bands to create brightness imagery, and implemented the conventional histogram threshold method to differentiate the unshadowed and shadowed segments (Dare 2005; Chen et al. 2007; Zhou et al. 2009).

In our study, the vegetation-high albedo-low albedo model was implemented for spectrally unmixing fine-scale urban imagery within different two-dimensional spectral mixing spaces constructed by a non-near-IR band and a near-IR band (Figure 9.1). Three vertices of this scatterplot triangle represent the end-members of vegetation, high albedo, and low albedo, respectively. To identify these three end-members, we measured the area of the triangle formed by any three pixels from the scatterplot, then selected the three pixels that resulted in the largest area (Yang et al. 2014). Yang et al. (2015) made use of the convex hull Graham's scan algorithm (Graham 1972) to expedite this procedure.

Mathematically, linear spectral unmixing is a process to solve the following equations:

$$DN_{\text{Non-Near-IR}} = f_V \times DN^V_{\text{Non-Near-IR}} + f_H \times DN^H_{\text{Non-Near-IR}} + f_L \times DN^L_{\text{Non-Near-IR}} \quad (9.1)$$

$$DN_{\text{Near-IR}} = f_V \times DN^V_{\text{Near-IR}} + f_H \times DN^H_{\text{Near-IR}} + f_L \times DN^L_{\text{Near-IR}} \quad (9.2)$$

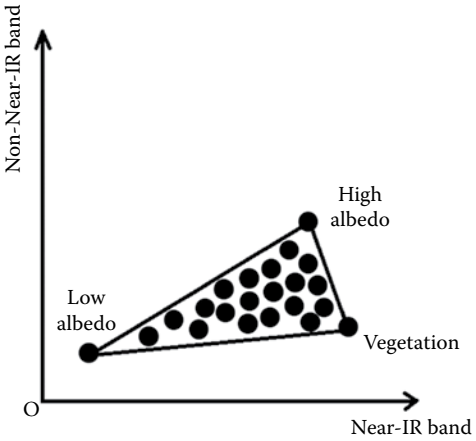


FIGURE 9.1 Scatterplot triangle in the two-dimensional spectral mixing space constructed by a non-near-IR band and a near-IR band. Black dots represent the unmixed pixels.

$$f_V + f_H + f_L = 1 \tag{9.3}$$

where $DN_{\text{Non-Near-IR}}$ and $DN_{\text{Near-IR}}$ are the DN values of any unmixed pixel in a non-near-IR band and a near-IR band, respectively. Additionally, $DN^X_{\text{Non-Near-IR}}$ and $DN^X_{\text{Near-IR}}$ are the DN values of each end-member ($X = V, H, L$, i.e., vegetation, high albedo, and low albedo) in a non-near-IR band and a near-IR band, while f_V, f_H, and f_L are the fractions of vegetation, high albedo, and low albedo, respectively. For any unmixed pixel located outside the scatterplot triangle, f_V, f_H, or f_L must be beyond the fraction range between 0 and 1. To solve this problem, we moved these outliers to their nearest vertices or edges of the scatterplot triangle prior to the solution of the preceding three equations (Yang et al. 2015). Two sets of end-members were identified from the unshadowed and shadowed areas, then used for spectrally unmixing the corresponding areas. Finally, both fraction maps were merged together to produce the final unmixing result.

To quantify the accuracy of linear spectral unmixing, we classified the pan-sharpened multispectral imagery by random forest (Breiman 2001) and calculated the reference proportion or amount for each end-member (Yang et al. 2014; Yang and Li 2015; Yang and He 2017). With respect to each end-member, the correct unmixed proportion (CUP) was obtained by dividing the number of correct unmixed pixels by the total number of pixels over the entire imagery (Yang et al. 2014; Yang and Li 2015; Yang and He 2017). The ALSU was defined as the average value of CUP by all three end-members.

9.2.2 Quantifying the Structure of the Scatterplot Triangle

Upon the identification of three end-members, the scatterplot triangle can be constructed in the two-dimensional spectral mixing space (Figure 9.1). In linear spectral unmixing, a good scatterplot triangle is always characterized by straight or convex edges (Small 2003). From this perspective, a good triangle should ensure that few unmixed pixels are located outside. Although our unmixing method was able to move the outliers to their nearest vertices or edges, the abundance of unmixed pixels outside the scatterplot triangle could substantially lower the unmixing accuracy. Thus, we should measure the outlier rate (OR) for the scatterplot triangle, as defined by

$$\text{OR} = \frac{\text{Number of Outliers}}{\text{Number of Total Unmixed Pixels}} \times 100\% \tag{9.4}$$

Meanwhile, we should also calculate the average outlier distance (OD) in different spectral mixing spaces to quantify how far the unmixed pixels are outside the triangle. Note that OD is measured between each outlier and its nearest vertex or edge of the scatterplot triangle. To integrate the merits of both OR and OD, we further proposed a new indicator (ORD [Integration of Outlier Rate and Distance]) to quantify the structure of the scatterplot triangle, as denoted by

$$\text{ORD} = \text{OR} \times \text{OD} \tag{9.5}$$

Due to the different scatterplot triangles for linear spectral unmixing, the values of OR, OD, and ORD were calculated for the unshadowed and shadowed areas, respectively. The values of OR, OD, and ORD, weighted by their respective areas, were summed up for the entire imagery. Hereafter, the OR, OD, and ORD values refer to the weighted indicators for both the unshadowed and shadowed areas in different spectral mixing spaces.

Furthermore, we proposed another indicator to quantify the similarity of the scatterplot triangles of unshadowed and shadowed areas in different spectral mixing spaces. Regardless of rotation, the similarity of scatterplot triangles (TS) can be expressed as

$$\mathrm{TS} = \left|\mathrm{MinIA}_{\mathrm{US}} - \mathrm{MinIA}_{\mathrm{S}}\right| + \left|\mathrm{SMinIA}_{\mathrm{US}} - \mathrm{SMinIA}_{\mathrm{S}}\right| \tag{9.6}$$

where $\mathrm{MinIA}_{\mathrm{US}}$ and $\mathrm{SMinIA}_{\mathrm{US}}$ are the smallest and second-smallest interior angles of the scatterplot triangle in the unshadowed areas, and $\mathrm{MinIA}_{\mathrm{S}}$ and $\mathrm{SMinIA}_{\mathrm{S}}$ are the smallest and second-smallest interior angles of the scatterplot triangle in the shadowed areas.

9.3 EXPERIMENTS

9.3.1 Study Site and Experimental Data Set

Our study site was located in downtown Toronto, Ontario, Canada, since Toronto is typically an urban mosaic of green space, sealed surface, and shadow. The WorldView-2 data set for the Greater Toronto Area (GTA), acquired on June 2, 2011, was used. It contains eight multispectral bands (i.e., coastal, blue, green, yellow, red, red edge, near-IR1, and near-IR2) with a spatial resolution of 2 meters, and a panchromatic band with a spatial resolution of 0.5 meters. To obtain the reference imagery, as mentioned previously, the multispectral and panchromatic bands were fused to produce pan-sharpened multispectral imagery, with a spatial resolution of 0.5 meters. The fusion process was implemented by NNDiffuse Pan Sharpening (Sun et al. 2014), which is a new algorithm in ENVI 5.3. Finally, a 1×1 kilometer subset of original multispectral imagery was clipped for our experimental data (Figure 9.2), as well as the corresponding subset of pan-sharpened multispectral imagery.

9.3.2 Experimental Results

We examined the histogram of brightness imagery (i.e., the mean of the eight spectral bands) based on objects generated from multiresolution segmentation (Figure 9.3). The brightness histogram is multimodal, with the shadowed segments occupying the lower end of the histogram. Therefore, the optimal threshold value of 190 (the circle in Figure 9.3) was determined by this histogram thresholding method. Consequently, the unshadowed and shadowed areas were separated (Figure 9.4), occupying 94.84% and 5.16% of the entire study site, respectively.

In this study, six non-near-IR bands (i.e., coastal, blue, green, yellow, red, and red edge) and two near-IR bands (i.e., near-IR1 and near-IR2) of WorldView-2

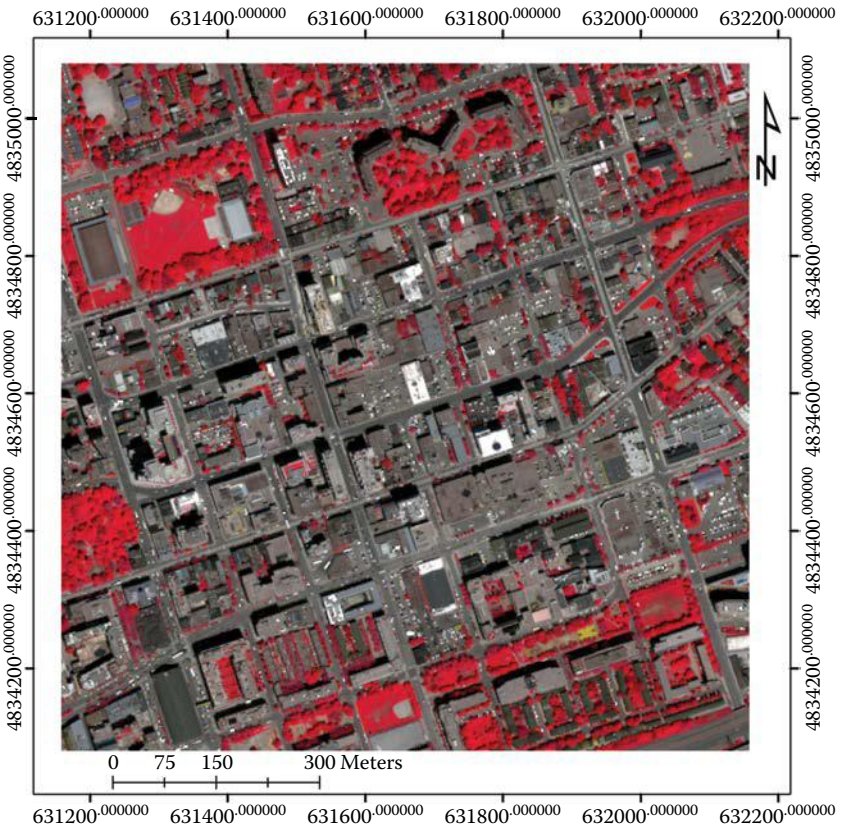


FIGURE 9.2 Clipped subset of original multispectral imagery used for experiment and analysis (near-IR1, red, and green bands as R, G, and B). It is under the WGS84 UTM coordinate system.

multispectral imagery constructed 12 spectral mixing spaces for the unshadowed (Figure 9.5) and shadowed (Figure 9.6) areas, respectively. Within each spectral mixing space, the scatterplot triangle was formed by three automatically identified end-members of vegetation, high albedo, and low albedo (the triangles in Figures 9.5 and 9.6). Compared to the others, the spectral mixing space constructed by the red and near-IR1 bands demonstrated the best triangle, matching the pixel cloud in this two-dimensional space well, especially for the unshadowed areas. To further demonstrate the triangle structures for all 12 spectral mixing spaces, three quantitative indicators of the scatterplot triangles (i.e., OR, OD, and ORD) are summarized in Table 9.1, as well as their corresponding accuracy of linear spectral unmixing (i.e., ALSU).

For the preceding 12 spectral mixing spaces, we plotted the accuracy of linear spectral unmixing (i.e., ALSU) versus our three proposed indicators of scatterplot triangles (i.e., OR, OD, and ORD) in Figure 9.7. All three indicators were negatively related to the unmixing accuracy, among which the ORD value performed the highest, with an R^2 of 0.65 (OR = 0.31; OD = 0.083). We therefore concluded that

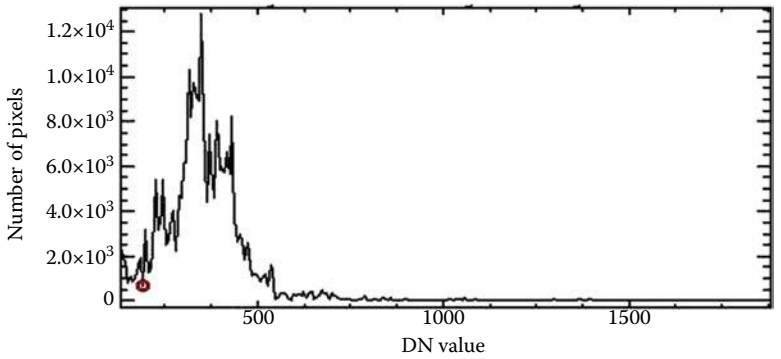


FIGURE 9.3 Histogram of brightness (i.e., the mean of the eight spectral bands), calculated using the number of pixels as statistical units.

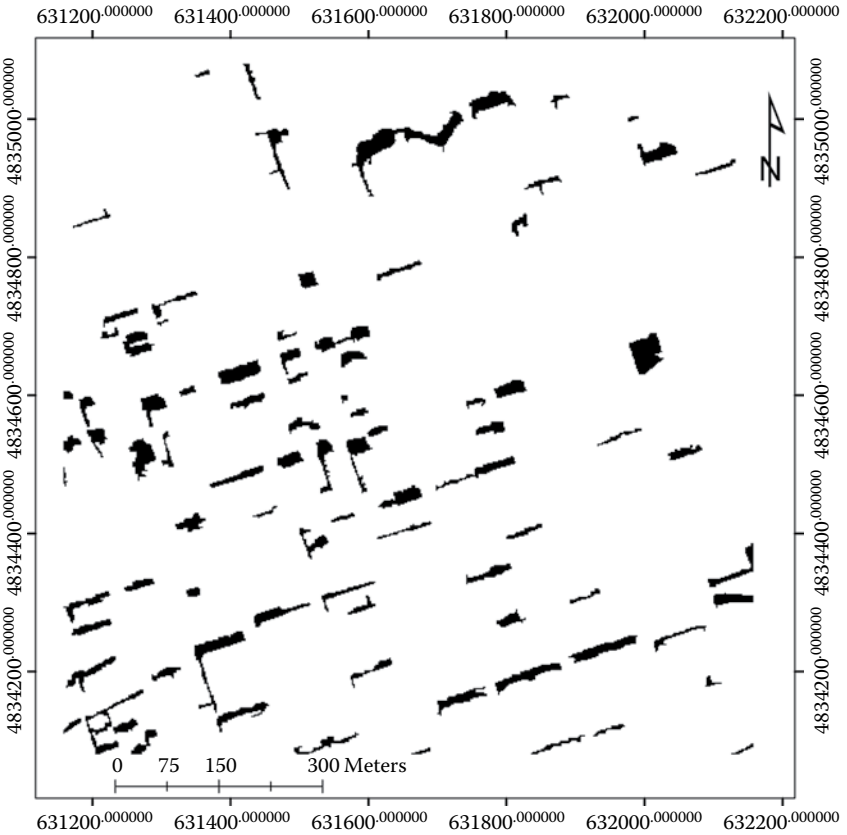


FIGURE 9.4 Unshadowed areas (white) versus shadowed areas (black) in our study site.

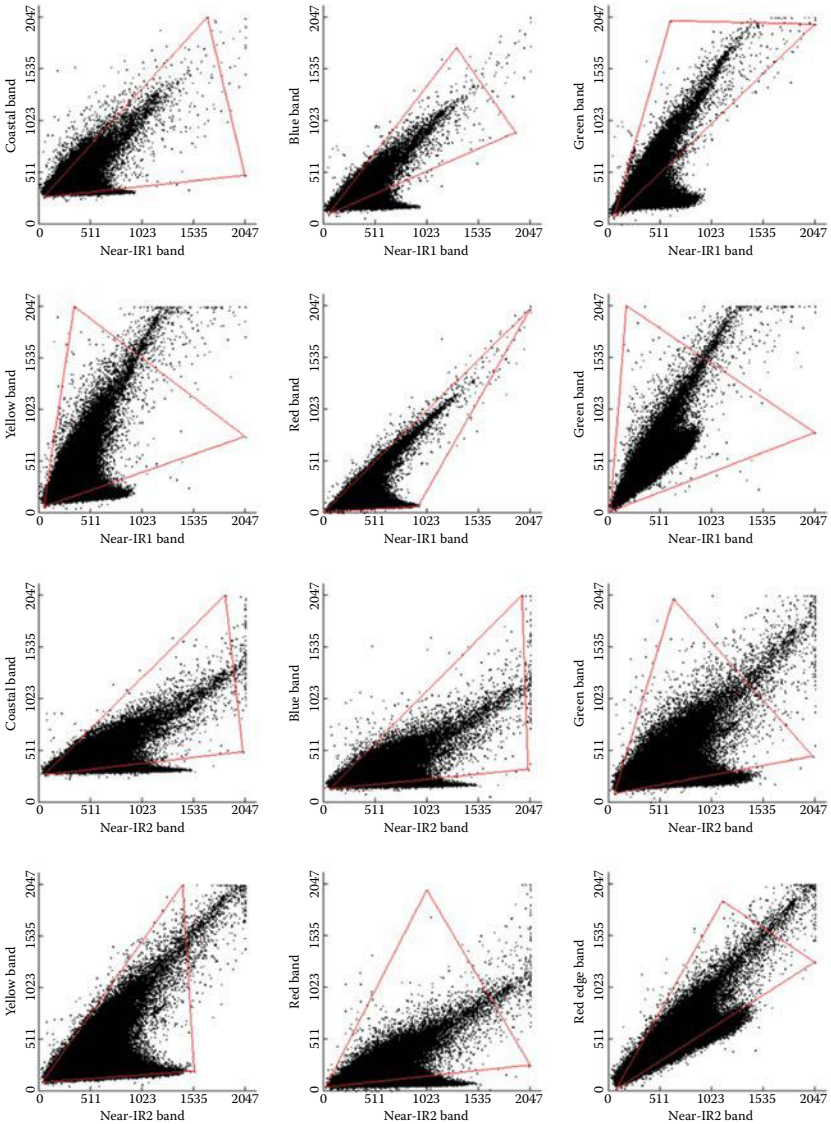


FIGURE 9.5 Scatterplot triangles (triangles) in different spectral mixing spaces constructed by a non-near-IR band and a near-IR band (unshadowed areas).

ORD was the best indicator for characterizing the structure of scatterplot triangles, and furthermore for representing the accuracy of linear spectral unmixing. Similar to our visual interpretation, the triangles formed in the spectral mixing space constructed by the red and near-IR1 bands were the best match for the pixel clouds with the lowest ORD value of 0.49, corresponding to the highest ALSU value of 75.24%. It was also found that the scatterplot triangles for the unshadowed and shadowed areas were not the most similar in this space. Conversely, its TS value

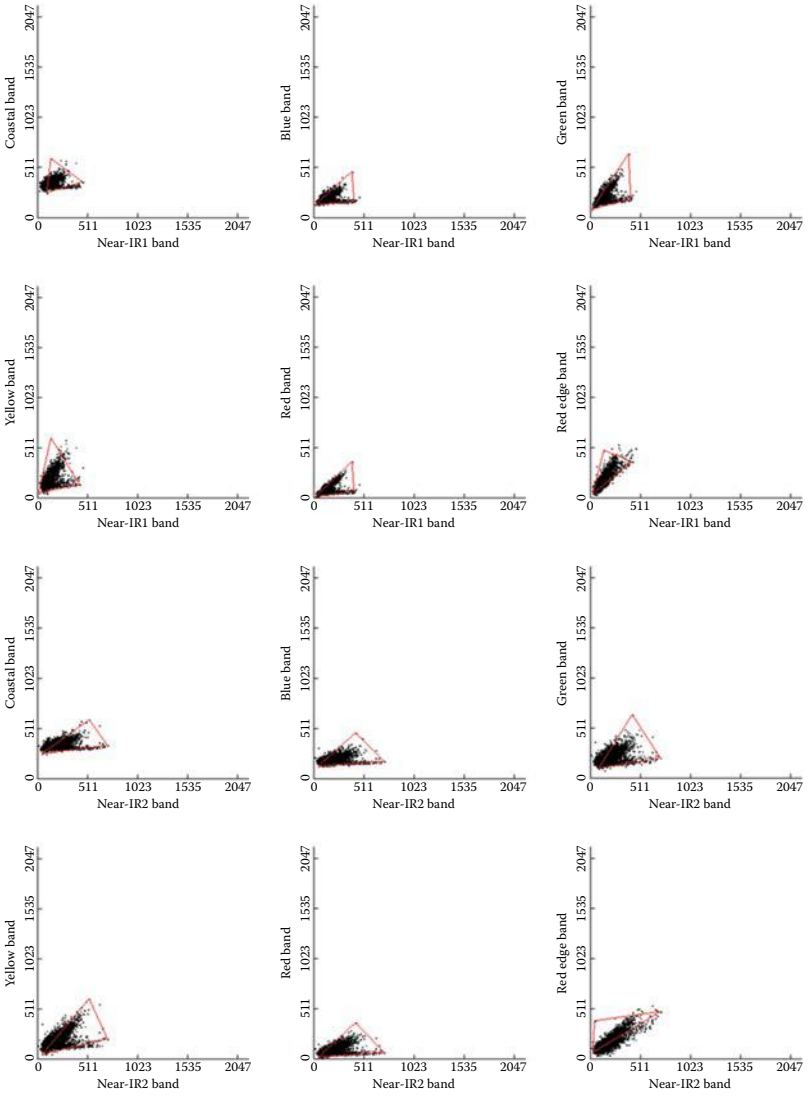


FIGURE 9.6 Scatterplot triangles (triangles) in different spectral mixing spaces constructed by a non-near-IR band and a near-IR band (shadowed areas).

was the highest, at 29.85. Within this space, the result of linear spectral unmixing is demonstrated in Figure 9.8.

9.4 DISCUSSION

As shown previously, our proposed new indicator, ORD, was negatively correlated to the unmixing accuracies of different two-dimensional spectral mixing spaces, with a high R^2 of 0.65. In the space constructed by the red and near-IR1 bands, the value of

TABLE 9.1
Summary of Three Indicators of Scatterplot Triangles (i.e., OR, OD, and ORD) for 12 Spectral Mixing Spaces and Their Corresponding Accuracy of Linear Spectral Unmixing (i.e., ALSU)

Spectral Mixing Space	OR (%)	OD	ORD	ALSU (%)
Coastal + Near-IR1	21.46	18.51	4.58	68.74
Blue + Near-IR1	30.54	65.18	19.77	59.41
Green + Near-IR1	26.68	134.14	33.94	53.25
Yellow + Near-IR1	11.13	52.63	6.06	65.32
Red + Near-IR1	6.99	7.30	0.49	75.24
Red Edge + Near-IR1	0.45	287.62	0.92	65.28
Coastal + Near-IR2	19.57	36.51	6.87	65.84
Blue + Near-IR2	16.94	30.65	4.93	64.32
Green + Near-IR2	11.41	48.12	4.86	61.23
Yellow + Near-IR2	5.25	85.33	2.15	66.21
Red + Near-IR2	17.86	44.24	7.46	62.12
Red Edge + Near-IR2	11.43	36.98	3.75	63.45

ORD (i.e., 0.49) was the lowest, thus corresponding to the highest accuracy of linear spectral unmixing (i.e., 75.24%). This conclusion was also consistent with the visual comparisons from existing studies (Yang et al. 2015; Yang and He 2017), which validated the effectiveness of our proposed new indicator to optimize the best spectral mixing space in two dimensions. On the other hand, it was clear that the scatterplot triangles for the unshadowed and shadowed areas were not the most similar (i.e., TS = 29.85 versus the lowest TS value of 4.82 from the blue + near-IR2 bands) in the spectral mixing space constructed by the red and near-IR1 bands. This observation was not beyond our expectations since the spectral similarity of unshadowed and shadowed areas was determined by their corresponding surface materials. Even identical surface materials would demonstrate different scatterplot triangles in different spectral mixing spaces. As long as both unshadowed and shadowed areas showed good scatterplot triangles, it would be reasonable to implement our linear spectral unmixing process in the unshadowed and shadowed areas separately.

When expanding the utilization of our proposed new indicator from a two-dimensional to an *n*-dimensional spectral mixing space, the unmixed pixels could form a scatterplot of a polyhedron with a vertex close to the coordinate origin, as long as this *n*-dimensional space was able to represent the corresponding *n* end-members. For example, the scatterplot of a three-dimensional space would be a tetrahedron, which could be also regarded as three triangles when projecting it to three two-dimensional spaces. As demonstrated by Yang et al. (2014), the vertices of a scatterplot polyhedron could be identified by the determinant of *n* *n*-dimensional vectors. In a two-dimensional spectral mixing space, we could directly calculate the number of unmixed pixels outside the triangle and their average distance to the triangle. The unmixed pixel should be outside its scatterplot triangle if any of three

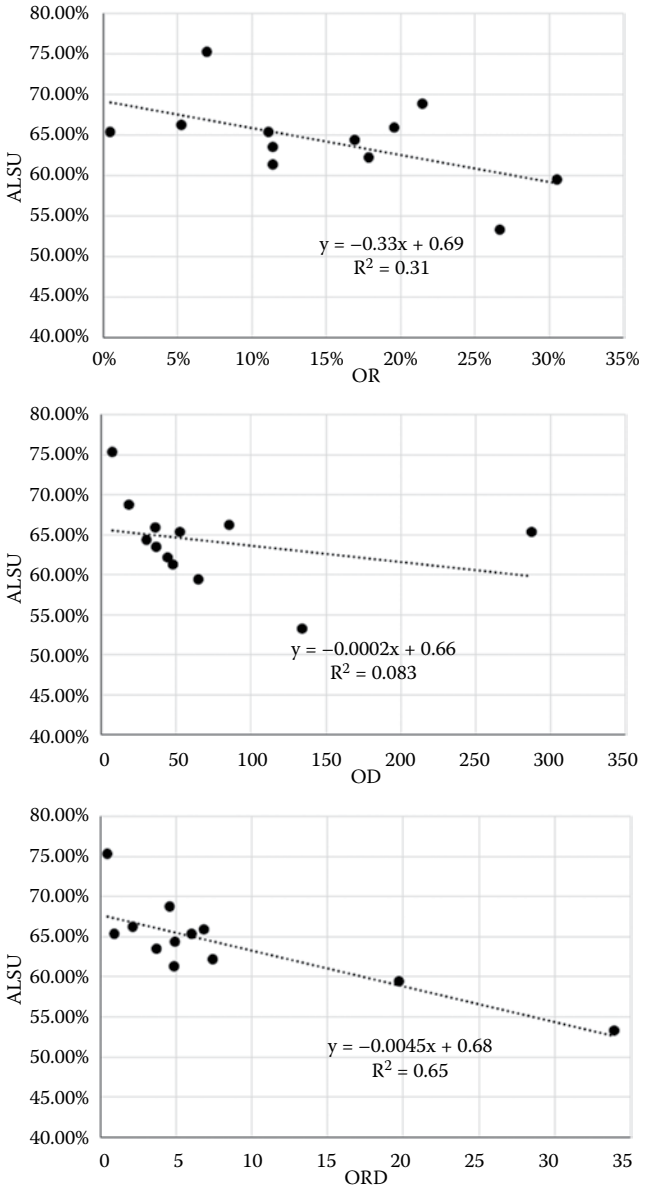


FIGURE 9.7 Accuracy of linear spectral unmixing (i.e., ALSU) versus three indicators of scatterplot triangles (i.e., OR, OD, and ORD) in 12 spectral mixing spaces.

unmixed fractions (i.e., f_V, f_H, and f_L) are not between 0 and 1. Similarly, the unmixed pixel should be outside the polyhedron in an n-dimensional spectral mixing space if any of the unmixed fractions are beyond the range of 0 to 1. However, it would not be a trivial task to obtain the OD or even the ORD value in higher dimensions, as with the OR value, since the distance from a point to a polyhedron in an n-dimensional

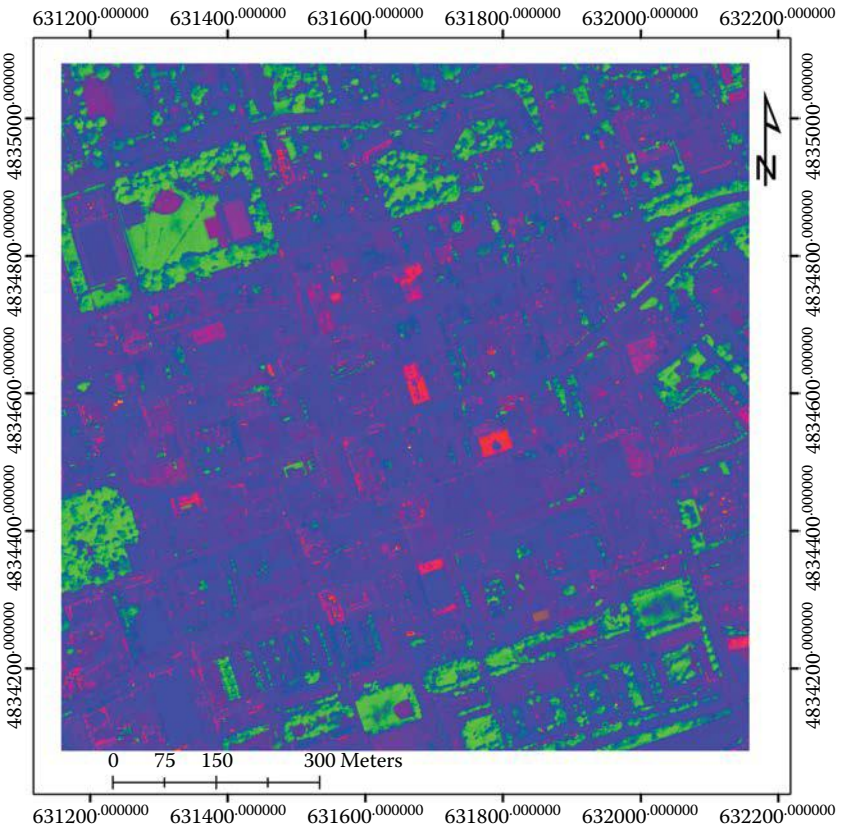


FIGURE 9.8 Fraction map from linear spectral unmixing (red = high albedo; green = vegetation; blue = low albedo).

space could be complicated to calculate. To simplify this calculation, we could project the polyhedron to a series of two-dimensional spaces and then compute the OD value for each two-dimensional space. In this case, the final OD value could be derived by averaging all of the OD values in different two-dimensional spaces.

9.5 CONCLUSIONS

To quantify the structure of a scatterplot triangle for spectrally unmixing high spatial resolution urban imagery, the current study proposed a new indicator, ORD, to represent the proportion of unmixed pixels outside the triangle and their average distance to the triangle. Our experimental results showed that the ORD value was strongly related to the unmixing accuracy, indicating that the quality of linear spectral unmixing was largely dependent upon the structure of scatterplot triangles in different two-dimensional spectral mixing spaces. Consequently, the spectral mixing space constructed by the red and near-IR1 bands was shown to be the best for the vegetation-high albedo-low albedo model, although its triangle similarity of unshadowed and shadowed areas was

not the highest among all of the two-dimensional spectral mixing spaces. The proposed indicator of ORD was proven effective at optimizing the two-dimensional spectral mixing space; however, more work should be conducted in higher dimensions.

REFERENCES

Benz, U. C., P. Hofmann, G. Willhauck, I. Lingenfelder, and M. Heynen. 2004. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry and Remote Sensing* 58 (3): 239–258.

Breiman, L. 2001. Random forests. *Machine Learning* 45 (1): 5–32.

Chen, Y., D. Wen, L. Jing, and P. Shi. 2007. Shadow information recovery in urban areas from very high resolution satellite imagery. *International Journal of Remote Sensing* 28 (15): 3249–3254.

Dare, P. M. 2005. Shadow analysis in high-resolution satellite imagery of urban areas. *Photogrammetric Engineering & Remote Sensing* 71 (2): 169–177.

Graham, R. L. 1972. An efficient algorithm for determining the convex hull of a finite planar set. *Information Processing Letters* 1 (4): 132–133.

Keshava, N. 2003. A survey of spectral unmixing algorithms. *Lincoln Laboratory Journal* 14 (1): 55–78.

Liu, T., and X. Yang. 2013. Mapping vegetation in an urban area with stratified classification and multiple endmember spectral mixture analysis. *Remote Sensing of Environment* 133: 251–264.

Lu, D., and Q. Weng. 2006. Use of impervious surface in urban land-use classification. *Remote Sensing of Environment* 102 (1): 146–160.

Lu, D., and Q. Weng. 2009. Extraction of urban impervious surfaces from an IKONOS image. *International Journal of Remote Sensing* 30 (5): 1297–1311.

Nichol, J., and M. S. Wong. 2007. Remote sensing of urban vegetation life form by spectral mixture analysis of high-resolution IKONOS satellite images. *International Journal of Remote Sensing* 28 (5): 985–1000.

Nichol, J. E., M. S. Wong, R. Corlett, and D. W. Nichol. 2010. Assessing avian habitat fragmentation in urban areas of Hong Kong (Kowloon) at high spatial resolution using spectral unmixing. *Landscape and Urban Planning* 95 (1): 54–60.

Phinn, S., M. Stanford, P. Scarth, A. T. Murray, and P. T. Shyy. 2002. Monitoring the composition of urban environments based on the vegetation-impervious surface-soil (VIS) model by subpixel analysis techniques. *International Journal of Remote Sensing* 23 (20): 4131–4153.

Ridd, M. K. 1995. Exploring a VIS (vegetation-impervious surface-soil) model for urban ecosystem analysis through remote sensing: comparative anatomy for cities. *International Journal of Remote Sensing* 16 (12): 2165–2185.

Shimabukuro, Y. E., and J. A. Smith. 1991. The least-squares mixing models to generate fraction images derived from remote sensing multispectral data. *IEEE Transactions on Geoscience and Remote Sensing* 29 (1): 16–20.

Small, C. 2001. Estimation of urban vegetation abundance by spectral mixture analysis. *International Journal of Remote Sensing* 22 (7): 1305–1334.

Small, C. 2003. High spatial resolution spectral mixture analysis of urban reflectance. *Remote Sensing of Environment* 88 (1): 170–186.

Small, C., and J. W. T. Lu. 2006. Estimation and vicarious validation of urban vegetation abundance by spectral mixture analysis. *Remote Sensing of Environment* 100 (4): 441–456.

Song, C. 2005. Spectral mixture analysis for subpixel vegetation fractions in the urban environment: How to incorporate endmember variability? *Remote Sensing of Environment* 95 (2): 248–263.

Sun, W., B. Chen, and D. W. Messinger. 2014. Nearest-neighbor diffusion-based pan-sharpening algorithm for spectral images. *Optical Engineering* 53 (1): 013107.

Van de Voorde, T., T. De Roeck, and F. Canters. 2009. A comparison of two spectral mixture modelling approaches for impervious surface mapping in urban areas. *International Journal of Remote Sensing* 30 (18): 4785–4806.

Weng, Q., X. Hu, and H. Liu. 2009. Estimating impervious surfaces using linear spectral mixture analysis with multitemporal ASTER images. *International Journal of Remote Sensing* 30 (18): 4807–4830.

Weng, Q., X. Hu, and D. Lu. 2008. Extracting impervious surfaces from medium spatial resolution multispectral and hyperspectral imagery: A comparison. *International Journal of Remote Sensing* 29 (11): 3209–3232.

Wu, C. 2009. Quantifying high-resolution impervious surfaces using spectral mixture analysis. *International Journal of Remote Sensing* 30 (11): 2915–2932.

Wu, C., and A. T. Murray. 2003. Estimating impervious surface distribution by spectral mixture analysis. *Remote Sensing of Environment* 84 (4): 493–505.

Yang, J., and Y. He. 2017. Automated mapping of impervious surfaces in urban and suburban areas: Linear spectral unmixing of high spatial resolution imagery. *International Journal of Applied Earth Observation and Geoinformation* 54: 53–64.

Yang, J., Y. He, and J. Caspersen. 2015. Fully constrained linear spectral unmixing based global shadow compensation for high resolution satellite imagery of urban areas. *International Journal of Applied Earth Observation and Geoinformation* 38: 88–98.

Yang, J., Y. He, and T. Oguchi. 2014. An endmember optimization approach for linear spectral unmixing of fine-scale urban imagery. *International Journal of Applied Earth Observation and Geoinformation* 27: 137–146.

Yang, J., and P. Li. 2015. Impervious surface extraction in urban areas from high spatial resolution imagery using linear spectral unmixing. *Remote Sensing Applications: Society and Environment* 1: 61–71.

Zhou, W., G. Huang, A. Troy, and M. L. Cadenasso. 2009. Object-based land cover classification of shaded areas in high spatial resolution imagery of urban areas: A comparison study. *Remote Sensing of Environment* 113 (8): 1769–1777.

10 Segmentation Scale Selection in Geographic Object-Based Image Analysis

Xiuyuan Zhang, Shihong Du, and Dongping Ming

CONTENTS

10.1 INTRODUCTION

10.1.1 Scale Issues in Geographic Object-Based Image Analysis

The past several years have witnessed the development of geographic object-based image analysis (GEOBIA) (Hay & Castilla, 2008), which has been widely used in dealing with very-high-resolution (VHR) images. Objects refer to homogenous image regions with continuous visual clues or consistent semantics, which are often represented as pixel clusters at multiple scales (Blaschke, 2010). Objects essentially serve as the bridge connecting geographic entities and image elements (Figure 10.1). As reported in Figure 10.1, the geographic entities (Figure 10.1a) have significant

FIGURE 10.1 Relationship among (a) geographic entities, (b) image objects, and (c) image pixels.

differences with image pixels (Figure 10.1c), but they have close relations with image objects (Figure 10.1b); thus, image objects are better than pixels to represent geographic entities, and play an important role in geographic studies. GEOBIA uses image objects to represent geographic entities both spatially and semantically, and extracts their representations and features through spatial, spectral, and temporal scales, so as to generate new geographic information in geographic information system (GIS)-ready format (Hay & Castilla, 2008). It has a wide range of applications (Table 10.1).

As reported in Figure 10.2, object segmentation aims to delineate remote sensing images into meaningful image objects. Among the existing image segmentations, multiresolution segmentation (MRS) provided by a commercial software, eCognition (Trimble, 2011), is the most widely used segmentation method in GEOBIA and it can produce meaningful objects at multiple scales (Baatz & Schäpe, 2000). *Scale* hereafter refers to the segmentation scale, which is an important parameter in the MRS algorithm and measures the largest heterogeneity of generated image objects (Smith, 2010).

Segmentation scale is the most important parameter, dominating the size of segmentation results. Generally, larger scales result in larger objects, and smaller scales result in smaller objects (Myint et al., 2011). Figure 10.3 reports the segmentation results of a building at the multiple scales of 30, 50, and 70. Generally, the 70-scale objects are much larger than those at the scale of 30, and each 70-scale object can be composed of several 30-scale objects because the MRS is boundary constrained. Based on visual interpretation, the building is divided into several image objects at the scale of 30, a phenomenon called oversegmentation (Figure 10.3a); while the

TABLE 10.1
Overview of GEOBIA's Application Field

Types	Application Fields	Examples
Relatively natural landscapes	Forests	Mallinis et al. (2008)
	Wetland	Powers et al. (2012), Bock et al. (2005)
	Shrub land	Stow et al. (2008), Laliberte et al. (2004)
Human-dominated landscapes	Crop	Peña-Barragán et al. (2011)
	Impervious surface	Im et al. (2012)
	Parkland	Yu et al. (2006)
	Urban land covers	Hu et al. (2015), Cleve et al. (2008)

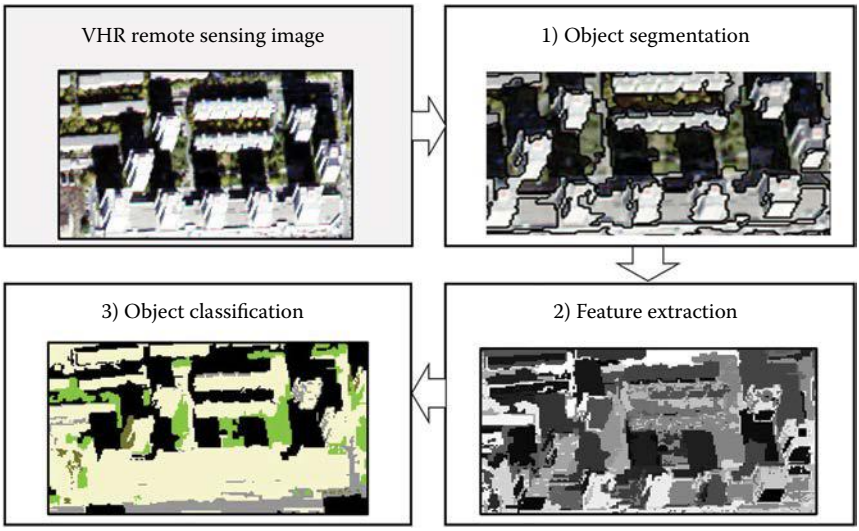


FIGURE 10.2 GEOBIA consists of three steps including object segmentation, feature extraction, and object classification.

FIGURE 10.3 Segmentation results at scales of (a) 30, (b) 50, and (c) 70.

building is accurately delineated at the scale of 50 (Figure 10.3b) because it can be solely segmented. On the contrary, at the scale of 70, the building is mixed with surrounding trees, which are represented as a single image object (Figure 10.3c), which is called undersegmentation (Meinel & Neubert, 2004). Both over- and undersegmentation results are inappropriate to subsequent classification, while only the objects that can exactly represent geographic entities are regarded as accurate results, and the corresponding scales are defined as the optimal scales (Drăguţ et al., 2010; Johnson & Xie, 2011).

10.1.2 Factors Influencing Scale Selection

It has been popularly recognized that the optimal scale is a variable (Huang et al., 2003; Drăguţ et al., 2010). Zhang and Du (2016) proposed that the optimal segmentation scales can be impacted by three factors: category system, surrounding contrast, and internal heterogeneity. First, objects' categories can impact segmentation scales, and usually different object categories are related to different optimal scales. Taking the image scene in Figure 10.4 as an example, the road is oversegmented, but the cars are undersegmented at the same scale. Second, different objects in the same class also

FIGURE 10.4 Segmentation results of a road with several cars where object boundaries are delineated by black lines.

correspond to different optimal scales because of the self-size diversity. Accordingly, different kinds of classes or even objects should be well segmented and classified at different scales. For example, the buildings having strong contrast with surroundings (Figure 10.5a) are oversegmented, while the buildings mixed with the surroundings (Figure 10.5b) are undersegmented at the same scale. Third, the different components within an object may need to be analyzed at different scales. As a component of the building, pixel P_1 (Figure 10.5a) on the building's boundary is mixed with the neighboring road pixels (undersegmentation); while pixel P_2 inside the building

FIGURE 10.5 Comparison of two sets of buildings with (a) strong and (b) weak surrounding contrasts, and their segmentation results at the same scale.

is segmented alone and distinguished from other components (oversegmentation). Therefore, the three influencing factors, different categories, surrounding contrasts, and internal heterogeneities of objects, often correspond to different optimal scales. The first factor is global, while the last two are local and represent local heterogeneity.

In the following, we select the optimal scales for object segmentation so as to generate accurate objects and avoid over- and undersegmentations. For selecting optimal scales of objects, the three influencing factors should be taken into consideration. The used terminologies in this chapter are outlined in Table 10.2.

10.2 STUDY AREA AND EXPERIMENTAL DATA

We selected a region of Beijing as the study area (Figure 10.6). This area, located at Peking University, contains many natural landscapes including vegetation, water, and soil, as well as numerous artificial structures, e.g., many kinds of buildings (commercial buildings, apartments, shantytowns, and stadiums) and roads. Accordingly, this study area with high heterogeneity can hardly be segmented with fixed scales. A WorldView-2 image, acquired in the summer of 2010, was used (Figure 10.6), and its multispectral bands were first merged with the panchromatic band to produce a pan-sharpened image of 0.5 m resolution with eight bands. This image is used in the following sections to evaluate the scale-selection methods.

10.3 POSTSEGMENTATION SCALE SELECTIONS

As demonstrated previously, segmentation scale can significantly impact GEOBIA results, but it can be variable owing to three influencing factors: categories, surrounding contrasts, and internal heterogeneities. Accordingly, we need to select

TABLE 10.2
Terminologies Repeatedly Used in This Chapter

Terminologies	Meanings
GEOBIA	Geographic object-based image analysis is a strategy for extracting image object information and complements traditional per-pixel image analysis.
Image objects	Homogenous regions of images that are collections of adjacent pixels and generated based on a certain rule.
Multiscale segmentation	Segmentation results at multiple scales organized in a hierarchical structure, with the finest scale on the bottom and the coarsest on the top.
Object classification	A technique using objects as basic units and features as clues to label each object with a defined category.
Object heterogeneity	An important criterion for measuring the internal nonstationarity of objects, which is basic to object segmentation.
Object segmentation	A technique for dividing an image into spatially continuous and non-overlapping patches, with each patch spatially delineating an object.
Segmentation scale	An important parameter in object segmentation and measures the largest heterogeneity of generated image objects.

FIGURE 10.6 WorldView-2 image (in band combination 5/3/2, true color) of study area.

the optimal one from multiple scales. Most existing scale-selection methods choose optimal scales from multiscale segmentation results, i.e., they first segment satellite images at multiple scales and then choose a scale level as the optimal one. These methods are so-called postsegmentation scale selections and mainly fall into two types: unsupervised and supervised (Table 10.3).

10.3.1 Unsupervised Scale Selection

Unsupervised methods essentially define several indicators to evaluate multiscale segmentation results, and select the most accurate one as the final segmentation results. These methods need no prior information, and thus are efficient for most GEOBIA. There are two indicators for evaluating segmentation results, i.e., local variance (LV) (Drăguţ et al., 2010, 2014) and global score (GS) (Johnson & Xie, 2011).

TABLE 10.3
Scale-Selection Methods

Types	Criterions for Scale Selection	Examples
Unsupervised	Local variance	Drăguţ et al. (2010, 2014)
	Local variance and correlations	Johnson and Xie (2011)
Supervised	Classification accuracy	Zhang and Du (2016)
	Spatial overlap	Clinton et al. (2010), Guo and Du (2016), Witharana and Civco (2014), Zhang et al. (2015a)
	Feature importance	Zhang and Du (2016)

10.3.1.1 Local Variance

LV roots in the idea of the relationship between the spatial structure of images, the size of the objects in the real world, and pixel resolution (Woodcock & Strahler, 1987). Woodcock and Strahler (1987) first calculated the value of the standard deviation (SD) in a small convolutional window (3 × 3), and then computed the mean of these values over the whole image. Accordingly, the obtained value is an indicator of the local variance in the image (Drăguţ et al., 2010). Kim et al. (2008) derived the SD from each segmented object instead of a moving window, and the value directly measures local variance of an object. The mean SD of all objects in the image is denoted as LV (Equation 10.1) (Drăguţ et al., 2010)

$$LV_l = \frac{\sum_{i=1}^{N_l} SD_{l,i}}{N_l} \tag{10.1}$$

where l refers to the lth scale level in the MRS results, and LV_l denotes the LV at this level; N_l is the number of segmented objects at the lth scale level, and $SD_{l,j}$ represents the standard deviation of the ith object. It is assumed that with the scale growing, the size of a segment is getting larger, and its SD increases continuously up to the point that it matches the object in the real world. Accordingly, the break point of LV indicates the optimal scale. To detect the break point, Drăguţ et al. (2010) then defined the rate of change (ROC) of LV (Equation 10.2) to access the dynamics of LV from one scale level to another, and hypothesized that the most obvious peaks in the ROC curve indicate the optimal segmentation scales

$$ROC_l = \frac{LV_{l+1} - LV_l}{LV_l} \times 100\% \tag{10.2}$$

This method was tested on the study area situated in Beijing (Figure 10.6), and the dynamics of changes in LV and ROC are reported in Figure 10.7, where LV increases

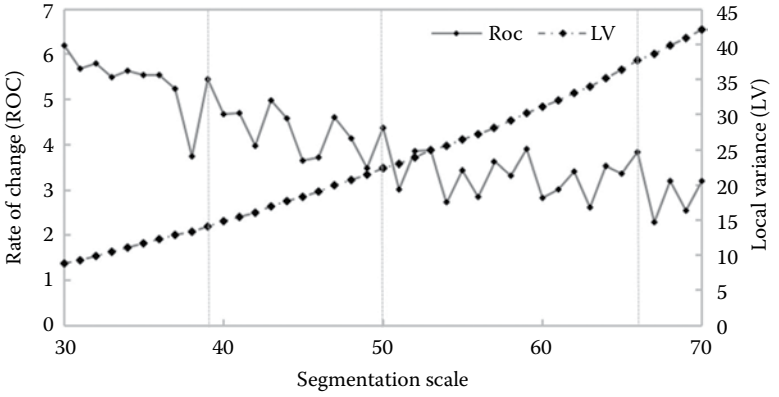


FIGURE 10.7 Dynamics of changes in LV and ROC with increasing scale parameter. Three selected scales are indicated by dotted vertical lines.

with the increasing scales while ROC follows an opposite trend. The scales of 39, 50, and 66 were manually selected as the optimal scale levels. At a scale of 39, most buildings are well segmented (Figure 10.8b). The shadow is best segmented at a scale of 50 (Figure 10.8c) because the shadow segments are easily merged with the dark trees or buildings at coarser scales. A scale of 66 generates accurate segments for soil and water (Figure 10.8d), while they are oversegmented at the finer scales. Therefore, ROC is an indicator for measuring the amount of change in LV from one level to another, and it is important to select the optimal scale level.

10.3.1.2 Global Score

Johnson and Xie (2011) proposed that scale selections should not consider only intrasegment heterogeneity, but also intersegment similarity. They gave a definition of a good segmentation that should meet four conditions: (1) segments should be uniform and homogeneous, (2) segments are significantly different from neighboring regions, (3) segments have a simple interior without many holes, and (4) segment boundaries should be simple, not ragged, and spatially accurate. However, most unsupervised evaluation methods solely involve intrasegment heterogeneity and ignore intersegment similarity measurements for each segment (Zhang et al., 2008). Accordingly, Zhang et al. (2008) proposed an indicator for scale selection, GS, which takes into consideration both intrasegment heterogeneity and intersegment similarity.

First, the intrasegment heterogeneity is calculated by an area-weighted variance (Equation 10.3)

$$WV_l = \frac{\sum_{i=1}^{N_l} a_{l,i} \times SD_{l,i}}{\sum_{i=1}^{N_l} a_{l,i}} \tag{10.3}$$

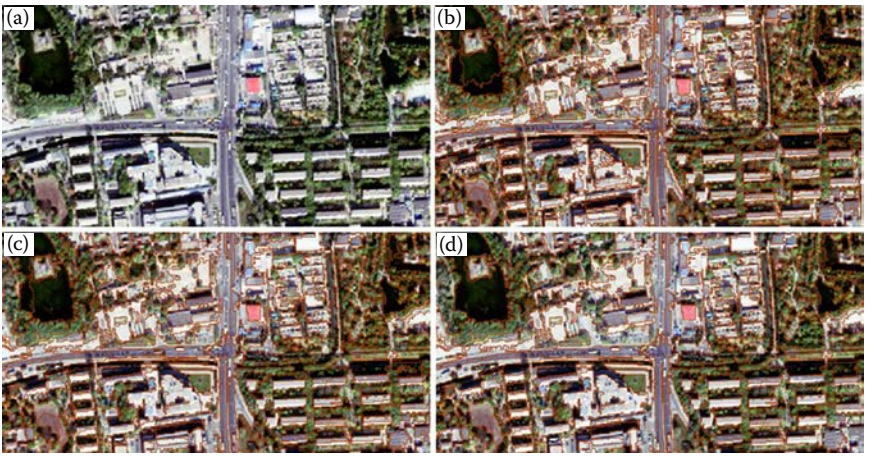


FIGURE 10.8 (a) Subimage of the study area, and three segmentation results based on different scale parameters of (b) 39, (c) 50, and (d) 66.

Similar to Equation 10.1, l also represents the lth scale level in multiscale segmentation results, while WV_l denotes the area-weighted variance at this level; N_l is the number of segmented objects at the lth scale level; and $a_{l,i}$ and $SD_{l,i}$ represent the area and the standard deviation of the ith object, respectively. Parameter WV_l is used for the global calculation so that large segments contribute more to the global score than small ones.

On the other hand, the intersegment similarity is measured by Moran's I index, a spatial autocorrelation metric that measures, on average, how similar a region is to its neighbors (Fotheringham et al., 2000), which was found as a good indicator of segmentation quality in previous segmentation evaluation studies (Espindola et al., 2006; Kim et al. 2008). For this study, Moran's I at the lth scale level (MI_l) was calculated using Equation 10.4

$$MI_l = \frac{N_l \times \sum_{i=1}^{N_l} \sum_{j=1}^{N_l} w_{ij}^{\ l} \times (y_i^{\ l} - \bar{y}^l) \times (y_j^{\ l} - \bar{y}^l)}{\sum_{i=1}^{N_l} (y_i^{\ l} - \bar{y}^l)^2 \times \left(\sum \sum_{i \neq j} w_{ij}^{\ l} \right)} \tag{10.4}$$

$y_i^{\ l}$ refers to the feature value of the ith object at the lth scale level, and $\bar{y}^l$ is the mean value in the image. Each weight $w_{ij}^{\ l}$ measures the spatial adjacency degree between the ith and jth objects. Accordingly, only neighboring objects are considered, with $w_{ij}^{\ l} = 1$; otherwise $w_{ij}^{\ l} = 0$. In general, low MI_l values indicate small intersegment similarity, which is desirable for an image segmentation. Both WV_l and MI_l are rescaled into the same range [0–1] by normalization. The GS is then calculated based on the two normalized indicators: $GS_l = WV_l + MI_l$. A low GS_l values corresponds to good segmentation results.

Five scales were selected to segment the WordView-2 image in the Beijing urban area (Figure 10.9a), and their corresponding GS values were calculated. The optimal segmentation result (Figure 10.9b) was selected among multiscale segmentations

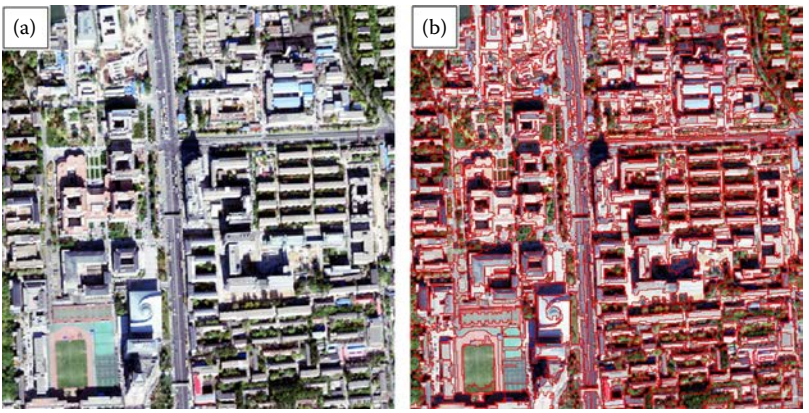


FIGURE 10.9 Comparison of (a) original WordView-2 image and (b) its segmentation results using the scale parameter of 40.

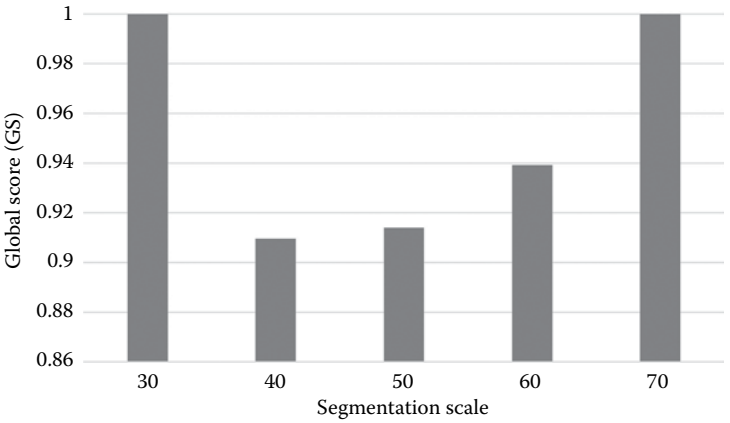


FIGURE 10.10 Dynamics of changes in GS with the increase of the scale parameter.

based on the GS curve (Figure 10.10). As shown in Figure 10.10, the image segmentation performed best at a scale of 40 and produced the lowest GS (0.9097). At the scale of 40, objects are, on average, most homogeneous internally and most different from their neighbors, and thus according to the GS value, 40 is the optimal scale for segmenting this image. In addition, the scales of 50 and 60 also produced good segmentation results with pretty high accuracies.

10.3.2 Supervised Scale Selection

The unsupervised selection methods mainly consider local heterogeneities, but they totally ignore a category's influence on scale selection. Accordingly, supervised methods are proposed to resolve this issue. The supervised scale selections fall into three types: classification, boundary, and feature based. They all aim to measure the discrepancy between segmentation results and reference objects, which can directly reveal segmentation quality. If the discrepancy is small, the segmentation quality is high. In an ideal case, if no discrepancy exists, the segmented regions are identical to the reference objects (Zhang et al., 2015a).

10.3.2.1 Selection Based on Classification Accuracy

Semantic discrepancies between generated objects and references are very important to measure the object quality (Zhang & Du, 2016). Accordingly, object classification results at diverse scales are first used to quantitatively evaluate multiscale segmentation results, and then to select the optimal scale for each category.

Object classifications are conducted at different scales, and the same training and testing samples are used for all scale layers. These samples are first collected per pixel, and then aggregated with the increase of scales. For example, a large-scale object may contain several training samples, and these samples can belong to different classes; while this object will be labeled as the class with the largest area ratio. Using these samples and machine learning methods (e.g., support vector machine, random forest (RF), and artificial neural network), object classification results and accuracies at

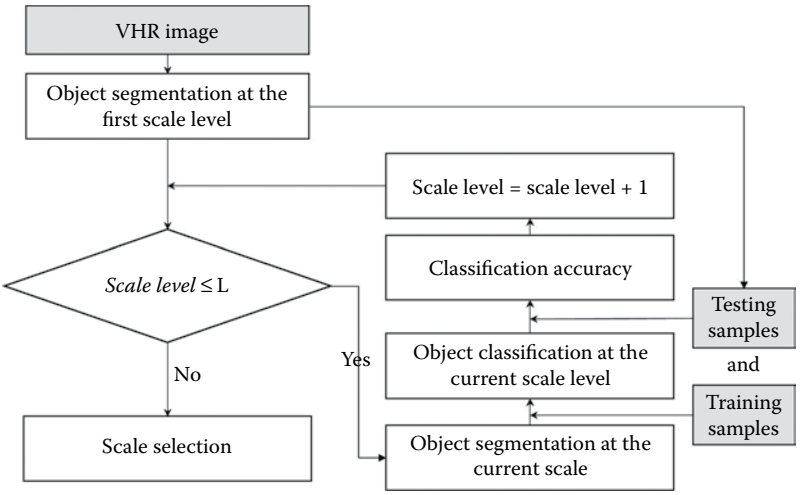


FIGURE 10.11 Scale selection based on object classification accuracies at multiple scales.

diverse scales can be obtained (Figure 10.11). Given an object category c_i, its producer's accuracy (Powers, 2011) at the lth scale level $PA_{c_i}^l$ is employed to measure the c_i's segmentation quality at this scale. Accordingly, the optimal scale of c_i is defined as the one with the largest accuracy, i.e., $OptScale_{c_i} = \arg\max_l \{PA_{c_i}^l \mid 1 \le l \le L\}$. For all categories, the overall accuracy at the lth scale level OA_l measures the average quality of all objects at this scale; thus, the optimal scale of all categories is defined as $OptScale = \arg\max_{l \le L} \{OA_l\}$.

In Zhang and Du (2016), this strategy was used to select the optimal scale parameter for the study area (Figure 10.12). Here, the satellite image is classified into six categories using random forest classifier, and their accuracies are shown in Table 10.4, and their overall accuracies at multiple scales are reported in Figure 10.13. By reference to the classification accuracy, the scale of 60 produced the most accurate classification results, and thus it is the optimal scale for GEOBIA in this study area. In detail, the roads are well segmented at scale 70, and the soils at scale 50, indicating that the different geographic entities have different intrinsic scales and should be segmented at different scales. In conclusion, the scale-selection method-based classification accuracy considers a category's impact, and can select the optimal scale for each category, but it ignores local heterogeneities within categories, so the selected scales cannot totally resolve the three influencing factors.

10.3.2.2 Selection Based on Spatial Overlap

Apart from semantic discrepancy, spatial discrepancy between segments and object references is also important to measure the qualities of generated objects. Accordingly, spatial overlapping degrees between segments and object references are presented to evaluate segmentation results, and the scale with the largest overlapping degree will be selected as the optimal scale from multiresolution segmentations (Zhang et al., 2015b). This kind of method is mainly divided into two steps, i.e., matching and calculating discrepancy. First, segments are matched to object references by

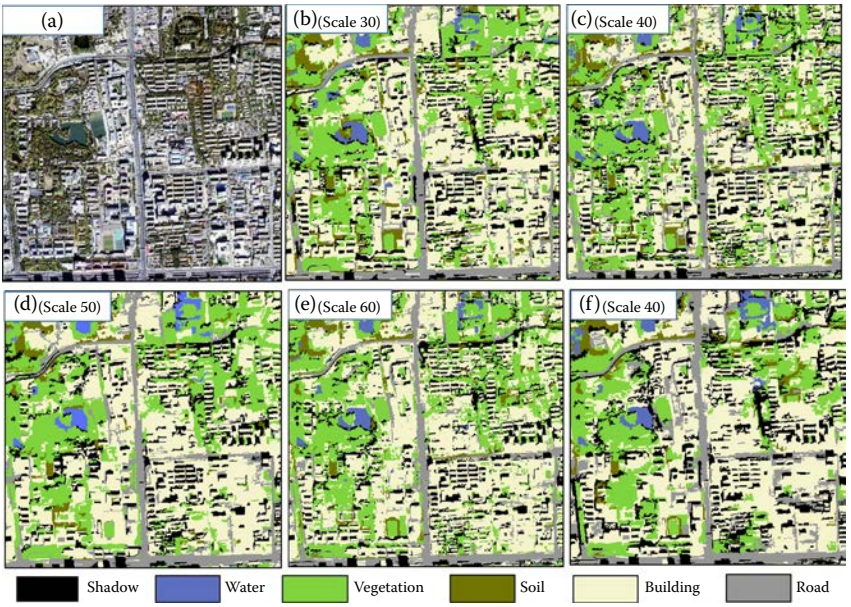


FIGURE 10.12 Object classification results of the (a) WorldView-2 image at (b–f) five scales. (From Zhang, X., & Du, S. 2016. *Remote Sensing of Environment*, 178, 172–190.)

boundary matching or region overlapping (Clinton et al., 2010). Then the discrepancy measures are calculated on an edge versus non-edge basis or by prioritizing the edge pixels according to their distance to the reference (Estrada & Jepson, 2009; Albrecht, 2010). However, in most cases, especially in remote sensing image segmentation, the region overlapping strategy is more often used (Witharana & Civco, 2014).

TABLE 10.4
Per-Category and Overall Classification Accuracies at Different Segmentation Scales

Per-Category Accuracy (%)	Scale 30	Scale 40	Scale 50	Scale 60	Scale 70	Selfhood Scale
Shadow	81.2	81.2	83.1	83.1	79.0	**85.6**
Water	**96.4**	95.6	95.6	90.6	86.9	95.0
Vegetation	79.8	81.7	85.3	86.7	85.4	**93.1**
Soil	61.0	70.1	80.7	76.2	74.3	**86.7**
Building	80.6	75.3	81.3	82.5	80.9	**92.5**
Road	71.3	71.3	76.7	81.4	83.3	**89.4**
Overall accuracy (%)	82.5	81.7	84.6	84.8	82.2	**92.4**

Source: Zhang, X., & Du, S. 2016. *Remote Sensing of Environment*, 178, 172–190.

Note: The bold numbers refer to the largest accuracy for each category, and the underlined numbers represent the second largest accuracy.

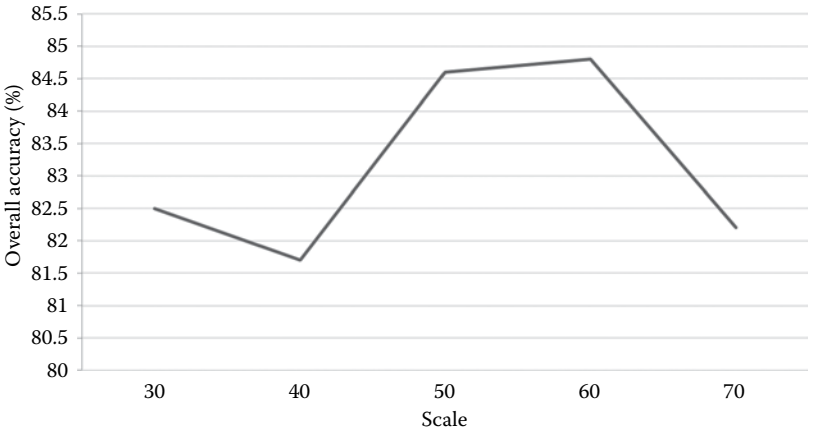


FIGURE 10.13 Overall accuracies of object classification results at multiple scales.

Zhang et al. (2015a) proposed region-based precision and recall to measure the boundary discrepancy between segmented regions and reference objects. The two indicators are calculated based on region overlapping and presented as a point or a curve in a precision-recall space. First, precision is calculated based on the reference-to-segment overlapping. In detail, the reference objects are matched to the segments, and the segment having the largest overlapping area with the reference object is identified. Given the segmentation result at the lth scale level S^l, i.e., N_l segments $\{S_1^l, S_2^l, \ldots, S_{N_l}^l\}$, and the reference R with M objects $\{R_1, R_2, \ldots, R_M\}$, the precision measure is calculated by matching $\{R_i \mid 1 \leq i \leq M\}$ to each segment $S_j^l (1 \leq j \leq N_l)$. For calculating the precision at the lth scale level (i.e., precision$_l$, Equation 10.5), the matched reference object ($R_{\max}$) of S_j^l should be first identified, where $R_{\max}$ has the largest overlapping area with S_j^l (Figure 10.14), and $|*|$ measures the area of the region.

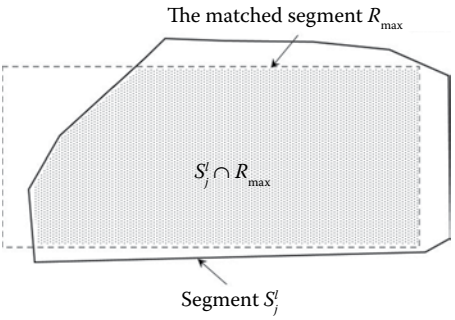


FIGURE 10.14 Relationship between the segment S_j^l and the matched reference object $R_{\max}$.

$$\text{precision}_l = \frac{\sum_{j=1}^{N_l} |S_j^l \cap R_{\max}|}{\sum_{j=1}^{N_l} |S_j^l|} \tag{10.5}$$

Similarly, recall at the lth scale level, recall_l, can be measured by matching $\{S_j^l \mid 1 \le j \le N_l\}$ to each segment $R_i (1 \le i \le M)$ (Equation 10.6), where the matched segment ($S_{\max}^l$) of R_i is first identified (Figure 10.15).

$$\text{recall}_l = \frac{\sum_{i=1}^{M} \left|R_i \cap S_{\max}^l\right|}{\sum_{i=1}^{M} \left|R_i\right|} \tag{10.6}$$

In practice, a limited number of reference objects are manually delineated, rather than the whole imagery, and the precision and recall values calculated based on these reference objects are used to evaluate the accuracies of different scale segmentation results. Generally, large precision and recall values indicate high segmentation quality. In addition, the two indicators can measure both over- and undersegmentations. When an image is mostly oversegmented at the lth scale level, the precision_l is large but recall_l decreases to incur the penalty. In an extreme oversegmentation case with each pixel being a segment, the precision_l is 1 and the recall_l is as low as M/N_{pixel}, where N_{pixel} is the number of pixels in the image. In contrast, when an image is mostly undersegmented, the recall_l is high but the precision_l decreases. In an extreme undersegmentation case in which the entire image is viewed as a segment, the recall_l is 1 and the precision_l decreases to $\left|R_{largest}\right|/N_{pixel}$, where $R_{largest}$ represents the largest reference object.

Based on the recall and precision, several indicators are presented to evaluate the lth scale segmentation results including Sum_l, ED_l, ED'_l, and F_l: $Sum_l = \text{precision}_l + \text{recall}_l$, $ED_l = \sqrt{\text{precision}_l^{\,2} + \text{recall}_l^{\,2}}$, $ED'_l = \sqrt{(1-\text{precision}_l)^2 + (1-\text{recall}_l)^2}$ (Clinton et al., 2010), and F_l is defined as

$$F_l = \frac{1}{\alpha \times \dfrac{1}{\text{precision}_l} + (1-\alpha) \times \dfrac{1}{\text{recall}_l}} \tag{10.7}$$

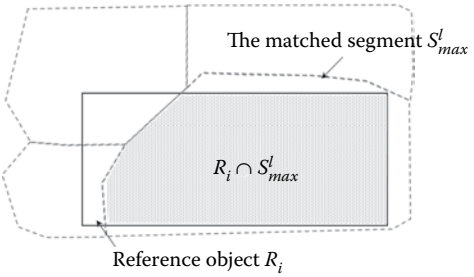


FIGURE 10.15 Relationship between the reference object R_i and its matched segment $S_{\max}^l$.

where the weight parameter α is usually set as 0.5. Large Sum_l, ED_l, and F_l values and a small ED'_l value indicate high segmentation quality.

The presented precision and recall can be used to measure the segmentation accuracy, but they cannot evaluate over- or undersegmentations directly. Accordingly, Guo and Du (2016) used three other indicators: OS_l for oversegmentation (Equation 10.8), US_l for undersegmentation (Equation 10.9), and their combination RMS_l (Equation 10.10).

$$OS_l = \sum_{j=1}^{N_l} \left(1 - \frac{\left| S_j^l \cap R_{\max} \right|}{\left| R_{\max} \right|} \right) \tag{10.8}$$

$$US_l = \sum_{j=1}^{N_l} \left(1 - \frac{\left| S_j^l \cap R_{\max} \right|}{\left| S_j^l \right|} \right) \tag{10.9}$$

$$RMS_l = \sqrt{\frac{OS_l^2 + US_l^2}{2}} \tag{10.10}$$

Both OS_l and US_l can measure how image segment S_j^l fits with the corresponding reference object R_{max}, and RMS_l integrates the two measurements into one. In the ideal case, $OS_l = 0$, $US_l = 0$, and $RMS_l = 0$ means the best segmentation, which can hardly be achieved. In practice, the image is either over- or undersegmented, and small values of the indicators refer to accurate segmentation results. Guo and Du (2016) used the three indicators to select the optimal scales for segmenting four kinds of buildings (Figure 10.16). As reported in Table 10.5, as segmentation scales increase,

FIGURE 10.16 (a) QuickBird image and (b) the reference building objects, which are classified into four categories. (From Guo, Z., & Du, S. 2016. *GIScience & Remote Sensing*, 54(1), 38–63.)

TABLE 10.5
Quantitative Evaluation on Multiscale Segmentations for Four Kinds of Buildings

Scale	Single Apartment			Multifunction Building			Tall Building			Shanty Town		
(l)	OS_l	US_l	RMS_l	OS_l	US_l	RMS_l	OS_l	US_l	RMS_l	OS_l	US_l	RMS_l
30	0.833	0.153	0.61	0.938	0.127	0.678	0.921	0.13	0.667	0.982	0.068	0.701
40	0.746	0.176	0.558	0.909	0.145	0.661	0.882	0.147	0.643	0.972	0.083	0.696
50	0.67	0.195	0.516	0.879	0.156	0.642	0.842	0.175	0.623	0.961	0.096	0.69
60	0.599	0.221	0.482	0.85	0.167	0.625	0.803	0.197	0.605	0.95	0.103	0.683
70	0.538	0.249	0.458	0.819	0.182	0.61	0.765	0.216	0.589	0.94	0.121	0.678
80	0.483	0.285	0.444	0.789	0.195	0.594	0.724	0.256	0.579	0.927	0.13	0.672
90	0.438	0.317	0.435	0.761	0.21	0.583	0.687	0.288	0.572	0.904	0.149	0.661
100	**0.404**	**0.353**	**0.434**	0.728	0.236	0.573	0.65	0.327	0.568	0.891	0.156	0.653
110	0.374	0.388	0.438	0.697	0.263	0.566	0.613	0.355	0.56	0.877	0.16	0.645
120	0.349	0.425	0.446	**0.666**	**0.287**	**0.556**	0.584	0.377	0.555	0.858	0.179	0.641
130	0.32	0.472	0.458	0.633	0.321	0.557	0.559	0.404	0.553	0.848	0.194	0.639
140	0.299	0.514	0.472	0.601	0.356	0.559	**0.531**	**0.43**	**0.551**	0.832	0.206	0.632
150	0.284	0.554	0.489	0.575	0.394	0.559	0.506	0.455	0.555	0.801	0.235	0.623
160	0.269	0.588	0.503	0.545	0.427	0.56	0.48	0.481	0.554	0.772	0.262	0.615
170	0.255	0.623	0.518	0.525	0.449	0.562	0.461	0.503	0.557	0.749	0.292	0.612
180	0.244	0.65	0.531	0.5	0.488	0.566	0.44	0.52	0.554	**0.731**	**0.317**	**0.61**
190	0.234	0.673	0.541	0.471	0.516	0.566	0.419	0.544	0.555	0.725	0.326	0.611
200	0.223	0.703	0.555	0.445	0.548	0.57	0.402	0.574	0.564	0.705	0.348	0.612

Source: Guo, Z., & Du, S. 2016. *GIScience & Remote Sensing*, 54(1), 38–63.

Note: The optimal scale for each building type is bold.

the values of US_l increase, while the values of OS_l decrease. This result is consistent with our theory that a small segmentation scale will result in oversegmentation and a large one will result in undersegmentation. The RMS_l values have a different trend, that is, as scales increase, the RMS_l values first decrease and then increase. The optimal segmentation scales are selected at the break points with lowest RMS_l values. Therefore, the optimal scales for single apartments, multifunction buildings, tall buildings, and shantytowns are selected as 100, 120, 140, and 180, respectively.

Accordingly, this method is effective to select the optimal scale for each reference object, and the selected scales can avoid the three influencing factors.

10.3.2.3 Selection Based on Feature Importance

Recently, a novel scale-selection method was proposed based on feature evaluation. In this strategy, scales are scored and ranked by reference to feature importance, and the optimal scale with the largest feature importance will be selected from multiple scales (Zhang & Du, 2016).

Zhang and Du (2016) employed an RF to measure feature importance, where feature importance refers to the contribution degree of a feature to classification

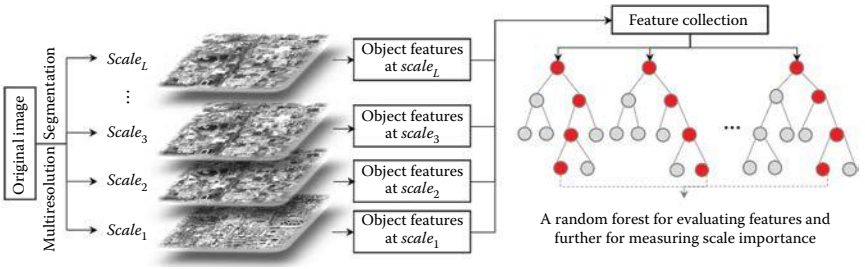


FIGURE 10.17 Procedure of feature-based scale selection using random forest. (From Zhang, X., & Du, S. 2016. *Remote Sensing of Environment*, 178, 172–190.)

(Figure 10.17). RF is an assembled classifier based on decision trees (Friedl & Brodley, 1997) and is skilled at measuring feature importance (Breiman, 2001). However, it ignores local heterogeneity during feature evaluation, which is an important influencing factor on scale. Accordingly, Zhang and Du (2016) advanced RF and proposed a local feature evaluation method named local Gini importance.

For this method, the original VHR image is first segmented at multiple scales ($Scale_1, Scale_2, \ldots, Scale_L$) using MRS (Baatz & Schäpe, 2000). Then spectral, geometrical, and texture features are extracted at each scale (Ramakant & Ramesh, 1980). For each image object, the features are assigned to its covering pixels, leading to each pixel being able to be characterized by M object features at each scale. Accordingly, a pixel has L (scale) $\times$ M(features/scale) features in the stacked L-scale image.

As demonstrated previously, the optimal scales can be learned from the L scales ($Scale_1, .., Scale_L$) based on the importance of $L \times M$ features. These features, however, are high dimension, redundant, and greatly correlated. Accordingly, the random forest algorithm, skilled in analyzing high-dimension and correlated features, is chosen to evaluate feature importance and further learn the optimal scales. Random forest is trained by supervised samples and is able to measure feature importance using the Gini importance approach (Breiman, 2001; Pal, 2005). Gini importance accumulates the decreases in Gini coefficients caused by a feature in all decision trees. The more Gini decrease a feature causes, the more important it is. Since we want to extract optimal scales considering local heterogeneity, a local Gini importance is developed to measure the feature importance for pixels instead of the whole image. When classifying a pixel $Pixel_p$ using random forest, its visited nodes in all decision trees are labeled (Figure 10.17). Each node stores the used feature and the decrease in Gini, which measures the importance of the used feature for classifying $Pixel_p$. Accordingly, the importance of the jth feature $FeatureImp_j(Pixel_p)$ ($1 \le j \le L \times M$) can be measured by accumulating Gini decreases of the jth feature in these visited nodes (Figure 10.17). Then, the importance of the lth scale ($1 \le l \le L$) for classifying $Pixel_p$, i.e., $ScaleImp_l(Pixel_p)$, is the sum of the importance of all the features at the lth scale (Equation 10.11).

$$ScaleImp_l(Pixel_p) = \sum_j FeatureImp_j(Pixel_p)(j \in Scale_l) \tag{10.11}$$

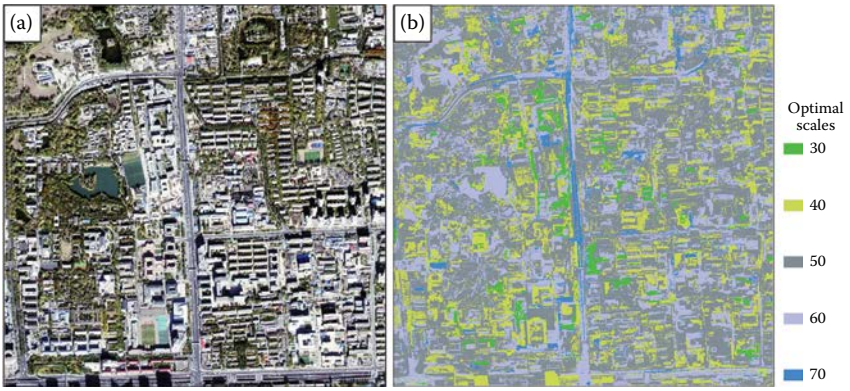


FIGURE 10.18 (a) Original WorldView-2 image and (b) the selected optimal scales. (From Zhang, X., & Du, S. 2016. *Remote Sensing of Environment*, 178, 172–190.)

Among the L scales, the most important scale with the largest $ScaleImp_l(Pixel_p)$ is regarded as the optimal scale of $Pixel_p$ and denoted by $SelfhoodScale(Pixel_p)$ (Equation 10.12).

$$SelfhoodScale(Pixel_p) = \arg\max_{l=1}^{L} ScaleImp_l(Pixel_p) \tag{10.12}$$

Zhang and Du (2016) used this method to select the optimal scale for each spatial unit of the VHR image in Beijing, and the selection result is shown in Figure 10.18.

The optimal scales change with categories (Table 10.6) according to training samples. For example, 69.2% of apartment buildings fit the scale of 40 because most buildings are oversegmented at scale 30, but undersegmented at scale 50. Accordingly, scale 40 is the optimal segmentation scale for segmenting an apartment building. Figure 10.18 shows that optimal scales not only change from category to category, but also vary among individual objects. For example, some road objects use scale 60 as their optimal scales, but the others choose the scale

TABLE 10.6
Proportions of Different Optimal Scales for Each Category

Proportion (%)	Scale 30	Scale 40	Scale 50	Scale 60	Scale 70
Shadow	0.2	26.2	**53.8**	11.7	8.1
Water	0	0.9	1.1	**96.7**	1.3
Vegetation	3.0	6.6	**57.2**	28.3	4.9
Soil	0.4	9.3	40.7	**41.2**	8.4
Building	16.7	**69.2**	12.1	1.2	0.8
Road	0.9	15.9	18.4	25.3	**39.5**

Note: The bold numbers represent the largest proportions for each category.

of 70, because the two kinds of roads have different conditions considering the distributions of moving cars. In addition, diverse optimal scales (30, 40, 50, and 60) are employed by building objects because buildings have different contrasts with their surroundings.

In summary, the feature-based method considers both categories and local heterogeneities, and is effective to select optimal scales for objects.

10.4 LATEST ADVANCES OF PRESEGMENTATION SCALE ESTIMATION

The past several years have witnessed the development of scale-selection methods in two aspects. First, the concept of segmentation scale has been generalized instead of only being a parameter in MRS, and it also appears in other segmentation methods, e.g., mean shift. Second, presegmentation scale estimation is proposed to select the scale by using spatial statics and it does not depend on multiscale segmentation results.

Ming et al. (2016b) generalized the commonly used segmentation scale parameters into three general aspects: spatial parameter h_s (spatial distance between classes or range of spatial correlation), area parameter M (the area or pixel number of the minimum meaningful object), and attribute/spectral parameter h_r (attribute difference between classes). Aiming to estimate the three scale parameters, Ming et al. (2012) proposed a spatial statistics-based scale estimation method for GEOBIA. The main idea of this proposed method is to use the average LV graph (Ming et al., 2015) or the semivariogram (Ming et al., 2012) to estimate the optimal h_s. Next, the estimation of the optimal M is based on simple geometric computation and the optimal h_r is based on the LV histogram. Taking mean shift segmentations as an example, Ming et al. (2015) used classical geospatial and spectral statistics to estimate the three parameters h_s, h_r, and M in GEOBIA. Figure 10.19 demonstrates the workflow of scale estimation based on the LV (Woodcock & Strahler, 1987) and spectral statistical methods, where $(I)_{xy}$ represents an image with x rows and y columns, i indicates the repeating times of the computation, and ALV refers to the average LV value.

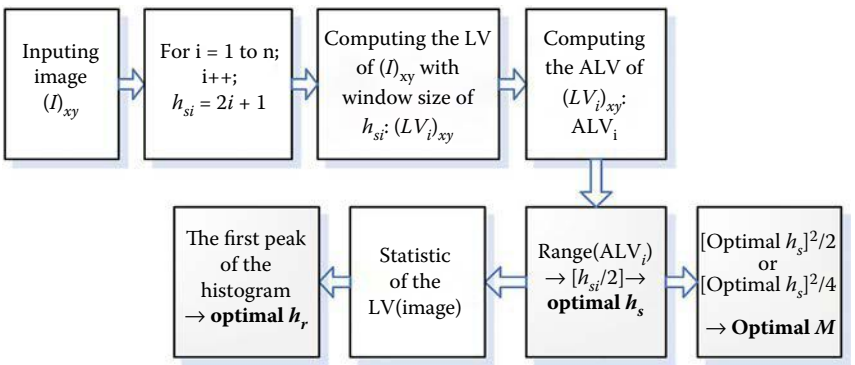


FIGURE 10.19 Workflow of estimating the optimal scale parameters.

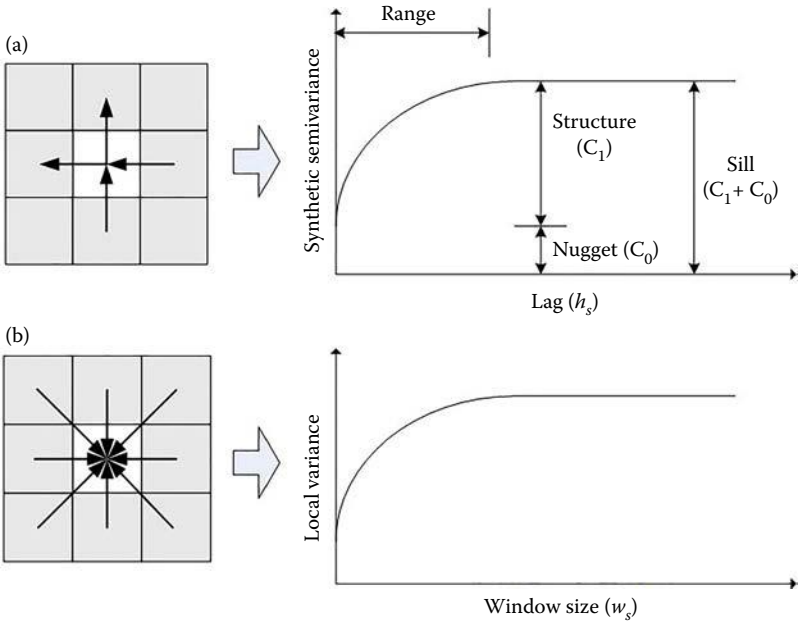


FIGURE 10.20 Schematic diagrams for the computation of (a) semivariance and (b) local variance.

As shown in Figure 10.20, computing the range of spatial correlation is the key foundation of the scale estimation. Ming et al. (2015) proposed that the selection of the optimal scale parameters could be achieved by statistical estimation considering spatial autocorrelation of objects. The optimal scale is essentially the critical point at which the spatial dependence exists or does not exist. The critical point just reflects the meaning of the spatial statistical term *range*. Figure 10.20 illustrates the schematic diagrams for computing semivariance and local variance to get the range. For more details on the computing formulae, please refer to Ming et al. (2012) and Ming et al. (2010). Semivariance (semivariogram) is highly recommended, as it considers the anisotropy of geographic objects.

Ming et al. (2015) used IKONOS, QuickBird, and aerial panchromatic images as the experimental data and employed mean shift segmentation method to verify the effectiveness and feasibility of the proposed method. Experiments based on quantitative multiscale segmentation evaluation testified to the validity of this method. Further, from the view of application, Ming et al. (2016a) combined this method in GEOBIA to extract cropland from a SPOT-5 satellite image. The experimental results indicated that the estimated scale parameters can generate objects with high accuracy and completeness. Ma et al. (2017) analyzed the relationship between MRS and spatial statistical characteristics, and then used the spatial statistical method to estimate the scale parameters in multiscale segmentation. The classification and the accuracy assessment results showed that

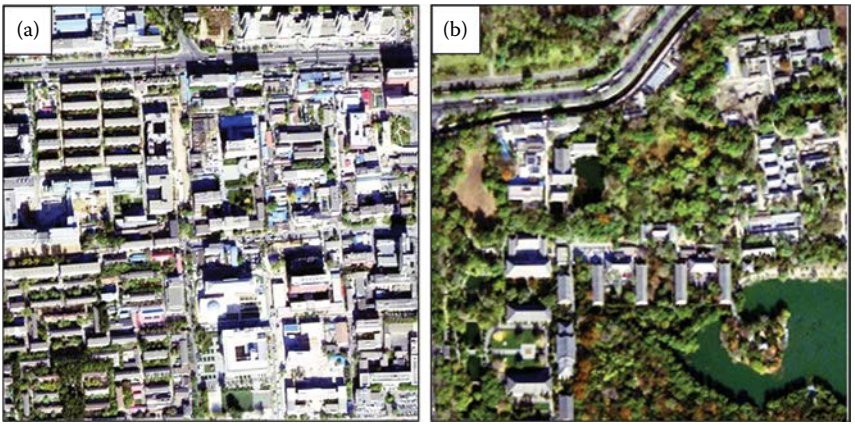


FIGURE 10.21 Two example areas of WorldView-2 image. (a) is mainly composed of buildings, and (b) is composed of vegetation.

the estimated scale is basically closed to the optimal scale in MRS-based multiscale segmentation. In addition, it is an essentially data-driven method that requires almost no prior knowledge; thus, it can enhance the efficiency and automatic degree of GEOBIA. However, this method is more suitable for local detailed information extraction or small areas. When the image area is large, scale estimation based on spatial statistics within the subregion is feasible.

This chapter used this method to extract urban information from the WorldView-2 VHR image. Because the dominant categories within this image are building and vegetation, two subimages shown in Figure 10.21 were cut from the original image and then used as the basis of scale estimation.

Considering that there is little difference between the semivariogram and ALvariogram (Ming et al., 2015), this study computes the semivariances of the experimental image and then determines the spatial domain parameter h_s. Further, attribute domain parameter h_r can be estimated by the proposed method. The essence of h_r is spectral heterogeneity, so it approximates or even equals the meaning of scale parameter (not considering the shape heterogeneity). Figure 10.22 demonstrates the increase of semivariance of these two subimages, and the h_s values for the two subimages are both 7 pixels.

Accordingly, the histograms of the LV image (with the estimated h_r as the window size) are shown in Figure 10.23, by which the estimated spectral parameters are respectively 68 and 53 for the images in Figure 10.21a and b.

When not considering the shape heterogeneity, the meaning of the estimated spectral parameter h_s is close to the term scale parameter in multiresolution segmentation in eCognition software. Compared with Table 10.6, the estimated scale parameter by spatial statistics is 53 for vegetation, which basically approximates the selected optimal scale 50 of vegetation. For buildings, the estimated scale parameter is 68, which somewhat differs from the selected scale in Section 10.3.2.3, i.e., 40. It

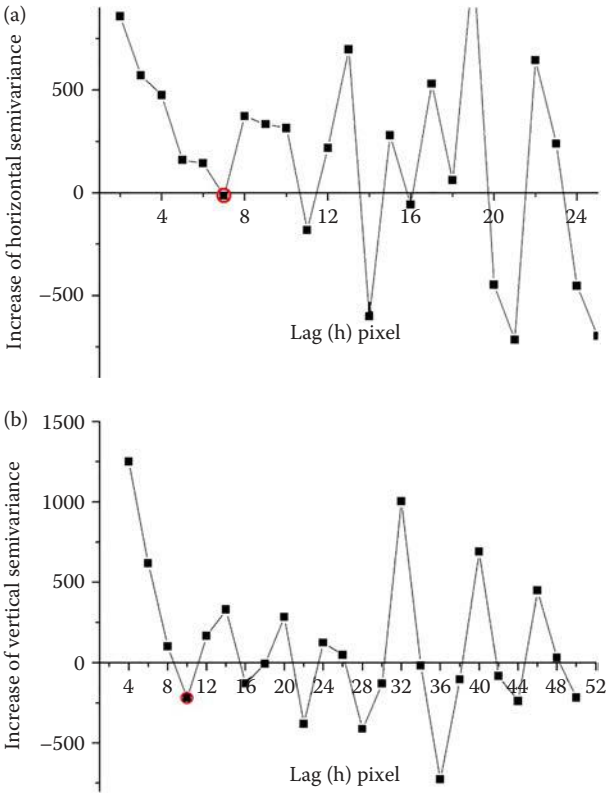


FIGURE 10.22 Semivariance-based spatial parameter estimation results of Figure 10.21a and b.

is because the image in Figure 10.21a is composed of diverse buildings that are very heterogeneous according to their spectral variances; thus, considering only local heterogeneity a large scale of 68 can be estimated. However, the large estimated scale will mix the building objects with their surroundings, while the selected small scale can avoid this issue. Second, the per-category proportions of different selfhood scales are statistically computed based on the number of image objects and they do not consider the size of the image object, which means the weight of the object size is ignored. However, the estimated scale parameter by spatial statistics is computed based on pixels. Third, the selfhood scales are based on samples, however the sample selection is inevitably influenced by personal subjectivity, and whether the selected samples cover all selfhood or whether they could reflect the probability distribution of selfhood scale should be further discussed. However, the proposed scale parameter estimation in this section is based on global spatial statistics and it is a data-driven method without human intervention. Fourth, the estimated scale parameter h_s does not contain the impact of shape heterogeneity, however the selfhood scale method does, which is also inevitable. Accordingly, it

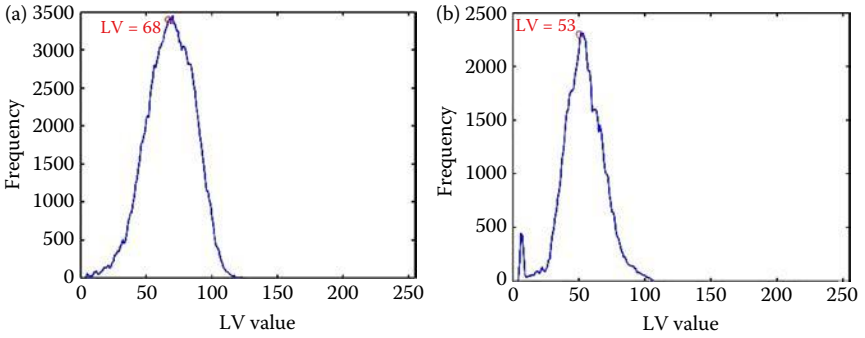


FIGURE 10.23 LV histogram-based spectral parameter estimation results of images in Figure 10.21a and b.

is explainable and acceptable that two different methods based on totally different ideas get different results.

10.5 DISCUSSION

This chapter reviews many scale-selection methods, including supervised and unsupervised ones. In addition, we compared these methods based on the same image data (Figure 10.6). Accordingly, the pros and cons of different scale-selection methods are discussed in this section.

As reported in Table 10.7, unsupervised scale selections including LV and GS need little manual intervention, and thus are much more efficient than supervised

TABLE 10.7
Pros and Cons of Different Scale-Selection Methods

Method	Output	Advantages	Disadvantages
LV and GS	Optimal scale for the whole image	Efficient, considers local heterogeneity, little manual intervention	Ignores category's impacts on scales
Classification-based selection	Optimal scale for each category	Considers object categories	Ignores local heterogeneity of scales, requiring training and testing samples
Boundary-based selection	Optimal scales for the objects with reference boundaries	Avoids three influencing factors on segmentation scales	Requires labor and time costs, mainly deals with the objects with reference boundary
Feature-based selection	Optimal scale for each object or pixel	Considers three influencing factors on scales	Requires a finite number of training samples
Spatial statistics-based estimation	Theoretically optimal scale for the whole image	Theoretically supported, efficient, data driven, without manual intervention	Only suitable for local detailed information extraction or image with small area

ones. They mainly consider local heterogeneity's impact on segmentation scales, but ignore another important factor of the object category. Supervised scale selections based on segmented objects need referenced polygons, so they are difficult to use in practical applications.

On the contrary, supervised selections based on classification accuracies take interesting categories into consideration, but ignore local heterogeneities. These methods need training and testing samples, and can select optimal scales for different categories. However, the impact caused by dynamic samples cannot be avoid, which is the defect of this method. Scale selections based on spatial overlap are able to select the optimal scales for the objects with reference boundaries and avoid three influencing factors. However, spatial-overlap-based scale selections have three limitations: (1) they are weak for dealing with the objects having no reference boundary, (2) the uncertainty on reference boundaries caused by different interpreters will impact selection results, and (3) the delineating of reference boundaries requires much labor and time cost. Finally, feature-based scale selection is proposed, which considers all three influencing factors. First, the surrounding contrast is encoded into high-level object features because each high-level object covers several neighboring objects and its texture features characterize the local contrasts among covered objects. Second, the inherent heterogeneity is measured by the low-level object features because the low-level objects are the inherent components of objects. Accordingly, the differences between the low-level object features can measure the inherent heterogeneity of an object. Third, category determines the visited nodes of the studied pixel in the random forest (Figure 10.17), and further impacts the local Gini importance for diverse scales. The feature-based scale selection needs a finite number of training samples, but it is still more efficient than boundary-based methods.

In addition, the scale estimation based on spatial statistics is theoretically supported by mathematics, spatial statistics, and pattern recognition theories. It is an essentially data-driven method that requires almost no prior knowledge. It can enhance the efficiency and automatic degree of GEOBIA; however, it is only suitable for local detailed information extraction or the image area will be small. In practical application, the subregion should be partitioned before scale selection.

According to the experimental results, we find that different scale-selection methods can produce different optimal scales even for the same data, for example, the scale 60 is selected by the feature- and classification-based methods, while 40 is selected by GS method. This is because these methods consider different aspects of influencing factors to scales and they use different criterions to evaluate segmentation results. Which method should be used depends on our application purpose and the availability of the used data set. For example, if we want to investigate the semantic proportions of land covers, we should use the selection method based on classification accuracy; if we want to map the land covers with accurate boundaries, the spatial-overlap method will be applied; if we want to extract the optimal features to characterize objects, the feature-based method is the best choice; if we just want to have an overall understanding of the land cover objects in the study area, we can use the unsupervised selections; and if we do not even have time to generate multiscale segmentation results, we should use the presegmentation scale estimation method.

10.6 CONCLUSIONS AND FUTURE DIRECTIONS

The scale issue has plagued GEOBIA studies for decades, and thus scale-selection methods are presented to resolve this issue. As discussed in the introduction, there are three influencing factors on segmentation scales, i.e., category, surrounding contrast, and internal heterogeneity, and they should be totally considered while selecting or estimating the optimal scales.

Existing studies on segmentation scales can be divided into two parts: postsegmentation scale selection and presegmentation estimation. This chapter gives a comprehensive review of these techniques, and four conclusions have been drawn:

- Generally, unsupervised selections are much more efficient than supervised ones because unsupervised selections do not need any samples; while supervised methods are more effective to select optimal scales than unsupervised ones because supervised methods use prior information based on human interpretation.
- Among these methods, the selection method based on spatial overlap is the most widely used, but is also the most time consuming; the feature-based method is the most effective for selecting the optimal scales for objects.
- A recent scale-estimation method based on spatial statistics was proposed. It is a fully data-driven method, and needs little manual intervention, but it is not suitable for images with complex landscapes or with large extents.
- Every scale-selection or scale-estimation method has unique advantages and limitations (Table 10.7), and which method to use mainly depends on our application purposes.

Despite the good performances of scale-selection methods, several issues remain to be studied further. First, all postsegmentation scale selections essentially aim to choose the optimal scale level from multiscale segmentations, but the multiscales are set manually; thus, there are two issues. How many scales will be employed? What is the interval between scales? Accordingly, an automatic method is required to resolve the two issues. Second, different data sources, study areas, and methods used may lead to different selected scales. It is hard to determine whether the geographic or data difference results in the scale differences. Third, the parameters used in scale selection will impact the selection results; thus, appropriate parameters should be chosen for improving the application accuracies of selected scales. Fourth, the scale-selection results (Table 10.7) are totally generated based on multiscale segmentation results, and thus they are discrete, but they should be continuous because the scales of geographic objects change continuously. Accordingly, a learning mechanism of continuous scales instead of the scale-selection method should be developed in the future.

ACKNOWLEDGMENTS

The work presented in this paper is supported by the National Natural Science Foundation of China (Nos. 41471315, 41371347, and 41671369). The comments from the anonymous reviewer are greatly appreciated.

REFERENCES

Albrecht, F. 2010. Uncertainty in image interpretation as reference for accuracy assessment in object-based image analysis. *Accuracy 2010. Proceedings of the Ninth International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences*, pp.13–16.

Baatz, M., & Schäpe, A. 2000. Multiresolution segmentation: An optimization approach for high quality multi-scale image segmentation. *Angewandte Geographische Informationsverarbeitung XII*, pp. 12–23.

Blaschke, T. 2010. Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65, 2–16.

Bock, M. et al. 2005. Spatial indicators for nature conservation from European to local scale. *Ecological Indicators*, 5(4), 322–338.

Breiman, L. 2001. Random forests. *Machine Learning*, 45, 5–32.

Cleve, C., Kelly, M., Kearns, F. R., & Moritz, M. 2008. Classification of the wildland–urban interface: A comparison of pixel- and object-based classifications using high-resolution aerial photography. *Computers, Environment and Urban Systems*, 32, 317–326.

Clinton, N., Holt, A., Scarborough, J., Yan, L., & Gong, P. 2010. Accuracy assessment measures for object-based image segmentation goodness. *Photogrammetric Engineering & Remote Sensing*, 76(3), 289–299.

Drăguţ, L., Csillik, O., Eisank, C., & Tiede, D. 2014. Automated parameterisation for multi-scale image segmentation on multiple layers. *ISPRS Journal of Photogrammetry and Remote Sensing*, 88, 119–127.

Drăguţ, L., Tiede, D., & Levick, S. R. 2010. ESP: A tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. *International Journal of Geographical Information Science*, 24(6), 859–871.

Espindola, G. M., Câmara, G., Reis, I. A., Bins, L. S., & Monteiro, A. M. 2006. Parameter selection for region-growing image segmentation algorithms using spatial autocorrelation. *International Journal of Remote Sensing*, 27(14), 3035–3040.

Estrada, F. J., & Jepson, A. D. 2009. Benchmarking image segmentation algorithms. *International Journal of Computer Vision*, 85(2), 167–181.

Fotheringham, A. S., Brunsdon, C., & Charlton, M. 2000. *Quantitative Geography: Perspectives on Spatial Data Analysis*. Sage.

Friedl, M. A., & Brodley, C. E. 1997. Decision tree classification of land cover from remotely sensed data. *Remote Sensing of Environment*, 61, 399–409.

Guo, Z., & Du, S. 2016. Mining parameter information for building extraction and change detection with very high-resolution imagery and GIS data. *GIScience & Remote Sensing*, 54(1), 38–63.

Hay, G. J., & Castilla, G. 2008. Geographic Object-Based Image Analysis (GEOBIA): A new name for a new discipline. *Object-Based Image Analysis*, 75–89. Berlin: Springer.

Hu, J., Ge, Y., Chen, Y., & Li, D. 2015. Super-resolution land cover mapping based on multiscale spatial regularization. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(5), 2031–2039.

Huang, H., Wu, B., & Fang, J. 2003. Analysis to the relationship of classification accuracy, segmentation scale, image resolution. *Geoscience and Remote Sensing Symposium*, 2003, pp. 3671–3673.

Im, J., Lu, Z., Rhee, J., & Quackenbush, L. J. 2012. Impervious surface quantification using a synthesis of artificial immune networks and decision/regression trees from multi-sensor data. *Remote Sensing of Environment*, 117, 102–113.

Johnson, B., & Xie, Z. 2011. Unsupervised image segmentation evaluation and refinement using a multi-scale approach. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(4), 473–483.

Kim, M., Madden, M., & Warner, T. 2008. Estimation of optimal image object size for the segmentation of forest stands with multispectral IKONOS imagery. *Object-Based Image Analysis*, 291–307. Berlin: Springer.

Laliberte, A. S., Rango, A., Havstad, K. M., Paris, J. F., Beck, R. F., McNeely, R., & Gonzalez, A. L. 2004. Object-oriented image analysis for mapping shrub encroachment from 1937 to 2003 in southern New Mexico. *Remote Sensing of Environment*, 93, 198–210.

Ma, Y., Ming, D., & Yang, H. 2017. Scale estimation of object oriented image analysis based on spectral spatial statistics: A fractal network evolution segmentation algorithm for example. *Journal of Remote Sensing*, 21, 566–578.

Mallinis, G., Koutsias, N., Tsakiri-Strati, M., & Karteris, M. 2008. Object-based classification using Quickbird imagery for delineating forest vegetation polygons in a Mediterranean test site. *ISPRS Journal of Photogrammetry and Remote Sensing*, 63, 237–250.

Meinel, G., & Neubert, M. 2004. A comparison of segmentation programs for high resolution remote sensing data. *International Archives of Photogrammetry and Remote Sensing*, 35, 1097–1105.

Ming, D., Ci, T., Cai, H., Li, L., Qiao, C., & Du, J. 2012. Semivariogram based spatial bandwidth selection for remote sensing image segmentation with mean-shift algorithm. *IEEE Geoscience and Remote Sensing Letters*, 9, 813–817.

Ming, D., Li, J., Wang, J., & Zhang, M. 2015. Scale parameter selection by spatial statistics for GeOBIA: Using mean-shift based multi-scale segmentation as an example. *ISPRS Journal of Photogrammetry and Remote Sensing*, 106, 28–41.

Ming, D., Luo, J., Li, L., & Song, Z. 2010. Modified local variance based method for selecting the optimal spatial resolution of remote sensing image. *Proceeding of the 18th International Conference on Geoinformatics*, IEEE.

Ming, D., Zhang, X., Wang, M., & Zhou, W. 2016a. Cropland extraction based on OBIA and adaptive scale pre-estimation. *Photogrammetric Engineering & Remote Sensing*, 82, 635–644.

Ming, D., Zhou, W., & Wang, M. 2016b. Scale parameter estimation based on the spatial and spectral statistics in high spatial resolution image segmentation. *Journal of Geo-Information Science*, 18(5), 622–631 (in Chinese).

Myint, S. W., Gober, P., Brazel, A., Grossman-Clarke, S., & Weng, Q. 2011. Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. *Remote Sensing of Environment*, 115, 1145–1161.

Pal, M. 2005. Random forest classifier for remote sensing classification. *International Journal of Remote Sensing*, 26, 217–222.

Peña-Barragán, J. M., Ngugi, M. K., Plant, R. E., & Six, J. 2011. Object-based crop identification using multiple vegetation indices, textural features and crop phenology. *Remote Sensing of Environment*, 115(6), 1301–1316.

Powers, D. M. W. 2011. Evaluation: From precision, recall and F-factor to ROC, informedness, markedness & correlation. *Journal of Machine Learning Technologies*, 2, 2229–3981.

Powers, R. P., Hay, G. J., & Chen, G. 2012. How wetland type and area differ through scale: A GEOBIA case study in Alberta's Boreal Plains. *Remote Sensing of Environment*, 117, 135–145.

Ramakant, N., & Ramesh, K .B. 1980. Linear feature extraction and description. *Computer Graphics and Image Processing*, 13, 257–269.

Smith, A. 2010. Image segmentation scale parameter optimization and land cover classification using the Random Forest algorithm. *Journal of Spatial Science*, 55, 69–79.

Stow, D., Hamada, Y., Coulter, L., & Anguelova, Z. 2008. Monitoring shrubland habitat changes through object-based change identification with airborne multispectral imagery. *Remote Sensing of Environment*, 112, 1051–1061.

Trimble. 2011. *eCognition Developer 8.7 Reference Book*. Munich Germany.

Witharana, C., & Civco, D. L. 2014. Optimizing multi-resolution segmentation scale using empirical methods: Exploring the sensitivity of the supervised discrepancy measure euclidean distance 2 (ED2). *ISPRS Journal of Photogrammetry and Remote Sensing*, 87(1), 108–121.

Woodcock, C. E., & Strahler, A. H. 1987. The factor of scale in remote sensing. *Remote Sensing of Environment*, 21(3), 311–332.

Yu, Q., Gong, P., Clinton, N., Biging, G., Kelly, M., & Schirokauer, D. 2006. Object-based detailed vegetation classification with airborne high spatial resolution remote sensing imagery. *Photogrammetric Engineering & Remote Sensing*, 72(7), 799–811.

Zhang, H., Fritts, J. E., & Goldman, S. A. 2008. Image segmentation evaluation: A survey of unsupervised methods. *Computer Vision and Image Understanding*, 110(2), 260–280.

Zhang, X., & Du, S. 2016. Learning selfhood scales for urban land cover mapping with very-high-resolution VHR images. *Remote Sensing of Environment*, 178, 172–190.

Zhang, X., Feng, X., Xiao, P., He, G., & Zhu, L. 2015a. Segmentation quality evaluation using region-based precision and recall measures for remote sensing images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 102, 73–84.

Zhang, X., Xiao, P., Feng, X., & Feng, L. 2015b. Toward evaluating multiscale segmentations of high spatial resolution remote sensing images. *IEEE Transactions on Geoscience & Remote Sensing*, 53(7), 1–13.

11 Computer Vision Methodologies for Automated Processing of Camera Trap Data
A Technological Review

Joshua Seltzer, Michael Guerzhoy, and Monika Havelka

CONTENTS

11.1 INTRODUCTION TO ANIMAL POPULATION MONITORING

Beneath the dense canopies of the Amazon rainforest, traditional remote sensing technologies such as aerial photography fail to capture the abundance of life. In lieu of such techniques, professional and amateur enthusiasts alike have begun using camera trap systems to monitor populations, which use hidden and nonintrusive cameras (triggered by motion and/or heat) placed strategically around natural parks and research stations. Camera traps provide unbiased and nonartificial glances into the lives of wild animals, which is highly conducive to ethological and populational studies. While camera trap technology has advanced significantly in recent years, methods of managing and analyzing the collected data have been largely bottlenecked by existing practices. The quality and quantity of photos and metadata have increased

dramatically with the advent of digital cameras—as well as advances in battery life—but the same advances have not occurred in managing and analyzing data from that rich source (Fegraus et al., 2011). Age-old methodologies have undergone only incremental improvements, rather than the revolutionary advances that so often drive progress.

Although some theoretical groundwork has been laid in the area of modeling the dynamics of animal populations, little work has focused on the integration of such models with the rapidly growing field of computer vision in order to automate the process. In particular, this chapter will examine population density models and computer vision techniques in order to evaluate existing and potential performance in the context of camera traps. Successful integration between camera traps, computer vision, and population modeling could be tested easily for accuracy against existing data sets (which have been compiled manually through observational data); and if proven to work could have lasting ramifications for research and methodology. Ultimately, such a system could provide insight into many ongoing questions in conservation biology. In addition to revealing ethological insights into the behavior of endangered species, automatic recognition of individuals and species for the purpose of estimating population densities could provide both long-term and real-time monitoring of threatened groups. Such information is invaluable to many fields of biology and conservation policies, and an effective software solution for automating recognition processes could be used to globally analyze animal populations, given the growing use of camera trap systems. In an age of fragmenting ecosystems, ever-increasing insularization, and rapid extinction, it is all the more important to act with haste.

11.2 CURRENT PRACTICES

Although there is great diversity of camera trap setups, situated in a wide range of habitats, the methodology for analyzing collected data is fairly consistent across each research station. While there is a range, to some extent, in the sophistication of their data management—such as the software used for organizing and sorting the photos—central to each system is the need for a human to manually tag each photo with the type of species present and other relevant data (Sundaresan et al., 2011). Even descriptions of state-of-the-art systems, e.g., Zaragozí et al. (2015), emphasize a need for modernizing data management, yet still focus on optimizing the speed at which humans can manually classify (label) the photos. The data may be processed by a small group of experts, or crowdsourced from a larger pool, but in either case it presents a significant financial cost and is very time consuming.

At Tiputini Biodiversity Station (TBS), a remote field research center in the Yasuní Biosphere Reserve of the Ecuadorian Amazon, camera traps have been an ongoing project since the days of film cameras. The surrounding park is a hot spot of biodiversity. Over the last 10 years, digital cameras have enabled the capture of more than 100,000 photos and videos of around 70 wild species, some of which are rare or endangered (D. Mosquera, personal communication, June 2016). Labeling such a massive collection is immensely laborious; and as photo capacity and sensor sensitivity progress even further, the data will only become less manageable. For some

species, where individual markings are clearer and consistent, members of the species can be individually identified, and thus population estimates for a certain region can be directly observed. However, for quite a few species this is not a viable technique: many animals lack persistent or prominent individual markings, and thus cannot be readily individuated. Such limitations prevent researchers from reliably identifying whether an individual is a repeat visitor (a recapture); however, identification at the species level is not precluded. In order to better extrapolate estimates of population density from the photos that they capture for these animals, the research team at TBS has begun using a recently developed gas particle model.

11.3 POPULATION DENSITY MODELING

In an effort to relax the complex reality of animal interactions, Rowcliffe et al. (2008) developed a two-dimensional ideal gas model to represent contact between individual animals. This method requires no individual recognition because it extrapolates from the number of times any member of a species is observed. The model can be used to estimate the number of particles in an area based on the number of collisions in which a single particle takes part. It assumes that the particles collide elastically and without friction. By modeling individual animals as particles, various properties of the population can be estimated using the model. In this particular application, the camera itself is also regarded as a particle, which is said to contact particle(s) whenever it sees an animal. The total number of particles in a given area, based on the model's assumptions, can then be computed as a function of several variables: total area, animal velocity, observed area of the camera, the number of contacts that it has with other particles, and time elapsed—all of which can be readily determined with data from camera traps. The density can therefore be computed by dividing the total area on either side, since density is the ratio between population and area.

While Rowcliffe et al. (2008) described the comparison of roaming animal population behaviors to ideal gas particles as unrealistic, the model nonetheless provides a robust estimation of population densities. By using camera trap data from Whipsnade Wild Animal Park in England (with less than 1000 photographs) and comparing the computed estimates with census data from the same area, Rowcliffe et al. performed an analysis on four mammal species. In three out of four of the examined species, they found that the model's estimates were within 22% of the census results, although the estimate for one species was 86% smaller than what was observed in the census. It is also worth noting that their estimates were distributed evenly on either side of the figures reported in the census.

Several other models have become popular in studies without the option of individually identifying species members, the most popular perhaps being the maximum entropy (Maxent) method. As Merow et al. (2013) explain, the model first requires researchers to define a landscape, breaking it into grid cells. Using presence-only data as input, which simply marks the locations at which a species has been seen, the Maxent algorithm attempts to generalize to background locations, which are regions of the landscape where the species' presence is unknown. Maxent outputs a relative occurrence rate (ROR), which marks the probability of an individual originating from any given cell. However, since this analysis depends

on the assumption that the data compiled from camera traps are representative of the whole landscape (i.e., a random sample)—a faulty assumption in most cases—additional data are used to correct the bias. For example, a feature will be selected that ideally varies with animal density, such as precipitation or temperature in a given area. Other factors include search effort in each cell within the landscape, allowing the sampling bias to be mitigated.

In order to test the efficacy of population models on small sample sizes, Pearson et al. (2006) adopted a jackknife statistical method in lieu of independent census data. This is also known as *leave one out*: for each of the n observed localities, they tested the models' predicted population density based on the remaining $n - 1$ localities. By averaging out the estimated versus observed populations, they were able to assess the performance of the Maxent against the Genetic Algorithm for Rule-Set Prediction (GARP)*. In particular, they assessed the models on presence-only data for an elusive species of nocturnal leaf-tailed geckos in Madagascar. Whereas GARP failed to accurately predict populations with very small (less than 10) localities, the Maxent model performed significantly better, reaching high success rates with as little as five measured localities. These results suggest that Maxent, if properly configured, may be a viable solution for modeling populations of endangered and elusive species, where sample size is a particularly troublesome problem.

By utilizing the gas particle model, the researchers at TBS merely have to identify the species present in their photographs, rather than distinguishing between individuals within a species in order to estimate the total population. However, the process of identifying animals at the species level is still time consuming. Because of the sheer mass of photographs being produced there and elsewhere, automating this process would provide the means of carrying out large-scale analyses without requiring the bottleneck of human input. Likewise, an automated approach to the Maxent model would simply require the presence locations of each species, which would be calculated (after species identification) from a camera's global positioning system (GPS) data, and some auxiliary data to mitigate sampling bias. In order to examine the potential of automating this process, the next sections will look at existing computer vision technologies and the extent to which they can be adapted for recognition of both species and individuals.

11.4 CONVOLUTIONAL NEURAL NETWORKS†

A feedforward neural network is a set of interconnected nodes (inspired by neural connections in the brain), with directed weighted connections. A feedforward neural network can be used for probabilistic categorization of inputs. The design of neural networks has sometimes been inspired by cognitive science. Computational methods, notably stochastic gradient descent, exist to find weights for a feedforward neural network such that the network is able classify instances in the training set. This

* GARP uses a process inspired by natural selection to develop conditional rules that correlate with factors indicative of population distributions.

† Much of the material and terminology in this section is informed by LeCun et al. (2015)'s extensive review of the subject.

process is referred to as training the neural network. A training set contains input (often images) with matching labels, or categories. When training, based on the network's success rate in guessing the correct label, its connections are modified to better optimize its accuracy. Neural networks learn by modifying connections throughout between the nodes in the network.

Convolutional neural networks (CNNs or ConvNets) are a category of feedforward neural networks. The connection patterns between their nodes have partly been inspired by the connection in the visual cortex of the mammalian brain. The connections between the nodes in CNNs are arranged in such a way as to allow detection of the same pattern in the entire image; furthermore, the connections are such that they allow for the CNN to model patterns at different levels of abstraction. This allows for the creation of representations of visual patterns, with layers of nodes that are further away from the nodes containing the input image encoding more general and abstract patterns. CNNs are able to represent and recognize very complex patterns as a result. The ability of CNNs to represent complex patterns leads to high performance in image classification tasks. For example, CNNs have reached more than 90% accuracy in high-quality photographs of fish, butterflies, and plants (Hernández-Serna & Jiménez-Segura, 2014), using ∼10,000 training images of 740 distinct species. Many CNNs can process hundreds of images each second, even on a modest laptop. They need not operate on any particular taxonomical level: that is, depending on the training sets used and ecological factors, CNNs may be trained to identify individual members within a species just as readily as they would identify one species from another. This elasticity is powerful because the ability to distinguish visual features stems directly from the features that humans have already observed. CNNs can therefore learn to distinguish either species or individual members, depending on the specificity of labels in the training sets; or any other property associated with visual changes. Adapting these systems to identify wild animals from camera traps is quite straightforward: by using existing, labeled camera trap photos as training sets (as well as online databases compiled for this purpose) and existing networks with deep representational power, a CNN can be trained to achieve adequately high performance in this area.

11.4.1 State-of-the-Art Performance

A network's performance classifying a labeled data set (i.e., the proportion of images correctly classified) is a popular type of metric for benchmarking a CNN's capability, generally using a test set—images not in the training set that the network tries to predict. One such metric is the annual ImageNet Large Scale Visual Recognition Challenge (ILSVRC), which uses a subset of the ImageNet (Deng et al., 2009) data set to analyze performance. In an extensive report on their challenge, Russakovsky et al. (2015) compared state-of-the-art object classification systems with human ability, using a Top-5 classification error, which defines an error as failing to guess the correct category within five attempts. The organizers trained two human annotators, A1 and A2, on an image-labeling task using hundreds of examples before testing each annotator on 1500 and 258 images, respectively. A1 achieved a Top-5 classification error rate of approximately 5.1%; whereas A2 performed much worse, with an error

rate of 12%. A1's error rate translates to a Top-5 classification accuracy of ∼95%, meaning a correct guess within five attempts for 95% of the test cases. One of the top machine classifiers in ILSVRC 2014, GoogLeNet (Szegedy et al., 2015), was tested on the same test set as A1, and achieved a Top-5 error rate of ∼6.8% (∼93% accuracy). Of the images in the test sets, 204 overlapped between A1 and A2; and, since the annotators did not show a strong overlap in their guesses, the researchers estimated an optimal human error rate of 2.4% by combining their answers. Various networks have since surpassed A1's performance, including He et al. (2015)'s CNN, which achieved a 4.94% Top-5 classification error on ImageNet, using ∼1.2 million training images and 100,000 images to test. Although this may still be far from the estimate of optimistic human accuracy, it is clear that machines have proven equally as effective in this task as the average human.

VGG16, another top competitor in ILSVRC 2014, is a "very deep" state-of-the-art CNN designed by Simonyan and Zisserman (2014). VGG16's base performance on several groups of taxonomic categories in the ILSVRC 2014 test set is shown in Figure 11.1. The test set contains 1000 unique categories, 398 of which classify some biological group (the others are assorted objects), each represented by 50 images of relatively high quality. For some groupings, such as birds (of which there are 37 categories), VGG16 correctly predicted the category in its first guess for more than 80% of test cases, and in nearly 95% cases classified it correctly in its Top-5 predictions. As shown in Figure 11.2, the network's misclassifications were often of genetically and visually similar animals, and were thus errors a human might make as well.

Figure 11.1 also shows performance when the test set is restricted based on the confidence rating. The confidence rating for a specified image and group of categories is, intuitively, the network's confidence that the image should be classified under any of the categories in the group. Restricting the test cases to instances where the confidence exceeds a certain threshold yields considerable improvements in classification performance. For example, by ignoring photos of marine animals for which VGG16 shows less than 50% confidence, and analyzing only the remaining 85% of the photos, the system sees an increase in performance of more than 10%. Such a restriction could be used to quickly filter out ambiguous images that will likely be misclassified, and that require expertise to classify with confidence.

11.4.2 Integrating with Ecological Paradigms

Because CNNs (and neural networks in general) are an emerging technology, very few attempts have been made to integrate them with field research. Gómez et al. (2016) reportedly found only two previous attempts. In their study they use the Snapshot Serengeti* data set, which provides more than a million labeled camera trap photos for training and testing sets, combined with fine-tuned off-the-shelf features, a special class of networks which use the feature recognition configurations

* Swanson et al. (2015) describe the Snapshot Serengeti project in depth. The project involves hundreds of cameras operating since 2010, with more than a million photos as of 2013. A citizen-science website, www.snapshotserengeti.org, has collected more than 10 million labels from online users.

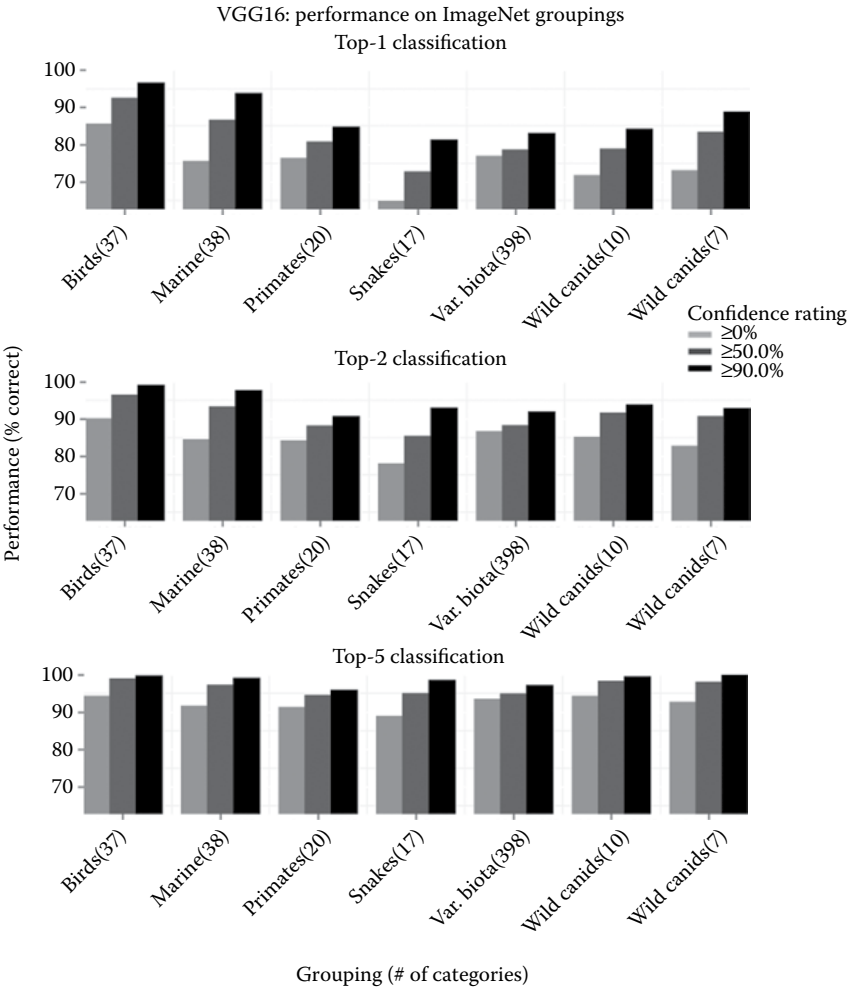


FIGURE 11.1 Top-1, Top-2, and Top-5 classification accuracy for the network VGG16 on the ILSVRC 2014 test set for images within a variety of groupings. Performance is also listed for subset(s) of each grouping's test set, where the network's confidence exceeds a specified confidence (0%, 50%, and 99%). The complete list of taxonomic classifications per grouping is listed, along with the percentage of images retained at each confidence threshold. (From Simonyan, K., & Zisserman, A. 2014. http://www.robots.ox.ac.uk/~vgg/research/very_deep/ (August 22, 2016); Deng, J. et al. 2009. *2009 IEEE Conference on Computer Vision and Pattern Recognition*.)

of previously trained CNNs (e.g., AlexNet or GoogLeNet) that already have rich representational power, with some training performed on them in order to acclimate them to the Snapshot Serengeti data set. Doing this, they achieved an approximately 90% success rate in distinguishing among 26 different species (using a training set with hundreds of thousands of images)—not entirely satisfactory, but quite incredible given the novelty of the techniques for this specific application. Gómez et al.'s work is

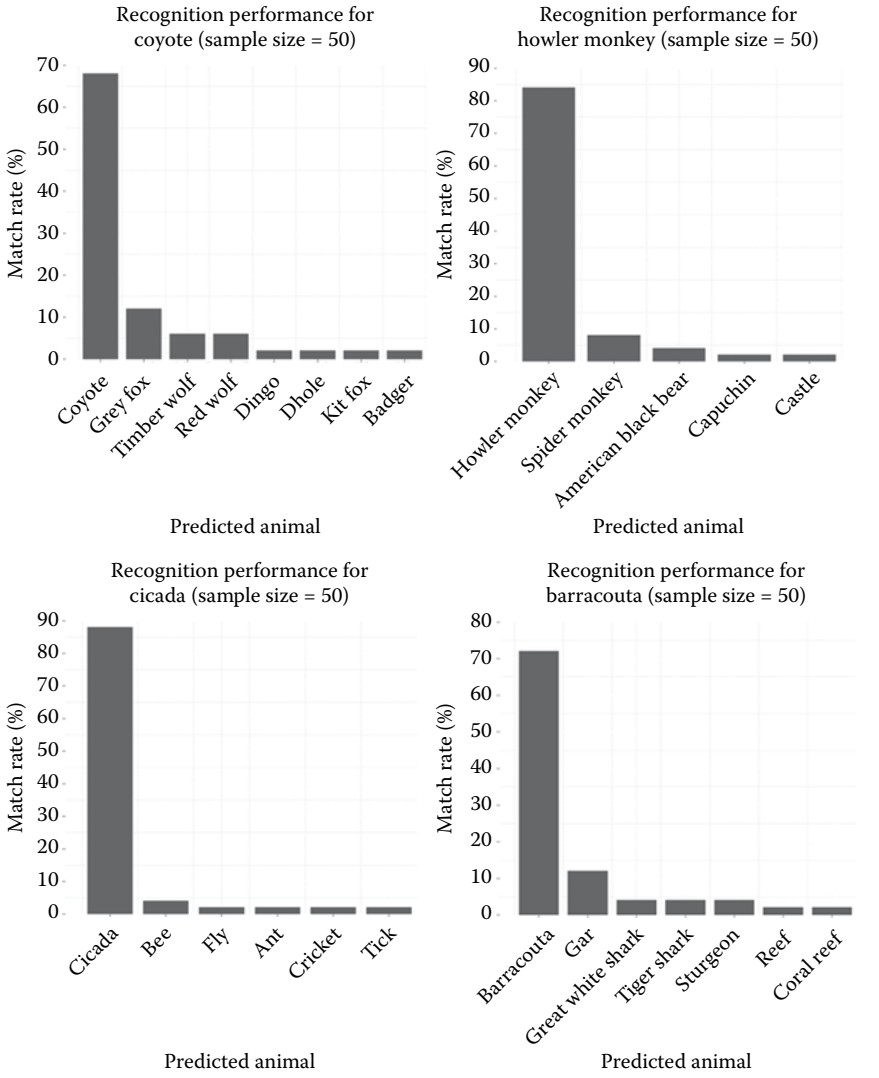


FIGURE 11.2 Classification rate by VGG16 for various animal categories in ImageNet, showing both success rate and rate of each misclassification. The sample size for each category is 50 images. (From Simonyan, K., & Zisserman, A. 2014. http://www.robots.ox.ac.uk/∼vgg/research/very_deep/ (August 22, 2016); Deng, J. et al. 2009. *2009 IEEE Conference on Computer Vision and Pattern Recognition*.)

also an improvement upon prior attempts they had found, indicating a large increase in performance in just a couple of years. As CNNs continue to rapidly improve, even greater gains will likely be seen.

One concerning issue with this methodology is that endangered and elusive species will be underrepresented in the data sets, and thus the networks will be less

tuned to recognize their features. This is known as one-shot or few-shot learning, wherein the target categories (which are to be matched with labels) are represented by sparse data—or even zero-shot learning, where only a description of a category, but no example of images from the category, is available. This might often be the case in ecological studies, where species lacking any preexisting photographic data might be uniquely recognized by a set of features a network can recognize. One attempt to eliminate biases in the data set by Hoffman et al. (2014) focused on adapting the CNN's internal architecture to better identify categories with sparse representation, of which endangered species are a prime example. They were able to increase performance on a subset of ImageNet from 66% accuracy to 77% using these techniques. Evidently, while identifying underrepresented species is a significant challenge, further research in network architectures could mitigate the problem.

Finally, there are some consequences that follow from the design of CNNs. There is no possibility for accidental errors as with humans: it is impossible for a neural network to mix up two photos it is analyzing because the process is entirely sequential and independent. Misclicking and mixing up photos is not a possibility. In all likelihood, poor performance on the edge cases, which truly do require human ingenuity to label, might be offset by a computer's lack of such trivial errors. That said, CNNs are susceptible to the same analytical mistakes that humans make. Just like the human propensity to see faces and patterns in random stimuli—known as pareidolia—neural networks are prone to the same hypersensitivity. Nguyen et al. (2015) report false positive predictions on the aforementioned AlexNet system using randomly perturbed images to produce what they call *fooling images*. These are pictures that are entirely unrecognizable to the human eye, but that the network nonetheless labels with more than 99% confidence. Clearly, these systems are far from perfect; but because the fooling images were purposefully contrived, this may not have a significant impact in the context of camera trap photos.

11.5 FACIAL AND INDIVIDUAL RECOGNITION

Face recognition is an important benchmark for prospecting the possible performance of other recognition systems, since such systems are the forefront of numerous prominent and emerging technologies—forensics, social media, security, and virtual reality, to name but a few—and thus much of the literature and state-of-the-art technical developments are based on facial modeling and matching. Although CNNs are now used extensively in face recognition, there are still a number of strategies from the field that may prove useful in identifying individuals. Moreover, the heuristics in use—which have been designed specifically for facial features—can give insight into the potential performance of modeling systems on animals, should the same tools be developed in that area. Tsao and Livingstone (2008) predict that "face perception is in many ways a microcosm of object recognition; and the solution to the particular problem of face recognition will undoubtedly yield insights into the general problem of object recognition." This section will overview three challenges and associated strategies in facial recognition that might be applied in the context of animal species recognition.

Unconstrained face recognition—the problem of matching faces to a database without any control over the lighting, colouration, or angle/extent at which each face is exposed—is a major open problem in computer vision. With major progress at the task of face recognition with monotonic backgrounds with normalized poses, the prevailing goal now is to recognize faces from any context or camera. Best-Rowden et al. (2014) call the subjects "uncooperative"—surely an apt descriptor of the wild animals roaming natural parks. In their study, focusing on the viability of unconstrained face recognition for forensic purposes, they use low-quality candid images and video. In order to analyze this media, they employ a "state-of-the-art face matcher and a separate face modeling SDK [development kit]." The latter is capable of generating pose-corrected faces, which are extrapolations of the original image recreated in a frontal view. In this particular study, Best-Rowden et al. probed several candid photo databases, using 13,000 images and more than 3000 videos. They achieved recognition scores of higher than 95% when the input media included multiple images and video tracks. While much lower error rates have been achieved for normalized data, this marks an impressive feat in recognizing entirely unconstrained input sources. However, Kemelmacher-Shlizerman et al. (2016) report that such promising results may not generalize to scale: in a data set with a million unique individual faces, state-of-the-art facial recognition algorithms degrade proportionally with the number of distractors (faces of individuals not present in the test set).

This capacity to artificially normalize an input is immensely facilitative to recognition. For example, even the fusiform face area, the neural region in the brain evolutionarily honed to represent faces, is famously prone to misidentifying upside-down faces (Tsao and Livingstone, 2008). This is the result of an infamous problem in cognitive science: the lack of invariance problem. The brain's visual processing systems seamlessly interpret complex patterns, objects, places, and people for us despite immense variation: poor lighting, distorted shapes, an odd angle or visual obstacles, among other things, result in an immense amount of misleading or missing visual information. Nonetheless, the brain can instantly associate a face with a person by comparing the face to its mental representation. Since faces are as variant as objects and animals, the elusive answer to this puzzle could be a breakthrough in representing and recognizing arbitrarily complex visual patterns that persist across many variations—exactly the obstacle preventing automated camera trap recognition from blooming. Thus high performance in state-of-the-art facial recognition on unconstrained sources is applicable as a benchmark for the camera trap paradigm, since the same concepts will eventually transfer to individual recognition of animals. For this to become a reality, however, much research and implementation is needed in developing the same heuristic modeling and matching algorithms as have been developed specifically for faces.

As in animal classification, underrepresentation in the data can be a major problem for recognizing individuals. One technique for one-shot learning in individual recognition, developed by Masi et al. (2016) for CNNs, could help to mitigate the bias. Starting from a relatively small set of training images, they created several copies of each with morphological variations in order to simulate a diversity in their input, in effect creating new training cases where there had been only one before. These mutations included pose variation, 3D shape variation, and expression

variation. For example, pose variation involved the re-creation of new viewpoints in order to create several versions of a human face from a single image—similar to the pose-corrected faces used by Best-Rowden et al. (2014). This artificial variation was effective, allowing the authors' neural network to perform as well as more sophisticated networks trained on up to a thousand more images. Although the same transformations may not apply to detection of distinguishing marks, the underlying concept transfers to camera trap photos. By creating transformations that sample from any possible variability in the photos (e.g., in lighting, angle, orientation), a CNN may be more effectively trained to differentiate species.

11.6 DISCUSSION

Having examined two models of population distribution that are currently in use for predicting population densities from camera trap data; and having looked at state-of-the-art species recognition (with CNNs) and facial recognition, the question remains: are existing computer vision technologies, coupled with these population models, sufficiently accurate to be used (or adapted) as part of an automated population density estimation system?

11.6.1 Ecological Considerations

First, it is worth reviewing the many ways in which the environment under consideration can alter the applicability of automated recognition and population models. It has been seen that the performance of VGG16 on the ImageNet data set varies significantly across categorical groupings of animals—and there are many reasons for this, because it reflects a huge variability in animal appearance. Perhaps the most apparent variable is the simplest: biological groups vary widely in the extent of their diversity. Another factor concerns scale—microscopic environments of course differ greatly from macroscopic environments, and a system that has learned to recognize patterns in the latter may experience great difficulty in analyzing the former, or in distinguishing the size of an individual. Thus the extent of physical differences across categories must be considered in order to evaluate the applicability of machine learning systems.

There is at least one benefit in employing CNNs for individual recognition in certain situations. The dynamic nature of CNN learning marks an abrupt departure from other feature recognition systems: for example, their predictive skills are able to adapt to changing frameworks of knowledge. Transient markings (visual characteristics that are temporary or seasonal) have made individual recognition impossible in some species: "until the technology exists that can keep up with change… you have to add human decision into the mix" (M. Povarova, personal communication, June 2016). However, their ability to reform previous connections make CNNs a suitable candidate for this task. Traditional face recognition techniques, on the other hand, apply best to scenarios in which transient markings are minimal since they perform best when there are distinct and recognizable features. Additionally, zero-shot learning and other aforementioned techniques can be used to reduce selection bias by allowing individuals with little or no existing photographic representation to still be identified.

There are also spatial modeling considerations to be made in applying the Maxent and gas particle models to animal populations. Because the gas particle model assumes a 2D distribution of animals, it is at best a weak approximation of real ecological environments. Generalizing the model to three dimensions and accounting for the biases of a species' patterns of movement could improve how well the model approximates the reality. Maxent models similarly suffer from a lack of knowledge about the distributional patterns of spatial cells—and these need to be addressed in order for the model to accurately represent the state of affairs. Adapting such a model to an environment would inevitably include an analysis of which cells in the landscape share similar properties, and to what extent. Territoriality among a group of animals may further confound these models since a group's distribution might fluctuate seasonally, and therefore must also be considered when applying the models to a study.

11.6.2 Conclusion

The success of object recognition and classification technology in the domains of face and generic-object recognition suggests that the technology could be applied to ecological studies: perhaps with enough work spent developing heuristics for modeling and matching images, the same successes seen in facial recognition can be transferred to camera traps. Due to the nature of the lack of invariance problem, wherein a recognition software must be able to encode core features of a visual object that are subject to modification and mutilation, the same underlying concepts are directly applicable to almost any area of computer vision. These recognition systems are therefore a likely approximation for individual animal recognition, and at the very least can provide a benchmark for potential performance. In addition, the dynamic learning ability of CNNs makes it a good candidate for individual recognition because transient markings would be considered one part of the network's constantly adapting knowledge base.

Since both the gas particle model and Maxent only require identification of species type for each captured photo, and the latter has effective results even in very small sample sizes, the only component missing for a fully automated system is the recognition of species. As shown here, convolutional neural networks have reached quite accurate species recognition scores, not only for high-quality photographs but even within the camera trap paradigm. Although there are some shortcomings in this approach, including an underrepresentation of endangered species and a predisposition toward false positive recognition, these problems are not irreparable: pose variations and other augmentations can be used to mitigate the underrepresentation bias. Finally, CNNs lack the accidental mislabeling that can plague manual systems; and their propensity for false positives is likely no worse in practice than it is for humans. They are thus well suited for a preliminary labeling task for high-confidence guesses, culling the more ambiguous data, which requires an expert or crowdsourced opinion. While machine processing at this stage is by no means perfect, these systems will only continue to improve. Crowdsourcing initiatives—such as Snapshot Serengeti—provide rich sources of labeled data, making training sets available to researchers globally. Cultivating and sharing this information allows for the training of CNNs to

better learn prominent patterns and distinctions between various species, to integrate a more diverse knowledge base, and to exploit automated understanding to better address pressing ecological concerns. For this reason, current research projects should be encouraged to share image and labeling data so as to collaboratively work on developing systems sophisticated enough to parallel human ability.

Thus, existing technologies for species recognition can currently be coupled with models of population density that do not require individuation, and can clearly do so in an accurate manner. There is, of course, room for improvement: performance in some domains is still inferior to humans. However, by adopting these innovations early on, an increased interest and collaboration can drive improvement in this area, namely, by producing more training data for CNNs to learn from. As shown from state-of-the-art facial recognition systems, our conceptual knowledge recognizing individual variation in humans has reached near-human (or perhaps even better) performance, and if the feature detection principles can indeed be transferred to individual animal recognition, the capability to accurately identify individual animals from camera trap data is likely on the horizon. Again, improvements will be driven by increased interest and work in this area, and thus putting confidence (and money) in the viability of such a system would be a major driver of progress. In the near future, these systems will undoubtedly be able to accurately monitor animal populations in a fraction of the time it currently takes for manual organization, labeling, and further analysis.

ACKNOWLEDGMENTS

The authors would like to thank Barbara Murck for reading and providing critical commentary, Maria Povarova and Diego Mosquera for their expertise on the subject of camera traps, and Davi Frossard for preparing the data used for VGG16 analysis.

REFERENCES

Best-Rowden, L., Han, H., Otto, C., Klare, B., & Jain, A. 2014. Unconstrained Face Recognition: Identifying a Person of Interest from a Media Collection (Technical Report MSU-CSE-14-1). http://www.cse.msu.edu/rgroups/biometrics/Publications/Face/BestRowdenetal_UnconstrainedFaceRecognition_TechReport_MSU-CSE-14-1.pdf (retrieved September 17, 2017).

Deng, J., Dong, W., Socher, R., Li, L., Li, K., & Fei-Fei, L. 2009. ImageNet: A Large-Scale Hierarchical Image Database. *2009 IEEE Conference on Computer Vision and Pattern Recognition*.

Fegraus, E. H., Lin, K., Ahumada, J. A., Baru, C., Chandra, S., & Youn, C. 2011. Data acquisition and management software for camera trap data: A case study from the TEAM Network. *Ecological Informatics*, 6(6), 345–353.

Gómez, A., Salazar, A., & Vargas, F. 2016. Towards Automatic Wild Animal Monitoring: Identification of Animal Species in Camera-trap Images using Very Deep Convolutional Neural Networks. http://arxiv.org/abs/1603.06169 (retrieved December 2, 2017).

He, K., Zhang, X., Ren, S., & Sun, J. 2015. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. *2015 IEEE International Conference on Computer Vision (ICCV)*.

Hernández-Serna, A. & Jiménez-Segura, L. F. 2014. Automatic identification of species with neural networks. *PeerJ*, 2, e563.

Hoffman, J., Tzeng, E., Donahue, J., Jia, Y., Saenko, K., & Darrell, T. 2014. One-Shot Adaptation of Supervised Deep Convolutional Models. http://arxiv.org/pdf/1312.6204.pdf (retrieved December 2, 2017)

Kemelmacher-Shlizerman, I., Seitz, S. M., Miller, D., & Brossard, E. 2016. The MegaFace Benchmark: 1 Million Faces for Recognition at Scale. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

LeCun, Y., Bengio, Y., & Hinton, G. 2015. Deep learning. *Nature*, 521(7553), 436–444.

Masi, I., Tran, A., Leksut, J., Hassner, T., & Medioni, G. 2016. Do We Really Need to Collect Millions of Faces for Effective Face Recognition? http://arxiv.org/abs/1603.07057 (retrieved September 6, 2017)

Merow, C., Smith, M. J., & Silander, J. A. 2013. A practical guide to MaxEnt for modeling species' distributions: What it does, and why inputs and settings matter. *Ecography*, 36(10), 1058–1069.

Nguyen, A, Yosinski, J, & Clune, J. 2015. Deep Neural Networks are Easily Fooled: High Confidence Predictions for Unrecognizable Images. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE.

Pearson, R. G., Raxworthy, C. J., Nakamura, M., & Townsend Peterson, A. 2006. Predicting species distributions from small numbers of occurrence records: A test case using cryptic geckos in Madagascar. *Journal of Biogeography*, 34(1), 102–117.

Rowcliffe, J. M., Field, J., Turvey, S. T., & Carbone, C. 2008. Estimating animal density using camera traps without the need for individual recognition. *Journal of Applied Ecology*, 45(4), 1228–1236.

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z. et al. 2015. ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision*, 115(3), 211–252.

Simonyan, K. & Zisserman, A. 2014. Very Deep Convolutional Networks for Large-Scale Visual Recognition. http://www.robots.ox.ac.uk/~vgg/research/very_deep/ (August 22, 2016).

Sundaresan, S. R., Riginos, C., & Abelson, E. S. 2011. Management and analysis of camera trap data: Alternative approaches (Response to Harris et al. 2010). *Bulletin of the Ecological Society of America*, 92(2), 188–195.

Swanson, A., Kosmala, M., Lintott, C., Simpson, R., Smith, A., & Packer, C. 2015. Snapshot Serengeti, high-frequency annotated camera trap images of 40 mammalian species in an African savanna. *Scientific Data*, 2, 150026.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. 2015. Going Deeper With Convolutions. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE.

Tsao, D. Y. & Livingstone, M. S. 2008. Mechanisms of face perception. *Annual Review of Neuroscience*, 31(1), 411–437.

Zaragozí, B., Belda, A., Giménez, P., Navarro, J. T., & Bonet, A. 2015. Advances in camera trap data management tools: Towards collaborative development and integration with GIS. *Ecological Informatics*, 30, 6–11.

Section III

Case Studies and Applications

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

12 UAV-Based Multispectral Images for Investigating Grassland Biophysical and Biochemical Properties

Bing Lu and Yuhong He

CONTENTS

12.1 INTRODUCTION TO UAV-BASED REMOTE SENSING

Satellites and airplanes have been extensively used in past decades to acquire remote sensing imagery. However, imagery from such platforms is typically restricted by weather conditions, operational complexity, acquisition cost, and technical constraints on spatial and temporal resolution (Zhang and Kovacs 2012). The lack of high spatial and temporal resolution imagery poses problems when studying grasslands with small plants, mixed communities, and rapid changes. In recent decades, unmanned aerial vehicles (UAVs) have been developed rapidly and were also applied as low-altitude

remote sensing platforms. UAV-based remote sensing shows many advantages over satellite- or airplane-based remote sensing, including but not limited to its low operation cost, low flight altitude (thus less limited by weather conditions), ability to acquire imagery at very high spatial resolution (centimeter scale), and high flexibility of deployment for repeat missions (Hunt et al. 2010; Laliberte et al. 2011).

In recent years, different types of UAVs have been utilized for remote sensing research; commonly used ones include unmanned multirotor copters (e.g., quadcopter, hexacopter, octocopter) (Nebiker et al. 2008; Córcoles et al. 2013; Del Pozo et al. 2014), unmanned helicopters (Nebiker et al. 2008; Swain et al. 2010), and fixed-wing planes (Hunt et al. 2008, 2010; Laliberte et al. 2011). Such UAVs have been utilized widely in various areas, such as precision agriculture, forest inventory, grassland mapping, fire detection and monitoring, invasive species detection, hydrology and riparian applications, pipeline inspection, movie production, and traffic monitoring (von Bueren et al. 2015). UAV systems typically integrate a module that controls autonomous flights based on predefined waypoints with global positioning system (GPS) information. It is thus convenient to fly over a targeted study area with limited human intervention. Various lightweight sensors, including digital cameras, modified digital cameras with a near-infrared (NIR) band, multispectral sensors, and hyperspectral sensors (Lelong 2008; Nebiker et al. 2008; Córcoles et al. 2013; Lucieer et al. 2014), have been developed to be deployed on UAVs. Since motion of a UAV is based on accurate stabilization of the whole system (e.g., rotor, frame), there are limitations for mount payloads (e.g., sensors) in terms of weight, size, or power supply. Adding payloads (e.g., using a heavy sensor) may cause imbalance to the UAV system. Digital cameras are therefore widely used because they can be lightweight, low cost, and easy to operate.

UAVs have been widely applied in various vegetation studies including precision agriculture and forest inventory (Nebiker et al. 2008; Laliberte et al. 2011; Gini et al. 2012; Zaman-Allah et al. 2015; Luna and Lobo 2016), but less so in grassland areas. A few existing grassland applications include rangeland mapping and species classification (Rango et al. 2006; Laliberte and Rango 2011; Lu and He 2017). Grasslands show distinct features compared with agricultural fields and forests, such as small plants and highly mixed species; therefore it is important to evaluate performance of UAV-based remote sensing for estimating grassland vegetation properties.

12.2 INVESTIGATING GRASSLAND PROPERTIES USING UAV

This study used a modified digital camera mounted on a multirotor UAV to investigate vegetation properties in a tall grassland. The goal was to establish a workflow for processing UAV-acquired images (e.g., mosaicking, geometric correction, and radiometric calibration), estimating vegetation biophysical and biochemical properties, and investigating spatial variations of these properties. The vegetation biophysical and biochemical properties that were investigated in this study include leaf area index (LAI) and chlorophyll content. LAI is a key canopy structural characteristic indicating vegetation physiological processes (e.g., photosynthesis, interception of radiation), while chlorophyll content is a fundamental factor governing vegetation photosynthetic capacity and productivity (Darvishzadeh et al. 2008).

12.2.1 Study Area

The study was conducted at Koffler Scientific Reserve (KSR) at Jokers Hill, located in Southern Ontario, Canada (Figure 12.1). The grassland in this reserve is temperate tall grassland, with grass height reaching up to 120 cm in the growing season (May to September). The mean temperature in this area ranges from −10°C (February) to 30°C (July), and monthly precipitation varies from 20 mm (March) to 100 mm (July) (Environment Canada 2016). The common grass species in this area are awnless brome (*Bromus inermis* Leyss), Canada goldenrod (*Solidago canadensis*), fescue (*Festuca rubra* L.), and milkweed (*Asclepias* L.).

12.2.2 UAV System, Sensor, and Image Acquisition

This study used a Tarot T15 Octocopter (Tarot RC, Wenzhou, China), owned and operated by the Flight Systems and Control Group at the University of Toronto Institute for Aerospace Studies (Figure 12.2). The UAV system weighs approximately 4.5 kg and has a maximum endurance of 20 minutes. Using an autopilot flight control system, the parameters of flight mission, such as area, path, altitude, and speed, were predetermined. Flights were conducted with the permission of the airspace authority. The imaging sensor mounted on the UAV was a modified Canon PowerShot ELPH 110 HS (Canon Inc., Tokyo, Japan). Typically, such digital cameras are limited to red, green, and blue optical bands. To capture the vegetation features in the NIR range,

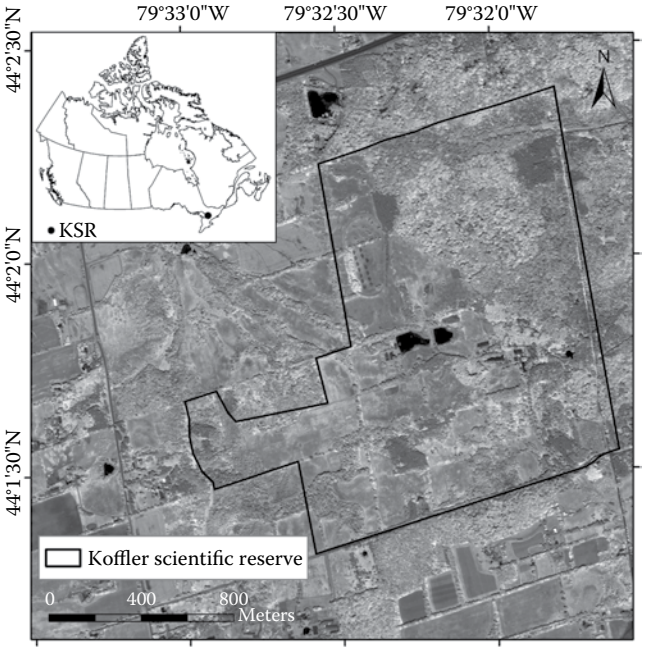


FIGURE 12.1 Map of Koffler Scientific Reserve. The background image was from a Quickbird near-infrared band, acquired on July 14, 2013.

FIGURE 12.2 UAV and camera. The arrow shows the camera mounting at the bottom of UAV.

the original red filter was replaced with an NIR filter to permit sensing in the NIR band and thus produce NIR-green-blue imagery (LDP LLC, Carlstadt, New Jersey). This camera weighed 135 g and acquired images with a resolution of 16.1 megapixels.

Seven UAV flights were conducted from April to December 2015 to investigate the complete grassland growing cycle. For each flight mission, the UAV was flown at an altitude of approximately 70 m above ground and a speed of 7 m/s. Imagery was acquired every second, so that the acquired images had approximately 85% forward overlap and 50% side overlap. The overlap is crucial to determining the imaging location and orientation for image mosaicking. Each flight mission was separated into three sub-missions due to the limits of the UAV's battery life (approximately 20 minutes). All flights were conducted under stable illumination condition with fixed camera settings in order to acquire images under uniform conditions. In this chapter, the imagery collected on July 9, 2015, was selected to investigate the fast growing season vegetation LAI and canopy chlorophyll content as well as their spatial variations.

12.2.3 Field Investigation

A total of 23 study sites, varying in species composition, vegetation growth status, and topographic conditions, were preselected. Each site was circular, with a 1 m diameter, and its boundary marked with colored flagging tape. Uniform white foam board (38×50 cm) labeled with the site identification was fixed on the ground adjacent to each study site. These foam boards were expected to be visible in the UAV-acquired imagery. The field survey was conducted simultaneously with UAV flights to measure vegetation properties and environmental factors, including LAI, species composition, canopy reflectance, vegetation height, and soil moisture. LAI was measured using an AccuPAR LP-80 ceptometer (Decagon Devices, Inc., Pullman, Washington). Canopy reflectance was collected using an ASD FieldSpec 3 Max field-portable spectroradiometer in the spectral range of 350–2500 nm (Analytical Spectral Devices Inc., Boulder, Colorado). The reflectance data were later used for image radiometric

calibration. Soil moisture was measured using HydroSense II (Campbell Scientific, Inc., Logan, Utah). GPS information for all 23 study sites and an additional 12 ground features (e.g., road intersections and individual trees) was collected using Trimble GeoExplorer (Trimble Navigation Limited, Sunnyvale, California) for image geometrical correction. Leaf samples for each site were also collected and taken to a laboratory for chlorophyll extraction following Minocha et al. (2009).

12.2.4 Imagery Process

A total of 700 images were acquired during the flight mission on July 9, 2015. The image processing steps included quality evaluation, orthorectification and mosaicking, geometric correction, and radiometric calibration (Figure 12.3).

12.2.4.1 Image Quality Evaluation and Process

Image quality evaluation involved a visual examination of image quality (e.g., oblique scenes, blurry) and evaluation of image radiometric distortion (e.g., assessing vignetting and atmospheric effects). Oblique scenes were identified in some images, largely due to the roll of the UAV system in flight, especially when changing direction. A small number of images were blurry, probably due to UAV vibration in response to wind gusts. Images with these issues were removed from analysis. Radiometric distortion can be introduced by various factors in the image acquisition process. Lebourgeois et al. (2008) discussed these factors and grouped them into two categories: camera-related factors (e.g., camera settings, vignetting effect) and environment-dependent factors (e.g., atmospheric effect, topographic condition). Our images were acquired

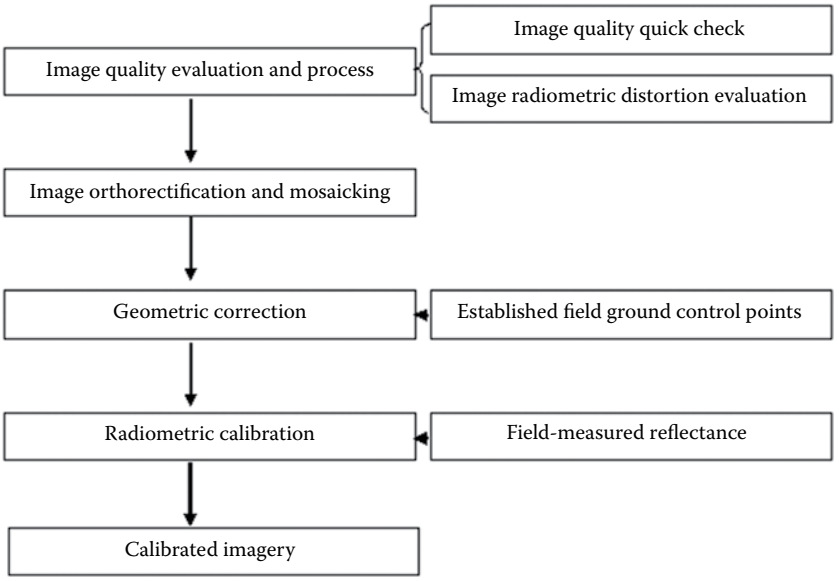


FIGURE 12.3 Processing steps for UAV-acquired images.

using invariant camera settings and under stable illumination conditions, so the impact of these factors on the images was constant. The UAV was flown at a low elevation (70 m) in a relatively flat grassland area; therefore, the influence of atmospheric and topographic variation was negligible (Hunt et al. 2008; Knoth et al. 2013; Del Pozo et al. 2014). This study emphasizes evaluating the vignetting effect.

The vignetting effect refers to brightness attenuation or gradual darkening of an image away from the image center and is an artifact prevalent in digital photography (Lebourgeois et al. 2008; Del Pozo et al. 2014). The vignetting effect results from the optical properties of camera lenses and is geometric in nature, in that less light is received by the camera sensor at the edges than in the center (Lebourgeois et al. 2008). Various approaches have been developed to evaluate the vignetting effect in previous studies. For instance, Lebourgeois et al. (2008) calculated an average from 500 images (separated bands) that were acquired in different locations and elevations in a field experiment. The resultant mean image was applied to determine the impact of the vignetting effect and image correction. This approach requires a great deal of computational power, and the accuracy of the mean image could be influenced by numerous ground features (e.g., highly reflective ground features distributed in the edge area of photos). Del Pozo et al. (2014) took a series of photographs of a white pattern with low-specular reflection in a laboratory under uniform illumination conditions, and then compared the pixel values in the photo center area with those in the edge area. However, it is a challenge to achieve such imaging conditions in field settings. In this study, the vignetting effect was evaluated by comparing pixel values of the same study site but extracted from different images. Since there is large overlap among adjacent photos, one study site can potentially be located at different locations in the images (i.e., located either in the image center, the edge, or the area between). Therefore, pixel values can be compared by the distance of the study site to the image center. The digital values of pixels within the study site area were extracted in ENVI (Visual Information Solutions Inc., Boulder, Colorado) and averaged, providing one mean pixel value of the study site for each band in each image. The results of one study site are shown in Figure 12.4.

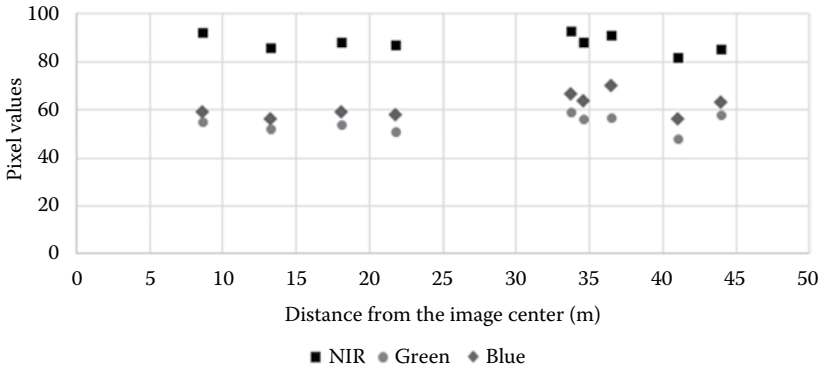


FIGURE 12.4 Comparison of pixel values of one study site by distance to image center in different images. NIR, green, and blue band pixel values were compared within each band.

If the image suffered severely from vignetting effects, the study site pixel values would decrease further away from the image center (Lebourgeois et al. 2008). However, there is no clear decrease in pixel values accompanying distance from the image center for all three spectral bands (Figure 12.4). The minor variations in pixel values for each band were potentially a result of the combined vignetting effect, different view angles, and the averaging of pixel values. We therefore concluded that the vignetting effects were minimal, and they were thus ignored. In addition, during the mosaicking step, pixel values in the images with more than 50% overlap were averaged, further reducing any potential vignetting effects.

12.2.4.2 Image Orthorectification and Mosaic

Orthorectification is an important processing step for UAV-acquired imagery (Turner et al. 2012). It was conducted to rectify images with the correct position relative to ground features (Laliberte et al. 2010). For each flight mission, the UAV can obtain hundreds to thousands of images, which must be mosaicked to produce an image covering the entire area. This study used PhotoScan software (Agisoft LLC, St. Petersburg, Russia), which is based on structure from motion (SfM) to generate orthorectified and mosaicked imagery (Stafford 2013; Whitehead and Hugenholtz 2014). General processing steps in PhotoScan include camera alignments, point cloud generation, 3D mesh, texture building, and orthomosaic generation (Agisoft LLC 2016). Camera alignment is used to find the position of each photo that relies on overlap among images. After camera alignment is complete, a point cloud is generated to represent the positions of photos. Mesh and texture are established sequentially, to represent the surfaces of objects. The orthomosaic is then generated.

12.2.4.3 Geometric Correction

The orthomosaic imagery generated in the previous step was geometrically corrected in ArcMap (ESRI, Redlands, California), using GPS information of ground control points (GCPs). The GCPs included both study sites and an additional 12 ground features.

12.2.4.4 Radiometric Calibration

Radiometric calibration was conducted to convert digital number (DN) values to reflectance, which can better represent spectral properties of ground features (Teillet 1986). Typically, DN values from remote sensing imagery are converted to radiances using sensor gain and bias values, which are then divided by solar radiance to produce reflectance (Jensen 2006). Since the gain and bias values were not available for our sensor, the widely used empirical line method was applied to correlate the DN values with ground-measured reflectance (von Bueren et al. 2015). DNs of pixels within the circle of each study site were extracted from imagery and averaged. Ground-measured reflectance of each site was averaged along each spectral band (i.e., NIR, green, and blue). Linear relationships between DNs and ground-measured reflectances were then established for each of the three bands (i.e., NIR, green, and blue) individually. Data from 12 sites were used for establishing regression lines, and data from the remaining 11 sites were used for validation. Figure 12.5 shows the linear relationships for three bands of the imagery acquired on July 9, 2015. Linear regressions were then applied to convert DNs to reflectance for the entire image; the final reflectance map is shown in Figure 12.6.

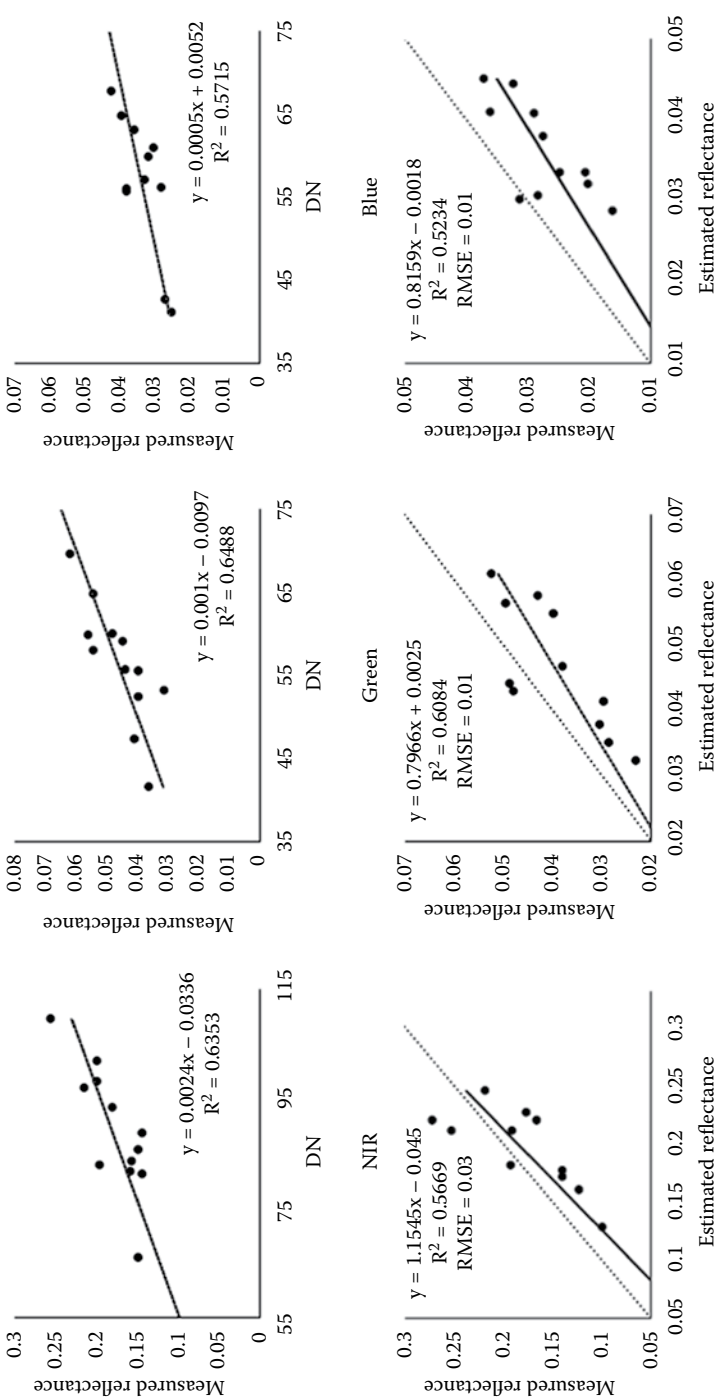


FIGURE 12.5 Linear relationships between imagery DNs and ground-measured reflectances. The top three figures show correlations built and the bottom three show validation results.

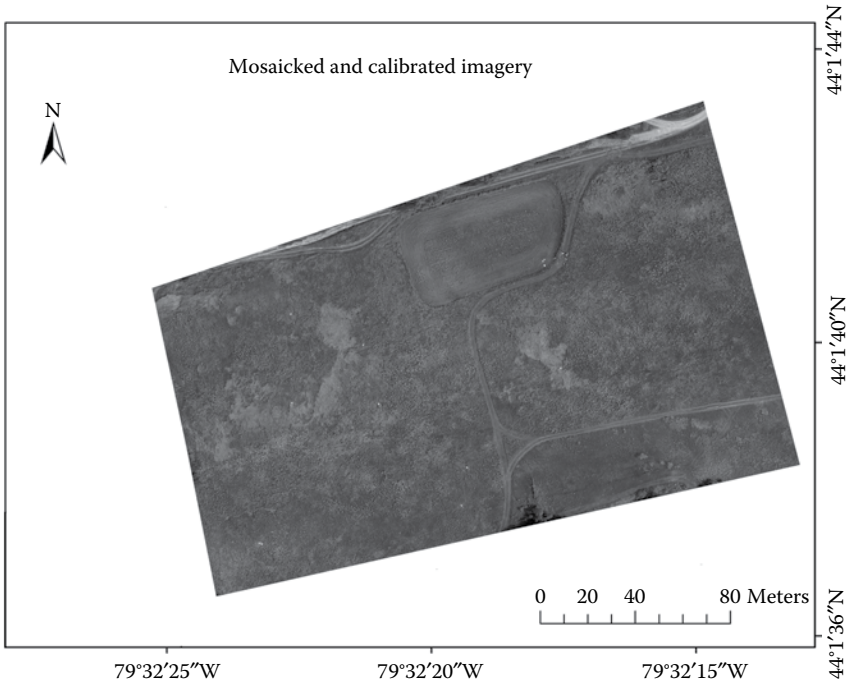


FIGURE 12.6 Mosaicked and calibrated imagery, using images acquired on July 9, 2015.

12.2.5 Vegetation Properties Estimation and Analysis

Spectral vegetation indices are one of the most important tools for estimating vegetation properties from remote sensing imagery (He et al. 2006). Various vegetation indices have been developed to enhance vegetation signals, reduce influence of backgrounds (e.g., soil), and investigate different vegetation properties (Tong 2014). Normalized difference vegetation index (NDVI) is one of the most widely used indices for estimating vegetation properties, and is calculated with spectral reflectance at the NIR and red bands using Equation 12.1 (Glenn et al. 2008)

$$\text{NDVI} = \frac{\text{NIR} - \text{red}}{\text{NIR} + \text{red}} \tag{12.1}$$

The camera used in this study does not measure in the red band and thus the NDVI cannot be calculated. To compensate, the camera manufacturer suggested using a blue NDVI (Equation 12.2) that replaces the red band with a blue band, since the blue band is also an absorption band for vegetation reflection that is similar to red (LDP LLC 2015). Good relationships between blue NDVI with vegetation features has been confirmed by the manufacturer.

$$\text{Blue NDVI} = \frac{\text{NIR} - \text{blue}}{\text{NIR} + \text{blue}} \tag{12.2}$$

Blue NDVI values for 12 study sites were calculated, and then correlated with corresponding field-measured LAI values (Figure 12.7). This relationship was validated using ground-measured LAI values from the remaining 11 sites (Figure 12.7). The LAI map was then produced using the established regression equation (Figure 12.8).

The spatial variation in LAI values is obvious in Figure 12.8. Area a in Figure 12.8 is a gravel road where there was no vegetation, and therefore LAI is 0. Areas b and c are human-disturbed areas (e.g., cultivation) where there was sparsely distributed vegetation. Area d is an upland area, where the elevation is 1 m higher

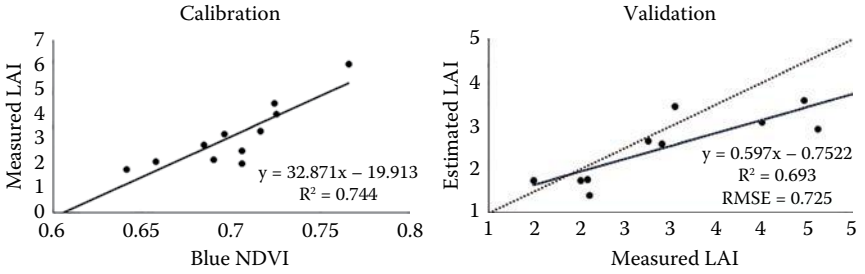


FIGURE 12.7 Correlation between blue NDVI and measured LAI (left) and results validation (right).

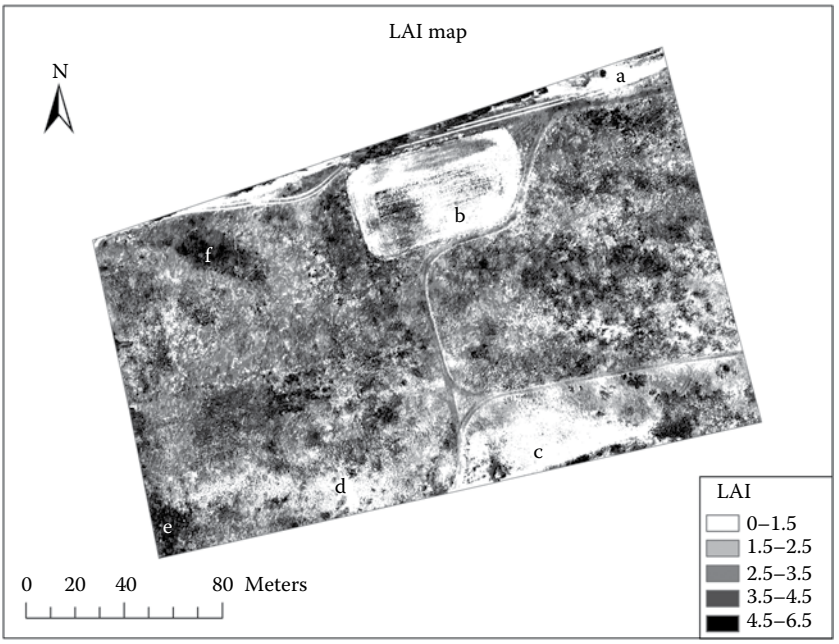


FIGURE 12.8 LAI map generated using blue NDVI (map accuracy 75.34%; imagery acquired on July 9, 2015; labels a–f show areas with different LAI values).

than the surroundings. Vegetation along this upland was sparse, likely due to less water availability. In contrast, Areas e and f were occupied with dense vegetation and had high LAI values. The dominant species in Areas e and f was Canada goldenrod (*Solidago canadensis*), which was flourishing in the summer with an average plant height of approximately 1 m. Field-measured LAI values in Areas e and f are typically above 5. Other areas in the map have moderate LAI values ranging from 1.5 to 4.5.

For chlorophyll content estimation, Gitelson et al. (2005) developed a green chlorophyll index (CI_{green}) (Equation 12.3), and found it correlated well with chlorophyll content. We thus applied CI_{green} to estimate grassland chlorophyll content from UAV-acquired imagery.

$$CI_{green} = \frac{NIR}{Green} - 1 \tag{12.3}$$

CI_{green} values from 12 study sites were correlated with measured chlorophyll content from ground sampling (Figure 12.9). The canopy chlorophyll contents were obtained by multiplying averaged leaf chlorophyll contents with canopy LAI. The regression model was validated using the measured chlorophyll content of the remaining 11 sites (Figure 12.9) and applied to the entire image to produce the chlorophyll map (Figure 12.10).

Spatial variation of canopy chlorophyll is clear (Figure 12.10), but not as obvious as the spatial variation of LAI. In Areas a, b, c, and d, where there was no vegetation or sparse vegetation with low LAI values, the canopy chlorophyll contents are also low. In Areas e and f, with dense vegetation and high LAI values, the canopy chlorophyll contents are not much higher than in the surrounding areas. This is likely due to the fact that the leaf chlorophyll contents of Canada goldenrod, the dominant species in Areas e and f, are generally lower than those of other species (e.g., awnless brome), which were widely distributed in adjacent areas. Area g, with abundant awnless brome, shows much higher canopy chlorophyll contents. We also found that the canopy chlorophyll content in Areas h and i is much lower than those of surrounding areas. This is because the grasses in Areas h and i fell to the ground due to strong winds as field observations were being made.

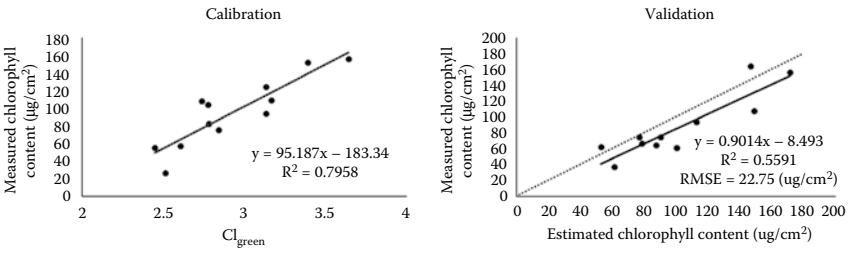


FIGURE 12.9 Correlation between CI_{green} and chlorophyll contents (left) and validation results (right).

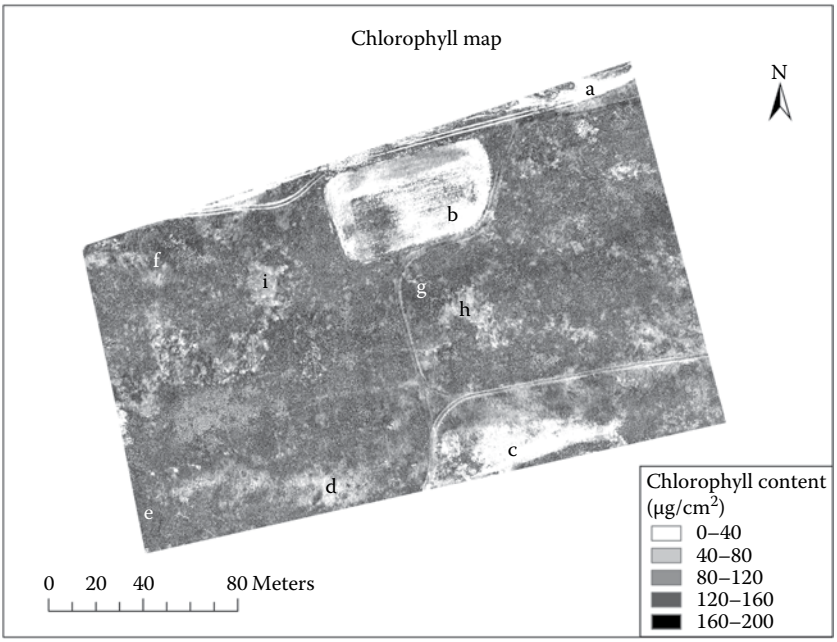


FIGURE 12.10 Canopy chlorophyll map produced using CI_{green} (map accuracy 74.13%; imagery acquired on July 9, 2015).

12.3 PROJECT SUMMARY

We used a UAV mounted with a modified digital camera to take images in a grassland area during the growing season to measure vegetation LAI and canopy chlorophyll content. The quality of the acquired images was evaluated, and they were orthorectified, mosaicked, geometrically corrected, and radiometrically calibrated. The vegetation indices blue NDVI and CI_{green} were used to map LAI and canopy chlorophyll content, respectively, and the spatial variations of these vegetation properties were then analyzed. This project demonstrated that a UAV mounted with a consumer digital camera is capable of remotely measuring vegetation biophysical and biochemical properties. Owing to the high spatial resolution of the acquired imagery, we were able to observe highly detailed variation in these properties. Species classification potentially can also be conducted at this very fine level. A UAV system can be deployed repeatedly to acquire images in different seasons for temporal analysis. We also collected imagery from this study area in the fall and winter of 2015, and the temporal variations of grassland properties from these periods will be discussed in future work.

We found that the most important step in the image processing workflow was to mosaic a large number of images together; this step can be completed effectively using software packages like PhotoScan. Other procedures (e.g., radiometric calibration) can be skipped or simplified to achieve a quick and rough analysis if field data are

not available. For instance, the distribution of different species or vegetation growing conditions can be analyzed roughly using raw mosaicked imagery. Vegetation indices can likely be calculated using raw DNs, rather than reflectance, as has been used in previous studies.

12.4 CHALLENGES AND FUTURE WORK WITH UAV APPLICATIONS

12.4.1 UAV Systems

Payload capacity is a major challenge for a wide application of UAV-based remote sensing. Limited by this capacity, only lightweight sensors can currently be mounted on UAVs, in turn limiting the potential to acquire high-quality multispectral or hyperspectral imagery. Therefore, there is a need to develop UAV systems with higher payload capacity, and to develop lightweight sensors that can provide more spectral channels for imaging. A standard protocol for processing UAV-acquired imagery (e.g., radiometric calibration) is also needed. Battery endurance is another major factor limiting the flight time of UAVs. In our study we used lithium polymer (LiPo) batteries, with a set of fully charged batteries enabling flying times of approximately 20 minutes. Another challenge was wind, which destabilizes the UAV system and causes oblique (off-nadir) views that may result in radiometric inconsistency. If the UAV is caught in turbulence, acquired imagery may be blurry and data lost.

12.4.2 UAV Flight Designs and Image Processing

UAV systems are capable of acquiring very high spatial resolution imagery. However, the lower-altitude flight and greater overlap between images result in a larger number of images for the same study area, which not only requires large memory capacity in the camera but also requires more processing time. Those interested in UAV imaging are advised to determine an optimal flight altitude and overlap proportion. In this study, the atmospheric effect and other effects (e.g., bidirectional reflectance distribution function [BRDF]) were not accounted for. Further analysis of these effects in UAV applications is needed.

ACKNOWLEDGMENTS

This study was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant (RGPIN-386183) to Dr. Yuhong He and Graduate Expansion Funds from the Department of Geography, University of Toronto Mississauga. The authors are grateful to the UAV flight crew from Arrowonics Technology Ltd. and to the managers of the study area, Koffler Scientific Reserve. Field assistance from a group of students enrolled in the Research Opportunity Program and Work Study Program at the University of Toronto Mississauga is acknowledged. Thanks also to Transport Canada for providing a Special Flight Operating Certificate (Numbers 5812-15-14-2016-1, 5812-15-14-2014-3).

REFERENCES

Agisoft LLC. 2016. Agisoft PhotoScan User Manual.

Córcoles, J. I., J F. Ortega, D. Hernández, and M. A. Moreno. 2013. Estimation of Leaf Area Index in Onion (*Allium cepa* L.) Using an Unmanned Aerial Vehicle. *Biosystems Engineering* 115 (1):31–42.

Darvishzadeh, R., A. Skidmore, M. Schlerf, and C. Atzberger. 2008. Inversion of a Radiative Transfer Model for Estimating Vegetation LAI and Chlorophyll in a Heterogeneous Grassland. *Remote Sensing of Environment* 112 (5):2592–2604.

Del Pozo, S., P. Rodríguez-Gonzálvez, D. Hernández-López, and B. Felipe-García. 2014. Vicarious Radiometric Calibration of a Multispectral Camera on Board an Unmanned Aerial System. *Remote Sensing* 6 (3):1918–1937.

Environment Canada. 2016. Historical Climate Data.

Gini, R., D. Passoni, L. Pinto, and G. Sona. 2012. Aerial Images from an UAV System: 3D Modeling and Tree Species Classification in a Park Area. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 39:361–366.

Gitelson, A. A., A. E. S. Vina, V. O. Nica Ciganda, D. C. Rundquist, and T. J. Arkebauer. 2005. Remote Estimation of Canopy Chlorophyll Content in Crops. *Geophysical Research Letters* 32 (8):1–4.

Glenn, E. P., A. R. Huete, P. L. Nagler, and S. G. Nelson. 2008. Relationship Between Remotely-Sensed Vegetation Indices, Canopy Attributes and Plant Physiological Processes: What Vegetation Indices Can and Cannot Tell Us About the Landscape. *Sensors* 8 (4):2136–2160.

He, Y., X. Guo, and J. Wilmshurst. 2006. Studying Mixed Grassland Ecosystems I: Suitable Hyperspectral Vegetation Indices. *Canadian Journal of Remote Sensing* 32 (2):98–107.

Hunt, E. R., W. D. Hively, C. S.T. Daughtry, G. W. McCarty, S. J. Fujikawa, T. L. Ng, M. Tranchitella, D. S. Linden, and D. W. Yoel. 2008. Remote Sensing of Crop Leaf Area Index Using Unmanned Airborne Vehicles. In *17th William T. Pecora Memorial Remote Sensing Symposium*. Denver, CO.

Hunt, E. R. Jr., W. D. Hively, S. J. Fujikawa, D. S. Linden, C. S. T. Daughtry, and G. W. McCarty. 2010. Acquisition of NIR-Green-Blue Digital Photographs from Unmanned Aircraft for Crop Monitoring. *Remote Sensing* 2 (1):290–305.

Jensen, J. R. 2006. *Remote Sensing of the Environment: An Earth Resource Perspective*: Prentice Hall.

Knoth, C., B. Klein, T. Prinz, and T. Kleinebecker. 2013. Unmanned Aerial Vehicles as Innovative Remote Sensing Platforms for High-Resolution Infrared Imagery to Support Restoration Monitoring in Cut-Over Bogs. *Applied Vegetation Science* 16 (3):509–517.

Laliberte, A. S., J. E. Herrick, A. Rango, and C. Winters. 2010. Acquisition, Orthorectification, and Object-Based Classification of Unmanned Aerial Vehicle (UAV) Imagery for Rangeland Monitoring. *Photogrammetric Engineering & Remote Sensing* 76 (6):661–672.

Laliberte, A. S., M. A. Goforth, C. M. Steele, and A. Rango. 2011. Multispectral Remote Sensing from Unmanned Aircraft: Image Processing Workflows and Applications for Rangeland Environments. *Remote Sensing* 3 (12):2529–2551.

Laliberte, A. S., and A. Rango. 2011. Image Processing and Classification Procedures for Analysis of Sub-Decimeter Imagery Acquired with an Unmanned Aircraft Over Arid Rangelands. *GIScience & Remote Sensing* 48 (1):4–23.

LDP LLC. 2015. Canon PowerShot ELPH 110 Hs 16.1 Megapixel Digital Camera User Manual.

Lebourgeois, V., A. Bégué, S. Labbé, B. Mallavan, L. Prévot, and B. Roux. 2008. Can Commercial Digital Cameras Be Used as Multispectral Sensors? A Crop Monitoring Test. *Sensors* 8 (11):7300–7322.

Lelong, C. C. D. 2008. Assessment of Unmanned Aerial Vehicles Imagery for Quantitative Monitoring of Wheat Crop in Small Plots. *Sensors* 8 (5):3557–3585.

Lu, B., and Y. He. 2017. Species Classification Using Unmanned Aerial Vehicle (UAV)-Acquired High Spatial Resolution Imagery in a Heterogeneous Grassland. *ISPRS Journal of Photogrammetry and Remote Sensing* 128:73–85.

Lucieer, A., Z. Malenovský, T. Veness, and L. Wallace. 2014. HyperUAS—Imaging Spectroscopy from a Multirotor Unmanned Aircraft System. *Journal of Field Robotics* 31 (4):571–590.

Luna, I., and A. Lobo. 2016. Mapping Crop Planting Quality in Sugarcane from UAV Imagery: A Pilot Study in Nicaragua. *Remote Sensing* 8 (6):500.

Minocha, R., G. Martinez, B. Lyons, and S. Long. 2009. Development of a Standardized Methodology for Quantifying Total Chlorophyll and Carotenoids from Foliage of Hardwood and Conifer Tree Species. *Canadian Journal of Forest Research* 39 (4):849–861.

Nebiker, S., A. Annen, M. Scherrer, and D. Oesch. 2008. A Light-Weight Multispectral Sensor for Micro UAV—Opportunities for Very High Resolution Airborne Remote Sensing. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 37:1193–1200.

Rango, A., A. Laliberte, C. Steele, J. E. Herrick, B. Bestelmeyer, T. Schmugge, A. Roanhorse, and V. Jenkins. 2006. Using Unmanned Aerial Vehicles for Rangelands: Current Applications and Future Potentials. *Environmental Practice* 8 (03):159–168.

Stafford, J. V. 2013. *Precision Agriculture '13*: Wageningen Academic Publishers.

Swain, K. C., S. J. Thomson, and H. P. W. Jayasuriya. 2010. Adoption of an Unmanned Helicopter for Low-Altitude Remote Sensing to Estimate Yield and Total Biomass of a Rice Crop. *Transactions of the ASABE* 53 (1):21–27.

Teillet, P. M. 1986. Image Correction for Radiometric Effects in Remote Sensing. *International Journal of Remote Sensing* 7 (12):1637–1651.

Tong, A. 2014. Estimating Grassland Chlorophyll Content for a Mixed Grassland: Exploring the Performance of the Empirical-Statistical and the Physical Modeling Approach. Master dissertation, Department of Geography, University of Toronto, Toronto, Canada.

Turner, D., A. Lucieer, and C. Watson. 2012. An Automated Technique for Generating Georectified Mosaics from Ultra-High Resolution Unmanned Aerial Vehicle (UAV) Imagery, Based on Structure from Motion (SfM) Point Clouds. *Remote Sensing* 4 (12):1392–1410.

von Bueren, S. K., A. Burkart, A. Hueni, U. Rascher, M. P. Tuohy, and I. J. Yule. 2015. Deploying Four Optical UAV-Based Sensors Over Grassland: Challenges and Limitations. *Biogeosciences* 12 (1):163–175.

Whitehead, K., and C. H. Hugenholtz. 2014. Remote Sensing of the Environment with Small Unmanned Aircraft Systems (UASs), Part 1: A Review of Progress and Challenges. *Journal of Unmanned Vehicle Systems* 02 (03):69–85.

Zaman-Allah, M., O. Vergara, J. L. Araus, A. Tarekegne, C. Magorokosho, P. J. Zarco-Tejada, A. Hornero, A. Hernandez Alba, B. Das, P. Craufurd, M. Olsen, B. M. Prasanna, and J. Cairns. 2015. Unmanned Aerial Platform-Based Multi-Spectral Imaging for Field Phenotyping of Maize. *Plant Methods* 11 (35):1–10.

Zhang, C., and J. M. Kovacs. 2012. The Application of Small Unmanned Aerial Systems for Precision Agriculture: A Review. *Precision Agriculture* 13 (6):693–712.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

13 Inversion of a Radiative Transfer Model Using Hyperspectral Data for Deriving Grassland Leaf Chlorophyll

Alexander Tong, Bing Lu, and Yuhong He

CONTENTS

13.1 INTRODUCTION

The investigation of leaf biochemical components, primarily foliar chlorophyll pigments, has been a major focus of research because they are responsible for the photosynthetic process that influences the physiological function of vegetation. Unlike other pigment groups including carotenoid and anthocyanin, which are essential leaf structural components and complementary to the photosynthetic process, chlorophyll

controls the amount of incident radiation that a leaf absorbs, and thus influences photosynthetic potential and primary production (Curran et al. 1990; Filella et al. 1995; Blackburn 2007). Additionally, chlorophylls can be used to detect vegetation stress by examining the total chlorophyll-to-carotenoid ratio; during periods of environmental stress and leaf senescence, chlorophylls are in lower concentration than carotenoids (Gitelson and Merzlyak 1994a,b; Merzlyak et al. 1999; Peñuelas and Filella 1998). Finally, chlorophylls can provide an indirect measure of nutrient status because leaf nitrogen, which is related to pigment formation and therefore leaf color, is contained within the structure of chlorophyll molecules (Filella et al. 1995; Moran et al. 2000).

Two different methods have been established for estimating chlorophyll at a range of spatiotemporal scales using remote sensing data. The empirical-statistical approach has been the most widely used, especially for estimating chlorophyll content at the leaf scale, using bivariate or multivariate analysis (i.e., regression models) to establish relationships between leaf chlorophyll content and spectral vegetation indices (SVIs) calculated from leaf spectral reflectance. SVIs are developed using linear combinations of reflectance in the visible (VIS), red edge, and near-infrared (NIR) spectral regions in which vegetation reveals distinct reflectance properties (Broge and Leblanc 2001; Darvishzadeh et al. 2008). The extrapolation of leaf-level relationships to larger scales has been challenging due to the confounding influence of canopy structure, background contributions, and nonphotosynthetic vegetation (Zarco-Tejada et al. 2000; Zhang et al. 2008). Indeed, many SVIs used to estimate chlorophyll content at the canopy level have been developed to minimize the spectral variability caused by external factors, such as background noise (e.g., rock, soil, litter), irradiance conditions, sun angle, and atmospheric composition (Broge and Mortensen 2002; Haboudane et al. 2008). Scaling chlorophyll content from the leaf to the canopy scale has been explored using three methods: direct extrapolation, canopy-integrated, and percent-cover-based methods. The empirical-statistical approach is often used because of its simplicity and computational efficiency. However, this approach lacks generality; that is, no universal relationships can be expected to exist between chlorophylls and spectral reflectance (Baret and Buis 2008). Consequently, statistically derived empirical relationships are sensor, species, location, and time specific (Baret and Guyot 1991; Houborg et al. 2007).

The alternative approach is physical modeling, which uses radiative transfer models (RTMs) based on complex mathematical methods to describe the two main physical processes of absorption and scattering in order to characterize the transfer and interaction of radiation at the leaf and canopy level (Kötz et al. 2004; Jacquemoud et al. 2009). Since RTMs are based on physical laws, they allow for an explicit connection between leaf or canopy spectral reflectance and a target biochemical (e.g., chlorophyll content) or biophysical [e.g., leaf area index (LAI)] parameter, respectively (Houborg et al. 2007). Models simulating leaf spectra have described the spectra as either the reflectance or transmittance of the light intercepted at a leaf's surface, which is a function of the concentration of light-absorbing compounds (e.g., chlorophyll, carotenoids, water, cellulose, lignin, starch, proteins), and the internal scattering of light that is either inefficiently or not absorbed (Jacquemoud and Ustin 2008). RTMs have simulated canopy reflectance as a function of foliage (e.g., leaves,

branches, stems) and background (e.g., soil, moss, litter) optical properties, canopy architecture, illumination conditions, and viewing geometry (Goel and Grier 1988; Chen et al. 2000). This approach does not suffer the same lack of generality as the empirical-statistical approach, and can be applied to a wide range of study areas and plant functional types and species over time, depending on the type of RTM (Houborg et al. 2007; Si et al. 2012).

RTMs can be employed in a direct (forward) or inverse mode. In the direct mode, simulated spectral reflectance values can be retrieved from either leaf or canopy models by parameterizing the input parameters (often from field measurements). By using the direct mode, spectral indices can be developed, especially ones optimized for a particular sensor (Ceccato et al. 2002; Jacquemoud et al. 2009). In the inverse mode, leaf models can be inverted to retrieve leaf biochemical parameters such as chlorophyll content, whereas canopy models have much more difficulty retrieving canopy biophysical parameters owing to the complexity of canopy structures (Schlerf and Atzberger 2006; Jacquemoud et al. 2009). Hence, leaf and canopy models are commonly coupled in order for inversion methods based on statistical and mathematical approaches to be effective for parameter retrieval. Different inversion algorithms include lookup table (LUT) methods (e.g., Weiss et al. 2000; Gastellu-Etchegorry et al. 2003), numerical optimization methods (e.g., Jacquemoud et al. 1995, 2000; Meroni et al. 2004), artificial neural networks (e.g., Kimes et al. 2002; Walthall et al. 2004), and support vector machines regression (e.g., Durbha et al. 2007). The limitation of model-based approaches for chlorophyll content estimation is the ill-posed inverse problem that not all inverted results may be unique. This issue can be minimized by adding additional input parameter information (e.g., biophysical measurements) into the model (Combal et al. 2003).

In this chapter, an endangered mixed-grass prairie ecosystem, located in Saskatchewan, Canada, was investigated using a physical modeling approach to estimate and predict leaf chlorophyll content. This study provides an exploratory evaluation of a coupled RTM, known as the Soil-Leaf-Canopy (SLC) model, for the retrieval of leaf chlorophyll via model inversion using a LUT-based algorithm. The SLC model had yet to be validated for leaf chlorophyll content estimation within a heterogeneous grassland environment, which contained vertically inclined heterogeneous canopies with fractional components of living and senesced material and exposed soil.

13.2 METHODS

13.2.1 Study Area

The study area was located in the West Block of Grasslands National Park (GNP) in southern Saskatchewan, Canada (Figure 13.1). The park was established in 1988 and is dedicated to preserving the largest intact area of northern mixed-grass prairie in North America (Csillag et al. 2001). This area is defined as a semiarid, mixed-grass prairie ecosystem which falls within the northern extent of the Great Plains (Black and Guo 2008). The GNP has a semiarid continental climate, with mean temperatures ranging from −22°C (January) to +33°C (July) and annual total precipitation of 270–460 mm (Lu et al. 2016). The dominant grass species found in the study

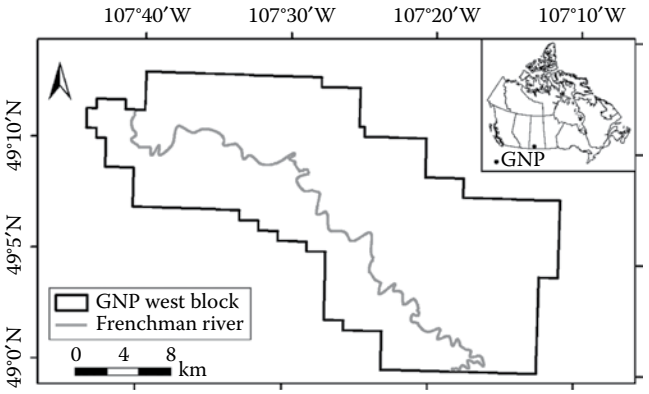


FIGURE 13.1 West Block of the Grasslands National Park, located in southern Saskatchewan, Canada.

area were junegrass (*Koeleria gracilis*), needle-and-thread grass (*Hesperostipa comata*), blue grama (*Bouteloua gracilis*), western wheatgrass (*Agropyron smithii*), northern wheatgrass (*Agropyron dasystachyum*), and crested wheatgrass (*Agropyron cristatum*) (Tong 2014).

13.2.2 Field Data

Fieldwork was conducted in June 2013, and measurements relating to vegetation biochemical content were collected within a 50 by 50 cm sampling frame representing the quadrats at each site. Measurements including percent coverage of the top canopy layer (e.g., grass, forbs, shrubs, standing dead vegetation) and lower canopy layer (e.g., litter, moss, lichen, rock, soil), and species composition were estimated and documented using photos. Soil moisture was recorded using a HydroSense II soil-water sensor (Campbell Scientific Canada, Edmonton, Alberta, Canada). LAI was indirectly measured using an AccuPAR LP-80 ceptometer (Decagon Devices, Inc., Pullman, Washington). Plant samples, including different species of grasses, forbs, and shrubs, were collected and kept on ice in a dark thermal cooler, to be immediately dispatched for leaf-level spectral reflectance measurements and chlorophyll extraction.

13.2.3 Field Spectroradiometric Measurements

Canopy spectral reflectance was measured approximately 1 m above each quadrat using the Analytical Spectral Devices (ASD) FieldSpec 3 Max field portable spectroradiometer (Boulder, Colorado). Additional items, including standing dead vegetation, litter, soil, and/or rocks, were also collected for spectral measurements. Care was taken to ensure that the fiber optic sensor was pointing at nadir to the surface (e.g., perpendicular) to avoid any spectral measurement discrepancies associated with differing offsets. The ASD spectroradiometer was frequently calibrated with a certified white reference to factor for changing irradiance conditions, thereby reducing measurement noise. Additionally, in order to minimize the effects of shadowing and

solar zenith changes, spectral measurements were only taken on days with clear sky conditions between 10:00 and 14:00 (two hours before and after solar noon).

13.2.4 Chlorophyll Measurements

Chlorophyll pigment was extracted from the collected plant samples with acetone, using a standard extraction protocol (Minocha et al. 2009). The extracts were then spectrophotometrically assayed at 662, 645, and 447 nm using a GENESYS 10S UV-Vis spectrophotometer (Thermo Fisher Scientific, Waltham, Massachusetts), corresponding to the absorbance peaks of chlorophyll *a*, chlorophyll *b*, and total carotenoids, respectively. The leaf pigment concentration for each sample was then calculated using the formula reported by Lichtenthaler (1987). The leaf pigment concentration was converted to leaf chlorophyll content using the area-to-weight ratio of each leaf sample:

$$Chl_{leaf_i} = CC_i R_i \tag{13.1}$$

where

Chl_{leaf_i} ($\mu g/cm^2$) = leaf-scale chlorophyll content for each species *i;*
CC_i ($\mu g/mg$) = chlorophyll concentration for each species *i*; and
R_i (g/cm^2) = area-to-weight ratio for each species *i*.

13.2.5 Model Inversion for Leaf Chlorophyll Content Retrieval

The estimation of leaf chlorophyll content using a physical modeling approach involves the inversion of a coupled RTM; a simplification of this method is described in Figure 13.2.

13.2.5.1 Model Selection: SLC Radiative Transfer Model

The SLC model by Verhoef and Bach (2007) is a recent development of the widely used PROSAIL model, which integrates a modified Hapke soil bidirectional reflectance distribution function (BRDF) model, a robust version of PROSPECT, and a 4SAIL2 model. Like PROSAIL, the SLC model has reasonable computation times and only requires a limited number of input parameters. Simulations are computed between 400 and 2500 nm at 10 nm step intervals, covering the VIS, NIR, and short-wave infrared (SWIR) regions. The model was developed to allow a canopy with two different leaf types, green and brown, in combination with soil to be modeled. Even though the SLC model was developed with forest scenes in mind, it is the only coupled model that is currently known to allow heterogeneous scenes with fractional coverage of these elements to be simulated, and for vertically inclined leaves to be modeled, which is characteristic of the GNP.

13.2.5.1.1 Submodel: Hapke Soil BRDF Model

Since most canopies exhibit fractional gaps that allow radiation to interact with the soil background, PROSAIL includes a soil reflectance input parameter. Consequently, PROSAIL treats the soil background as a Lambertian surface (e.g., diffuse isotropic reflector), which is not ideal for sparsely vegetated canopies, such as in grasslands

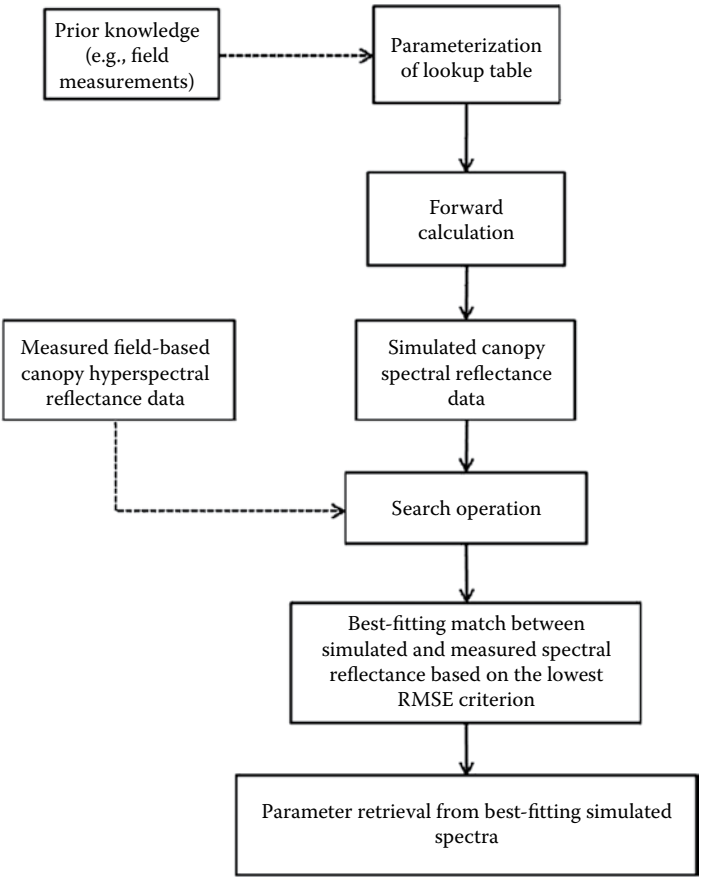


FIGURE 13.2 Overview of the physical modeling approach for retrieving chlorophyll content estimates at the leaf scale.

(Verhoef and Bach 2007). The original Hapke model (Hapke 1981; Hapke and Wells 1981) is able to calculate the hot spot effect of soil and expresses the interaction of radiation at the surface of a soil medium using an isotropic scattering phase function. Verhoef and Bach (2007) modified the model to include the use of a non-isotropic phase function to describe the soil background in terms of a non-Lambertian surface. Also included in the modified Hapke model is the spectral soil moisture effect. In order to simulate the effect of soil moisture, Verhoef and Bach (2007) made use of a physically based model by Bach and Mauser (1994), which considers two criteria that influence soil reflectance: (1) the specific absorption of water, and (2) the interactive effect of water on soil particles.

13.2.5.1.2 Submodel: PROSPECT Leaf Optical Model

The PROSPECT model has been updated several times since its inception by Jacquemoud and Baret (1990), and has evolved to include new leaf biochemical constituents (e.g., cellulose and lignin). The model simulates the directional-hemispherical reflectance

and transmittance of various green and senescent monocotyledon and dicotyledon species. The model is based on the transparent plate model developed by Allen et al. (1969), which represents a leaf as one or more absorbing plates with rough parallel Lambertian surfaces. These rough surfaces are specified in order to simplify the light-to-plate interaction, such that incident light is assumed to be partially isotropic. The model uses two types of input parameters: (1) a leaf structural parameter, and (2) leaf biochemical constituents. The leaf structural parameter specifies the number of compact transparent plates; by adding additional plates, a corresponding increase in the average number of air/cell wall interfaces is created (Jacquemoud et al. 2009). In other words, the model attempts to mimic the discontinuities of the internal structure of a leaf, which is made up of spongy mesophyll cells; leaves that contain more spongy mesophyll tissue will produce greater structurally induced reflectance. When the specific absorption coefficients derived from the leaf biochemical contents are added to these transparent plates, the concentrations of the various absorbers will reduce the reflectivity in certain regions of the spectrum (e.g., chlorophyll absorption in the blue and red region), thus generating reflectance spectra typical of vegetation. The PROSPECT version used in the SLC model is based on the specific absorption coefficients of chlorophyll, brown pigments, dry matter, and water and the refractive index of leaf material from PROSPECT-4 (Feret et al. 2008).

13.2.5.1.3 Submodel: 4SAIL2 Canopy Reflectance Model

The canopy reflectance model 4SAIL2 (Verhoef and Bach 2007) is an advanced two-layer version of the SAIL model (Verhoef 1984), which features a number of improvements in comparison with earlier versions. The original SAIL model is an extension of the 1D model developed by Suits (1972), and is the only true turbid medium model, simulating the bidirectional reflectance of turbid medium plant canopies by utilizing four radiative flux equations. Two diffuse fluxes (upward and downward) are considered to model the scattering of radiative fluxes to and from a canopy. A direct solar flux (e.g., sun) is then considered, to determine the amount of radiation that is intercepted by the canopy. The upward and downward fluxes are combined with the direct solar flux to derive a directional flux in the direction of the observer (e.g., sensor). The succeeding version of SAIL was a development by Kuusk (1991), and was appropriately named SAILH since it includes a function to describe the hotspot effect (H) as the ratio of leaf size to canopy height. SAILH would be the basis for future versions of PROSAIL, including GeoSAIL (Verhoef and Bach 2003), SAIL++ (Verhoef 2002), 4SAIL (Verhoef et al. 2007), and 4SAIL2 (Verhoef and Bach 2007). In order to describe vertically inclined heterogeneous canopies that feature vertical leaf color gradients, GeoSAIL was developed using a two-layer structure, which is described by a dissociation factor and fraction brown leaf area. The 4SAIL2 model is an extension of GeoSAIL, inheriting the two-layer structure, in addition to the singularity removal and speed optimization features from SAIL++ (Verhoef and Bach 2007). Thus, 4SAIL2 is the most recent version developed for use with the SLC model, and includes the crown clumping effect, which is a function of vertical crown cover and tree shape factor, described as the ratio of crown diameter to height (Verhoef and Bach 2007). The crown clumping and the 2-layer structure allow heterogeneous scenes with green and brown elements to be better represented, such as in the case of the GNP.

13.2.6 Scenes Evaluated Using the SLC Model

The West Block of the GNP is characteristic of a heterogeneous landscape, defined in this study as a canopy with multiple layers consisting of green vegetation which may include greater than one dominant species, senesced material, litter, and/or exposed soil. In order to test the capacity of the model to estimate leaf chlorophyll content, four quadrats with corresponding field-based canopy hyperspectral measurements were tested (Figure 13.3). Although unusual and not representative of the typical GNP landscape, a homogeneous canopy of one dominant species was selected to validate the ability of the model to simulate a typical green canopy using a grasslands species (Figure 13.3a). As shown in Figure 13.3b, a heterogeneous canopy with a single dominant species interspersed with standing dead vegetation and litter was selected to represent the typical canopies found throughout the GNP. Other areas in the GNP also exhibited exposed soil in addition to green and brown fractions, as in the case of Figures 13.3c and 13.3d, which have more than one dominant species within the quadrats.

13.2.7 LUT Inversion

Many methods have been proposed to address the ill-posed inverse problem, with the LUT inversion approach being one of the more popular ones, owing to its simplicity and efficient computation times. For the inversion, a LUT was created in advance

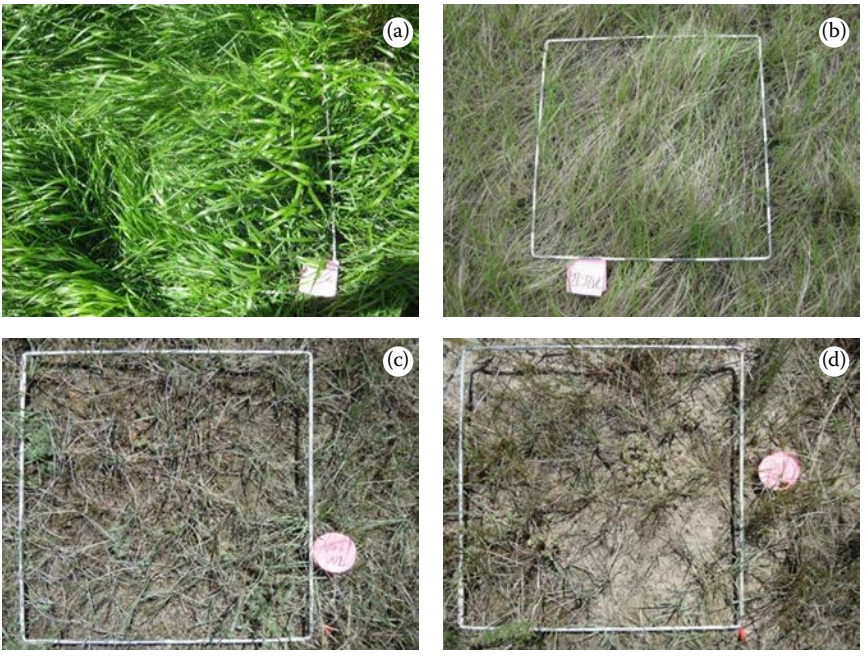


FIGURE 13.3 Scenes used to test and validate the SLC model: (a) green canopy; (b) mixed green and brown canopy; (c, d) mixed green, brown, and soil canopy.

of the inverse of the SLC model through forward calculations using MATLAB and Statistics Toolbox Release 2012b (The MathWorks, Inc., Natick, Massachusetts), which computed all possible parameter combinations from the LUT and generated simulated reflectance curves. From there, the best parameter combination was identified by matching the best fit between measured and simulated reflectance. Only the range between the VIS and NIR (400 to 1300 nm) of the spectra curves were assessed since they correspond to the dominant control of leaf pigments (VIS) and internal leaf structure (NIR), and are thus representative of the discrete variations in vegetation. The criterion used to determine the best-fit spectra as a function of wavelengths was the lowest root-mean-square error (RMSE) criterion, which was calculated using the R statistical software programming language (R Foundation for Statistical Computing, Vienna, Austria):

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(R_{\mathrm{measured}} - R_{\mathrm{LUT}}\right)^2} \tag{13.2}$$

where

R_{measured} = measured reflectance spectra at wavelength λ;
R_{LUT} = modeled reflectance spectra at wavelength λ in the LUT; and
n = total number of wavelengths.

Traditionally, the optimal solution is regarded as the set of input parameters corresponding to the best fit between measured and simulated spectra based on the lowest RMSE. In order to ensure that parameter estimation from the SLC model was robust and accurate, regularization strategies were considered. These strategies included the use of prior knowledge for certain parameter inputs when creating a LUT and exploring the use of multiple solutions because the most optimal solution may not have been unique (Combal et al. 2003). Instead of only reporting the optimal solution, the top five solutions were retrieved. Comparing the top five solutions and ensuring that the corresponding sets of parameter values were realistic provided confidence and reinforced that the solutions were indeed robust.

13.2.8 Sensitivity Analysis

By successively changing the input parameters, a simple sensitivity analysis was performed. This was done by changing one parameter at a time, while keeping all other parameters constant, to determine how much variation was attributable to *x* parameter. Seven variables, including soil moisture, Cab_green (green leaf chlorophylls *a* and *b* content), Cs_brown (brown leaf senescent material concentration), LAI, LIDFu (average leaf inclination), fB (fraction of brown leaf area), and Cv (vertical crown cover percentage) were noted to have substantial control over the shape and/or magnitude of the simulated reflectance curves (Table 13.1). Of course, not all of these variables were applicable to every scene, depending on the fractional coverage of either green, brown, or soil components. For all scenes, simulations were shown to be extremely sensitive to soil moisture, such that it had substantial control over the overall shape and

TABLE 13.1
Specific Ranges for All Input Parameter Values Used for Generating the Final LUT Results for the SLC Model Simulations

Parameter	Abbreviation in Model	Unit	Minimum	Maximum
Soil moisture	Soil moisture	Unitless	0.1	0.3
Green leaf chlorophyll content	Cab_green	$\mu g/cm^2$	50	60
Green leaf water content	Cw_green	g/cm^2	0.002	0.008
Green leaf dry matter content	Cdm_green	g/cm^2	0.002	0.002
Green leaf senescent material concentration (brown pigments)	Cs_green	Unitless	0	0
Green leaf structural parameter	N_green	Unitless	1.5	1.7
Brown leaf chlorophyll content	Cab_brown	$\mu g/cm^2$	0	3
Brown leaf water content	Cw_brown	g/cm^2	0	0.002
Brown leaf dry matter content	Cdm_brown	g/cm^2	0	0.008
Brown leaf senescent material concentration (brown pigments)	Cs_brown	Unitless	0	0.9
Brown leaf structural parameter	N_brown	Unitless	0	1.7
Leaf area index	LAI	m^2/m^2	0.37	9
Leaf distribution function parameter: controls average leaf inclination	LIDFu	Unitless	−1	−0.1
Leaf distribution function parameter: controls bimodality of distribution	LIDFv	Unitless	0	0
Hot spot size parameter, ~ leaf width / canopy height	hot	Unitless	0.05	0.05
Fraction of brown leaf area	fB	Unitless	0	0.65
Layer dissociation factor	Diss	Unitless	0	1
Clumping effect—vertical crown cover percentage	Cv	Unitless	0.3	1
Clumping effect—tree shape factor	zeta	Unitless	0.2	0.2
Solar zenith angle	tts	Degrees	25.9	30.15
Observer/viewing zenith angle	tto	Degrees	0	0
Relative azimuth angle	psi	Degrees	0	0

magnitude of the simulated reflectance curves. Hence, it was determined that a scene should have at least 25% exposed soil coverage for the soil moisture parameter to even be considered. Cab_green was important in explaining the majority of the reflectance variations in the VIS (e.g., green peak with the blue and red absorption feature of green leaves). In the VIS and NIR, LAI and LIDFu are understood to be the most important contributors to reflectance variation (Bacour et al. 2002a). It was determined that LAI is a key parameter for shaping and controlling the magnitude of curves, whereas the Cv and LIDFu parameters only had a magnitude effect. For mixed green and brown scenes, LAI was particularly important for controlling the red absorption feature of green leaves in the VIS, which would otherwise be obscured by the high-reflectance contribution in the same region by Cs_brown and fB. Since the regions associated with leaf pigment and internal structure control were of primary interest, it is unknown how important the contribution of Cw (leaf water content) was in the SWIR region because this was not tested; however, it has been reported to be the most important driver in this region (Bacour et al. 2002a). In sum, the appropriate proportion of each input specified within realistic ranges would produce the most optimal result.

Indeed, more robust sensitivity analysis methods that use rigorous statistical techniques have been successfully tested with PROSAIL, such as the design of experiments for simulation method (DOES) (Bacour et al. 2002b) or the extended Fourier amplitude sensitivity test (EFAST) (Bowyer and Danson 2004). Such methods are able to identify the appropriate parameters that explain the most observed variability in a simulated reflectance field at any wavelength and/or viewing direction (Jacquemoud et al. 2009). However, it was not in the scope of this study to perform such an analysis.

13.2.9 Parameterizing the LUT

Multiple LUTs were created for each scene using the SLC model. It would have been impractical and computationally expensive to generate only a single LUT to account for all of the scenes due to the extreme contrasting differences in green and/or brown leaf and/or soil fractions. In total, 212,885 parameter combinations were generated. Not reported are the multitude of parameter combinations generated from numerous LUTs that were built for testing and sensitivity analysis of the SLC model. The ranges of values (minimum and maximum) for each model parameter are presented in Table 13.1.

Certain parameters, including soil moisture, Cab_green, LAI, LIDFu, LIDFv (bimodality of distribution), hot (hot spot size parameter), fB, Cv, tts (solar zenith angle), tto (observer/viewing zenith angle), and psi (relative azimuth angle), could be effectively estimated or fixed based on prior knowledge from field measurements. Initial parameterization using chlorophyll values below 50 $\mu g/cm^2$ generally produced poor results for mixed scenes. As a result, the range of Cab_green values were constrained to high chlorophyll values. The dissociation parameter (Diss) was given a value of 0–1, based on the scene (e.g., Diss = 1 when scenes approximated having all green leaves in the top layer and all brown leaves in the bottom layer; Diss = 0 for scenes with a homogeneous mixture of green and brown leaves). The tree shape factor (zeta) was fixed as the default value given by the model, which was purported by Verhoef and Bach (2007) to have minimal impact on the modeled results. The

other input parameters, including Cw_green and Cw_brown (green and brown leaf water content, respectively), Cdm_green and Cdm_brown (green and brown leaf dry matter content, respectively), Cs_green and Cs_brown (green and brown leaf senescent material, respectively), and N_green and N_brown (green and brown structural parameter, respectively), were fixed to nominal values. As specified by Darvishzadeh et al. (2008), the N parameter was set at a value of $\geq$1.5 since grasses have thin leaves. The range of all input parameters were in agreement with existing literature (e.g., Atzberger et al. 2003; Atzberger 2004; Chaurasia and Dadhwal 2004; Haboudane et al. 2004; le Maire et al. 2004; Schlerf and Atzberger 2006; Houborg et al. 2007; Darvishzadeh et al. 2008; Laurent et al. 2011a,b; Clevers and Kooistra 2012).

13.3 RESULTS

The top five best-fitting simulated spectral reflectance curves were compared against the measured curves in accordance with the lowest RMSE criterion for four different canopies (Figure 13.4). Site information for each scene, with the corresponding top five RMSE solutions and values of the sets of parameters that had predominant control of the shape and magnitude of the simulated reflectance curves, are presented in Table 13.2.

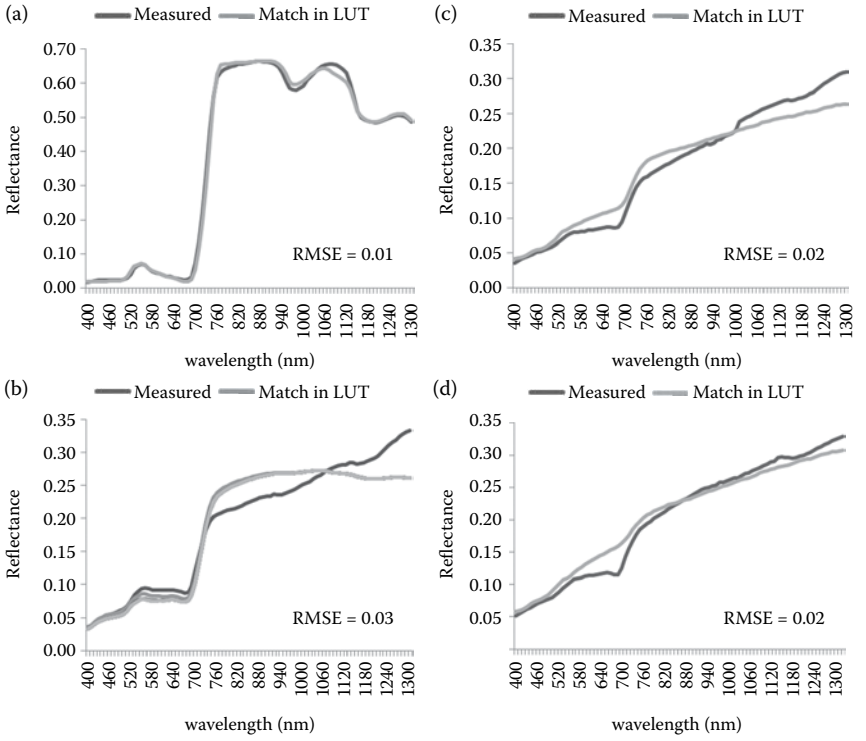


FIGURE 13.4 Top five best-fitting simulated spectra curves, based on the lowest RMSE criterion: (a) green canopy; (b) mixed green and brown canopy; (c, d) mixed green, brown, and soil canopy.

TABLE 13.2
RMSE between the Top Five Best-Fitting Simulated and Measured Spectra Curves, with Estimates for the Dominant Controlling Variables

Scene	RMSE	Soil Moisture	Cab_green (μg/cm²)	Cs_brown	LAI	LIDFu	fB	Cv
Pure green (Figures 13.3a and 13.4a)	0.01	N/A	50	0	9	−0.7	0	1
	0.01	N/A	51	0	9	−0.7	0	1
	0.01	N/A	52	0	9	−0.7	0	1
	0.01	N/A	53	0	9	−0.7	0	1
	0.01	N/A	54	0	9	−0.7	0	1
Mixed green with brown—result (Figures 13.3b and 13.4b)	0.03	N/A	57	0.4	2	−0.7	0.58	1
	0.03	N/A	57	0.5	2	−0.7	0.58	1
	0.03	N/A	58	0.5	2	−0.7	0.58	1
	0.03	N/A	59	0.5	2	−0.7	0.58	1
	0.03	N/A	60	0.5	2	−0.7	0.58	1
Mixed green with brown and exposed soil—Scene 1 (Figures 13.3c and 13.4c)	0.02	0.2	58	0.6	0.58	−0.8	0.15	0.6
	0.02	0.2	59	0.6	0.58	−0.8	0.15	0.6
	0.02	0.2	59	0.6	0.58	−0.8	0.15	0.6
	0.02	0.2	60	0.6	0.58	−0.8	0.15	0.6
	0.02	0.2	60	0.6	0.58	−0.8	0.15	0.6
Mixed green with brown and exposed soil—Scene 2 (Figures 13.3d and 13.4d)	0.02	0.1	59	0.6	0.37	−0.8	0.1	0.4
	0.02	0.1	60	0.6	0.37	−0.8	0.1	0.4
	0.02	0.1	60	0.6	0.37	−0.8	0.1	0.4
	0.02	0.1	60	0.6	0.37	−0.8	0.1	0.4
	0.02	0.1	60	0.6	0.37	−0.8	0.1	0.4

The first scene tested was a green homogeneous canopy (Figure 13.3a). A modified soil spectrum with zero reflectance was used for this scene since there was no exposed soil present. The best-fitting simulated spectra were in agreement with the measured spectra given by the RMSE of 0.01 (Figure 13.4a). The corresponding input parameter values, such as Cab_green (50 to 54 μg/cm^2), LIDFu (−0.7), and Cv (1), were comparable with *in situ* measured values, but LAI (9) was overestimated. The fit of the simulated and measured spectra demonstrated that the PROSPECT model was robust to simulate native grass species from the GNP, even though the database did not contain any biochemical data corresponding to any of the native grass species from the GNP. However, the scene only had one native grass species, and thus the result should be taken with caution (e.g., the SLC model may not fit other native GNP species equally well).

A heterogeneous canopy, consisting of both green and brown fractions (standing dead vegetation and/or litter), was tested (Figure 13.3b). A LUT was built using a zero-reflectance soil spectrum; the simulated curves had relatively good fit (RMSE = 0.03), as shown in Figure 13.4b, which was able to capture the red absorption feature of the green vegetation in the scene. However, even when all of the input parameter values were specified within realistic ranges, including Cs_brown (0.4 to 0.5), LAI (2), and LIDFu (−0.7)—with the exception of leaf chlorophyll content being overestimated (57 to 60 μg/cm^2)—the SLC model was not able to capture the discrete variations in the NIR region of the measured reflectance curve. The best-fitting simulated curves demonstrated plateauing in the NIR, with the reflectance gradually increasing in the NIR of the measured curve, indicating that the SLC model could not correctly model brown material.

In order to determine the utility of the SLC model for heterogeneous scenes with fractional coverage of green and brown in addition to having a substantial component of exposed soil, two canopies were selected, with one having greater exposed soil (Figure 13.3d) than the other (Figure 13.3c). The best-fitting simulated curves for both canopies had great fit (RMSE = 0.02), with corresponding parameter values that were within realistic ranges of values and/or comparable with *in situ* measured values, including soil moisture (0.2 in Figure 13.4c and 0.1 in Figure 13.4d), LAI (0.58 in Figure 13.4c and 0.37 in Figure 13.4d), LIDFu (−0.8 in Figure 13.4c and −0.8 in Figure 13.4d), fB (0.15 in Figure 13.4c and 0.1 in Figure 13.4d), and Cv (0.6 in Figure 13.4c and 0.4 in Figure 13.4d). For Figure 13.4c, the SLC model was able to capture the chlorophyll absorption with good accuracy, but in the NIR the trajectory of the curve fell short of the measured curve, indicating that the combination of the brown and soil fraction was not being modeled correctly. On the other hand, the scene with greater exposed soil (Figure 13.4d) captured the reflectance in the NIR accurately, which was attributable to less brown coverage and greater soil influence, but failed to represent the green fraction correctly by not simulating the chlorophyll absorption. Like the mixed green and brown scene, the chlorophyll content estimates (58 to 60 μg/cm^2) are likely overestimated and inaccurate for both soil dominated canopies.

13.4 DISCUSSION

The results from this study have demonstrated the utility of the SLC model and its application for heterogeneous canopies with mixed fractions of green, brown, and soil fractions. Leaf chlorophyll content was accurately estimated for a green canopy,

but overestimated for mixed scenes with brown and exposed soil fractions. It was determined during the initial parameterization of the LUTs that if high green leaf chlorophyll content values were not specified for mixed scenes, the red absorption feature (as observed in the measured curves) would be masked and overcompensated for by the reflectance contribution from the brown and/or soil components in the simulated curves. This was likely an issue caused by the SLC model not being sensitive enough to detect the contribution of the discrete variation in low chlorophyll content values. Hence, the chlorophyll content values from the best-fitting spectra for the mixed scenes were likely overestimated. Even so, all results were determined to be highly accurate based on the lowest RMSE criterion between the best-fitting simulated and measured curves (e.g., RMSE = 0.01–0.16). Consequently, the use of the lowest RMSE criterion only considers the average error between each predicted (simulated) and observed (measured) wavelength value. In this respect, the lowest RMSE criterion may potentially match a simulated curve that is based on unrealistic parameter combinations with a measured curve. However, owing to the use of prior knowledge from the extensive collection of field measurements, this issue was avoided because all of the best-fitting matches had realistic parameter estimates. Hence, even when the simulated curves that represented the mixed scenes were generated using realistic parameter values and compensated for using high leaf chlorophyll content values, they still could not accurately match the measured curves, displaying discrete reflectance differences in both the VIS and NIR. This problem was likely caused by the inability of the SLC model to correctly model heterogeneous canopies.

The SLC model cannot accurately simulate heterogeneous scenes with green, brown, and/or soil fractions, and will require a recalibration based on *in situ* measurements from the GNP. Based on the findings for the green homogeneous canopy (Figure 13.4a) and the mixed green and brown canopy (Figure 13.4b), the SLC model was capable of capturing the green reflectance and red absorption features of green vegetation. However, for the mixed green and brown canopy (Figure 13.4b), the reflectance in the NIR could not be correctly modeled, indicating that the issue was likely an inaccurate representation of brown pigments via the PROSPECT model. In the same way, the two mixed scenes (Figure 13.4c and d) with exposed soil also demonstrated that the soil spectrums of the Hapke model were not fully representative of GNP soil types. Accordingly, the PROSPECT model will need to be updated to include biochemical compound and spectral reflectance data of brown leaves (both standing dead vegetation and litter) from the GNP. Also of importance, but not as essential as the brown pigment issue, would be the inclusion of GNP species data into the PROSPECT species database. Furthermore, the Hapke model will require soil spectral reflectance data from the GNP to be able to correctly simulate the different soil types within the GNP. Despite the shortcomings of the PROSPECT and Hapke models, the 4SAIL2 canopy model accurately represented canopy heterogeneities, such as vertical leaf color gradients (i.e., fB and Diss) and clumping effects (i.e., Cv and zeta). The 4SAIL2 model accurately explained the canopy structure (i.e., leaf angle distribution) governed by LIDFu and LIDFv, and canopy biophysical parameters such as Cv for all scenes and fB for mixed scenes. However, the 4SAIL2 model may have issues accurately assessing LAI for very dense grass canopies (e.g., overestimated for Figure 13.4a) since the retrieved LAI estimates for the mixed scenes were reasonably accurate.

A study by Darvishzadeh et al. (2008) on estimating LAI and chlorophyll content from green heterogeneous grassland canopies through the inversion of PROSAIL determined that the accuracy for estimating biochemical and biophysical variables decreased as the number of species in a scene increased. Presumably, the issue with PROSAIL was a result of the SAILH canopy model being unable to address vertically inclined canopies characteristic of grass species. This discrepancy would later be addressed by the 4SAIL2 model, as used in the SLC model. Consequently, it is unknown whether the SLC model truly rectifies parameter retrieval for vertically inclined green canopies featuring multiple species because a green canopy with more than one dominant species could not be tested in this study. The closest representation were the results from the soil-dominated scenes that had multiple species. Even when accounting for additional heterogeneous discontinuities, such as multiple layers (green, brown, soil components) with multiple species, the parameter retrieval was highly accurate, with the exception of leaf chlorophyll content. Hence, the effectiveness of the SLC model for canopies that have multiple layers and species is excellent, despite the shortcomings discussed previously. To this end, Darvishzadeh et al. (2008) suggested that 3D models may be able to more accurately portray heterogeneous grassland canopies. However, since 3D models are conceptually and computationally complex and require a high number of input parameters, this option is not practical, especially for operational monitoring programs. Darvishzadeh et al. also suggested that the then recently introduced SLC model could provide improved parameter retrieval, which, in the case of this study, has been tested and validated with success.

The SLC model has been successfully validated for homogeneous (e.g., Laurent et al. 2011a,b) and heterogeneous (e.g., Hernández-Clemente et al. 2012) forest canopies, homogeneous precision agriculture canopies (Migdall et al. 2009), and homogeneous grassland canopies (e.g., Dasgupta and Qu 2009; Dasgupta et al. 2009). The literature review revealed no previous studies that had attempted the estimation of vegetation biochemical and/or biophysical variables from heterogeneous canopies with mixed proportions of green and brown, with or without soil fractions. Any studies pertaining to the use of RTMs for retrieval of biochemical and/or physical parameters from heterogeneous canopies (e.g., Darvishzadeh et al. 2008; Vohland and Jarmer 2008; Si et al. 2012; Hernández-Clemente et al. 2012) have only been done for green canopies without the influence of brown material. However, of the studies that used the SLC model in the inverse mode to estimate biochemical and/or biophysical variables, two did retrieve green and brown parameter estimates from homogeneous canopies. For instance, a homogeneous canopy of winter wheat was assessed by Migdall et al. (2009) in a precision agriculture study. They successfully employed the SLC model to estimate green LAI and fraction of senescent material. The other study was an assessment of three homogeneous Norway spruce stands by Laurent et al. (2011a). They successfully used the SLC model to retrieve accurate biochemical and biophysical parameters, such as green leaf chlorophyll content and plant area index, respectively, as well as estimates for bark material. In the case of both studies, the authors were likely able to estimate brown leaves (senescing leaves and bark material) accurately due to the PROSPECT database being developed with biochemical and reflectance data of brown leaves that were more representative of forest and agricultural crop species, which evidently was not characteristic of brown material in the GNP.

13.5 CONCLUSION

This study tested the utility of the SLC model to estimate chlorophyll content in the inverse mode for heterogeneous canopies characteristic of a mixed-grass prairie ecosystem. The use of prior knowledge greatly improved processing times and helped circumvent the ill-posed inversion problem. By using the lowest RMSE criterion, leaf chlorophyll content was accurately estimated for a green canopy (50 to 54 $\mu g/cm^2$), but overestimated for mixed canopies with green and brown fractions (57 to 60 $\mu g/cm^2$) and for mixed canopies with green, brown, and soil fractions (58 to 60 $\mu g/cm^2$). Canopy biophysical parameters, such as Cv for all scenes and fB for mixed scenes, were accurately estimated. The results confirmed the potential of model inversion for estimating leaf biochemical parameters and canopy biophysical parameters for heterogeneous canopies with multiple layers and species using hyperspectral measurements.

The SLC model will require refinements to be able to accurately assess and estimate the radiation field for heterogeneous canopies with multiple layers (e.g., green, brown, and soil components) and multiple species. Results from this study demonstrated that the SLC model was able to accurately simulate a homogeneous green canopy in terms of leaf biochemical parameters. Accurate canopy biophysical parameter estimates were also retrieved, with the exception of LAI, which indicated that the SLC model may have issues simulating very dense green canopies. On the other hand, when the SLC model was tested for heterogeneous canopies with multiple layers and species, it had issues capturing the discrete reflectance variations when compared to the measured reflectance data. Nonetheless, the parameter retrieval was highly accurate in terms of canopy biophysical parameters for mixed scenes, but demonstrated an overestimation of leaf biochemicals such as leaf chlorophyll content. The limitation of the SLC model is most likely a result of the model not being specifically calibrated for GNP species (specifically brown pigments) and soil types.

Any future work with the SLC model for heterogeneous grassland canopies with green, brown, and/or soil fractions will need to consider updating the PROSPECT and Hapke models. Once the appropriate changes are implemented, improvements in parameter retrieval should be expected, especially for leaf chlorophyll content estimation; thus, the use of the SLC model for temporal work can in principle be applied to different remote sensing data over the GNP and similar heterogeneous environments. Recent progress in the modification of the SLC model has involved the recalibration of the stand-alone PROSPECT model; work on the Hapke model has yet to be addressed. A preliminary examination of the PROSPECT model revealed poor performance in simulating leaf spectra of senescing and senesced leaves. This was expected because the PROSPECT model was calibrated using a green leaf data set (Feret et al. 2008). Since green and senesced leaves have different biophysical and biochemical properties, a new set of parameters needs to be considered for simulating spectra of senesced leaves in the PROSPECT model. For instance, chlorophyll or water content is very minimal in senesced leaves. However, senesced leaves have biochemical components that do not exist in green leaves, such as humic and fulvic acid. For senesced leaves, the biochemical parameters that are considered in calibrating PROSPECT are initial and advanced decay parameters, including humic and fulvic acid, brown pigments, water, and dry matter.

The recalibration of PROSPECT has included the collection of leaf samples of 33 monocot species (e.g., grasses, forbs, and crops) at five senescence stages (e.g., from freshly senesced to fully decayed) as experimental material. Half of the samples were used to recalibrate the model, and the other half were used for validation. Reflectance and transmittance of all leaf samples were measured using an integrating sphere equipped with an ASD spectroradiometer. The biochemical components were analyzed following the protocol of Van Soest et al. (1991) and Sparks et al. (1996). After acquiring the spectra and biochemical components of leaf samples, the model recalibration was conducted based on the procedure proposed by Feret et al. (2008). The recalibration of PROSPECT was divided into two steps: (1) calibrating the leaf thickness for each sample, and (2) determining the specific absorption coefficients for each chemical component and quantifying the refraction index. After the PROSPECT model is recalibrated, it will be integrated into the SLC model, which will then be evaluated for heterogeneous canopies.

REFERENCES

Allen, W. A., Gausman, H. W., Richardson, A. J., and Thomas, J. R. 1969. Interaction of isotropic light with a compact plant leaf. *Journal of the Optical Society of America,* 59, 1376–1379.

Atzberger, C. 2004. Object-based retrieval of biophysical canopy variables using artificial neural nets and radiative transfer models. *Remote Sensing of Environment,* 93(1–2), 53–67.

Atzberger, C., Jarmer, T., Schlerf, M., Kötz, B., and Werner, W. 2003. Retrieval of wheat biophysical attributes from hyperspectral data and SAILH + PROSPECT radiative transfer model. In M. Habermeyer, A. Müller, and S. Holzwarth (Eds.), *Proc. 3rd EARSeL Workshop on Imaging Spectroscopy*, Herrsching, Germany, 473–482.

Bach, H., and Mauser, W. 1994. Modelling and model verification of the spectral reflectance of soils under varying moisture conditions. In *Proc. IGARSS '94 Symposium*, Pasadena, CA, 4, 2354–2356.

Bacour, C., Baret, F., and Jacquemoud, S. 2002a. Information content of HyMap hyperspectral imagery. In *Proc. 1st International Symposium on Recent Advances in Quantitative Remote Sensing*, Valencia, Spain, 503–508.

Bacour, C., Jacquemoud, S., Tourbier, Y., Dechambre, M., and Frangi, J. P. 2002b. Design and analysis of numerical experiments to compare four canopy reflectance models. *Remote Sensing of Environment,* 79(1), 72–83.

Baret, F., and Buis, S. 2008. Estimating canopy characteristics from remote sensing observations: Review of methods and associated problems. In *Advances in Land Remote Sensing*, pp. 173–201. Springer.

Baret, F., and Guyot, G. 1991. Potentials and limits of vegetation indices for LAI and APAR assessment. *Remote Sensing of Environment,* 35(2–3), 161–173.

Black, S. C., and Guo, X. 2008. Estimation of grassland CO_2 exchange rates using hyperspectral remote sensing techniques. *International Journal of Remote Sensing,* 29(1), 145–155. doi:10.1080/01431160701253220

Blackburn, G. A. 2007. Hyperspectral remote sensing of plant pigments. *Journal of Experimental Botany,* 58(4), 855–867.

Bowyer, P., and Danson, F. M. 2004. Sensitivity of spectral reflectance to variation in live fuel moisture content at leaf and canopy level. *Remote Sensing of Environment,* 92(3), 297–308.

Broge, N. H., and Leblanc, E. 2001. Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. *Remote Sensing of Environment,* 76(2), 156–172.

Broge, N. H., and Mortensen, J. V. 2002. Deriving green crop area index and canopy chlorophyll density of winter wheat from spectral reflectance data. *Remote Sensing of Environment,* 81(1), 45–57.

Ceccato, P., Flasse, S., and Gregoire, J. 2002. Designing a spectral index to estimate vegetation water content from remote sensing data: Part 2. Validation and application. *Remote Sensing of Environment,* 82, 198–207.

Chaurasia, S., and Dadhwal, V. K. 2004. Comparison of principal component inversion with VI-empirical approach for LAI estimation using simulated reflectance data. *International Journal Remote Sensing,* 25(14), 2881–2887.

Chen, J. M., Li, X., Nilson, T., and Strahler, A. 2000. Recent advances in geometrical optical modelling and its applications. *Remote Sensing Reviews,* 18(2–4), 227–262.

Clevers, J. G. P. W., and Kooistra, L. 2012. Using hyperspectral remote sensing data for retrieving canopy chlorophyll and nitrogen content. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* 5(2), 574–583.

Combal, B., Baret, F., Weiss, M., Trubuil, A., Macé, D., Pragnère, A., Mynenic, R., Knyazikhin, Y., and Wang, L. 2003. Retrieval of canopy biophysical variables from bidirectional reflectance using prior information to solve the ill-posed inverse problem. *Remote Sensing of Environment,* 84(1), 1–15.

Csillag, F., Kertesz, M., Davidson, A., and Mitchell, S. 2001. On the measurement of diversity-productivity relationships in a northern mixed grass prairie (Grasslands National Park, Saskatchewan, Canada). *Community Ecology,* 2, 145–159.

Curran, P. J., Dungan, J. L., and Gholz, H. L. 1990. Exploring the relationship between reflectance red-edge and chlorophyll content in slash pine. *Tree Physiology*, 7, 33–48.

Darvishzadeh, R., Skidmore, A., Schlerf, M., and Atzberger, C. 2008. Inversion of a radiative transfer model for estimating vegetation LAI and chlorophyll in a heterogeneous grassland. *Remote Sensing of Environment,* 112(5), 2592–2604.

Dasgupta, S., and Qu, J. J. 2009. Soil adjusted vegetation water content retrievals in grasslands. *International Journal of Remote Sensing,* 30(4), 1019–1043.

Dasgupta, S., Qu, J. J., and Bhoi, S. 2009. Constrained radiative transfer inversions for vegetation moisture retrievals in grasslands. *Journal of Applied Remote Sensing,* 3(1), 031503.

Durbha, S. S., King, R. L., and Younan, N. H. 2007. Support vector machines regression for retrieval of leaf area index from multiangle imaging spectroradiometer. *Remote Sensing of Environment,* 107(1–2), 348–361.

Feret, J.-B., François, C., Asner, G. P., Gitelson, A. A., Martin, R. E., Bidel, L. P. R., Ustin, S. L., le Maire, G., and Jacquemoud, S. 2008. PROSPECT-4 and 5: Advances in the leaf optical properties model separating photosynthetic pigments. *Remote Sensing of Environment,* 112(6), 3030–3043.

Filella, I., Serrano, L., Serra, J., and Peñuelas, J. 1995. Evaluating wheat nitrogen status with canopy reflectance indices and discriminant analysis. *Crop Science,* 35(5), 1400–1405.

Gastellu-Etchegorry, J. P., Gascon, F., and Estève, P. 2003. An interpolation procedure for generalizing a look-up table inversion method. *Remote Sensing of Environment,* 87(1), 55–71.

Gitelson, A., and Merzlyak, M. N. 1994a. Quantitative estimation of chlorophyll-a using reflectance spectra: Experiments with autumn chestnut and maple leaves. *Journal of Photochemistry and Photobiology, B: Biology,* 22(3), 247–252.

Gitelson, A. A., and Merzlyak, M. N. 1994b. Spectral reflectance changes associated with autumn senescence of *Aesculus hippocastanum* L. and *Acer platanoides* L. leaves. Spectral features and relation to chlorophyll estimation. *Journal of Plant Physiology,* 143, 286–292.

Goel, N. S., and Grier, T. 1988. Estimation of canopy parameters for inhomogeneous vegetation canopies from reflectance data: III. Trim: A model for radiative transfer in heterogeneous three-dimensional canopies. *Remote Sensing of Environment,* 25(3), 255–293.

Haboudane, D., Miller, J. R., Pattey, E., Zarco-Tejada, P. J., and Strachan, I. B. 2004. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modelling and validation in the context of precision agriculture. *Remote Sensing of Environment,* 90(3), 337–352.

Haboudane, D., Tremblay, N., Miller, J. R., and Vigneault, P. 2008. Remote estimation of crop chlorophyll content using spectral indices derived from hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing,* 46(2), 423–436.

Hapke, B. 1981. Bidirectional reflectance spectroscopy. 1. Theory. *Journal of Geophysical Research,* 86, 3039–3054.

Hapke, B., and Wells, E. 1981. Bidirectional reflectance spectroscopy. 2. Experiments and observations. *Journal of Geophysical Research,* 86, 3055–3060.

Hernández-Clemente, R., Navarro-Cerrillo, R. M., and Zarco-Tejada, P. J. 2012. Carotenoid content estimation in a heterogeneous conifer forest using narrow-band indices and PROSPECT + DART simulations. *Remote Sensing of Environment,* 127, 298–315.

Houborg, R., Soegaard, H., and Boegh, E. 2007. Combining vegetation index and model inversion methods for the extraction of key vegetation biophysical parameters using terra and aqua MODIS reflectance data. *Remote Sensing of Environment,* 106(1), 39–58.

Jacquemoud, S., Bacour, C., Poilvé, H., and Frangi, J.-P. 2000. Comparison of four radiative transfer models to simulate plant canopies reflectance: Direct and inverse mode. *Remote Sensing of Environment,* 74(3), 471–481.

Jacquemoud, S., and Baret, F. 1990. PROSPECT: A model of leaf optical properties spectra. *Remote Sensing of Environment,* 34(2), 75–91.

Jacquemoud, S., Baret, F., Andrieu, B., Danson, F. M., and Jaggard, K. 1995. Extraction of vegetation biophysical parameters by inversion of the PROSPECT + SAIL models on sugar beet canopy reflectance data application to TM and AVIRIS sensors. *Remote Sensing of Environment,* 52(3), 163–172.

Jacquemoud, S., and Ustin, S. L. 2008. Modeling of leaf optical properties. http://www.photobiology.info/Jacq_Ustin.html (December 5, 2012).

Jacquemoud, S., Verhoef, W., Baret, F., Bacour, C., Zarco-Tejada, P. J., Asner, G. P., François, C., and Ustin, S. L. 2009. PROSPECT + SAIL models: A review of use for vegetation characterization. *Remote Sensing of Environment,* 113(Suppl. 1), S56–S66.

Kimes, D., Gastellu-Etchegorry, J., and Estève, P. 2002. Recovery of forest canopy characteristics through inversion of a complex 3D model. *Remote Sensing of Environment,* 79(2–3), 320–328.

Kötz, B., Schaepman, M., Morsdorf, F., Bowyer, P., Itten, K., and Allgöwer, B. 2004. Radiative transfer modelling within a heterogeneous canopy for estimation of forest fire fuel properties. *Remote Sensing of Environment,* 92(3), 332–344.

Kuusk, A. 1991. The hot-spot effect in plant canopy reflectance. In R. B. Myneni and J. Ross (Eds.), *Photon Vegetation Interaction*, pp. 140–159. Berlin: Springer-Verlag.

Laurent, V. C. E., Verhoef, W., Clevers, J. G. P. W., and Schaepman, M. E. 2011a. Estimating forest variables from top-of-atmosphere radiance satellite measurements using coupled radiative transfer models. *Remote Sensing of Environment,* 115(4), 1043–1052.

Laurent, V. C. E., Verhoef, W., Clevers, J. G. P. W., and Schaepman, M. E. 2011b. Inversion of a coupled canopy-atmosphere model using multi-angular top-of-atmosphere radiance data: A forest case study. *Remote Sensing of Environment,* 115(10), 2603–2612.

le Maire, G., François, C., and Dufrêne, E. 2004. Towards universal broad leaf chlorophyll indices using PROSPECT simulated database and hyperspectral reflectance measurements. *Remote Sensing of Environment,* 89, 1–28.

Lichtenthaler, H. K. 1987. Chlorophylls and carotenoids: Pigments of photosynthetic biomembranes.. *Methods Enzymology*, 148, 350–382.

Lu, B., He, Y., and Tong, A. 2016. Evaluation of spectral indices for estimating burn severity in semiarid grasslands. *International Journal of Wildland Fire,* 25, 147–157.

Meroni, M., Colombo, R., and Panigada, C. 2004. Inversion of a radiative transfer model with hyperspectral observations for LAI mapping in poplar plantations. *Remote Sensing of Environment,* 92(2), 195–206.

Merzlyak, M. N., Gitelson, A. A., Chivkunova, O. B., and Rakitin, V. Y. 1999. Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening. *Physiologia Plantarum,* 106(1), 135–141.

Migdall, S., Bach, H., Bobert, J., Wehrhan, M., and Mauser, W. 2009. Inversion of a canopy reflectance model using hyperspectral imagery for monitoring wheat growth and estimating yield. *Precision Agriculture,* 10(6), 508–524.

Minocha, R., Martinez, G., Lyons, B., and Long, S. 2009. Development of a standardized methodology for quantifying total chlorophyll and carotenoids from foliage of hardwood and conifer tree species. *Canadian Journal of Forest Research,* 39(4), 849–861.

Moran, J. A., Mitchell, A. K., Goodmanson, G., and Stockburger, K. A. 2000. Differentiation among effects of nitrogen fertilization treatments on conifer seedlings by foliar reflectance: A comparison of methods. *Tree Physiology,* 20(16), 1113–1120.

Peñuelas, J., and Filella, L. 1998. Technical focus: Visible and near-infrared reflectance techniques for diagnosing plant physiological status. *Trends in Plant Science,* 3(4), 151–156.

Schlerf, M., and Atzberger, C. 2006. Inversion of a forest reflectance model to estimate structural canopy variables from hyperspectral remote sensing data. *Remote Sensing of Environment,* 100(3), 281–294.

Si, Y., Schlerf, M., Zurita-Milla, R., Skidmore, A., and Wang, T. 2012. Mapping spatio-temporal variation of grassland quantity and quality using MERIS data and the PROSAIL model. *Remote Sensing of Environment,* 121, 415–425.

Sparks, D. L., Page, A. L., Helmke, P. A., Loeppert, R. H., Soltanpour, P. N., Tabatabai, M. A., Johnston, C. T., and Sumner, M. E. et al. 1996. *Methods of Soil Analysis. Part 3—Chemical Methods.* Soil Science Society of America Inc.

Suits, G. H. 1972. The calculation of the directional reflectance of a vegetative canopy. *Remote Sensing of Environment,* 2, 117–125.

Tong, A. 2014. *Estimating Grassland Chlorophyll Content for a Mixed Grassland: Exploring the Performance of the Empirical-Statistical and the Physical Modeling Approach*. Master's thesis, University of Toronto, Toronto, Canada.

Van Soest, P. J., Robertson, J. B., and Lewis, B. A. 1991. Methods for dietary fiber neutral detergent fiber and nonstarch polysaccharides in relation to animal nutrition. *Journal of Dairy Science,* 74, 3583–3597.

Verhoef, W. 1984. Light scattering by leaf layers with application to canopy reflectance modelling: The SAIL model. *Remote Sensing of Environment,* 16(2), 125–141.

Verhoef, W. 2002. Improved modelling of multiple scattering in leaf canopies: The model SAIL++. In J. A. Sobrino (Ed.), *Proc. 1st International Symposium on Recent Advances in Quantitative Remote Sensing*, Universitat de Valencia, Valencia, Spain, 11–20.

Verhoef, W., and Bach, H. 2003. Simulation of hyperspectral and directional radiance images using coupled biophysical and atmospheric radiative transfer models. *Remote Sensing of Environment,* 87(1), 23–41.

Verhoef, W., and Bach, H. 2007. Coupled soil-leaf-canopy and atmosphere radiative transfer modelling to simulate hyperspectral multi-angular surface reflectance and TOA radiance data. *Remote Sensing of Environment,* 109(2), 166–182.

Verhoef, W., Jia, L., Xiao, Q., and Su, Z. 2007. Unified optical-thermal four-stream radiative transfer theory for homogeneous vegetation canopies. *IEEE Transactions on Geoscience and Remote Sensing,* 45(6), 1808–1822.

Vohland, M., and Jarmer, T. 2008. Estimating structural and biochemical parameters for grassland from spectroradiometer data by radiative transfer modelling (PROSPECT + SAIL). *International Journal of Remote Sensing,* 29(1), 191–209.

Walthall, C., Dulaney, W., Anderson, M., Norman, J., Fang, H., and Liang, S. 2004. A comparison of empirical and neural network approaches for estimating corn and soybean leaf area index from Landsat ETM+ imagery. *Remote Sensing of Environment,* 92(4), 465–474.

Weiss, M., Baret, F., Myneni, R. B., Pragnere, A., and Knyazikhin, Y. 2000. Investigation of a model inversion technique to estimate canopy biophysical variables from spectral and directional reflectance data. *Agronomie,* 20(1), 3–22.

Zarco-Tejada, P. J., Miller, J. R., Mohammed, G. H., Noland, T. L., and Sampson, P. H. 2000. Chlorophyll fluorescence effects on vegetation apparent reflectance: II. Laboratory and airborne canopy-level measurements with hyperspectral data. *Remote Sensing of Environment,* 74(3), 596–608.

Zhang, Y., Chen, J. M., Miller, J. R., and Noland, T. L. 2008. Leaf chlorophyll content retrieval from airborne hyperspectral remote sensing imagery. *Remote Sensing of Environment,* 112(7), 3234–3247.

14 Wetland Detection Using High Spatial Resolution Optical Remote Sensing Imagery

Amy B. Mui

CONTENTS

14.1 INTRODUCTION

The applications of remote sensing research extend to a variety of different ecosystems found across the globe. Wetlands are particularly important despite covering only 4%–6% of Earth's surface, or 5.3 to 12.8 million km^2 (Zedler &

Kercher, 2005). In some areas, they can cover >20% of the landscape, such as the vast peatlands of the boreal region of Canada (Mitsch & Gosselink, 1993). The high rate of biological productivity of wetlands, strong natural selection pressures, and diverse aquatic environments have produced many species of plants and animals that are not found anywhere else (Gibbs, 1993). As both sources and sinks of carbon depending on their age, composition, and surrounding environmental conditions (Kayranli et al., 2010; Roulet 2000), wetlands also play a strong role in global climate projections.

Obvious benefits to human society are provided such as flood abatement, climate change mitigation, and the filtration of ground water (Mitsch & Gosselink, 2000). An estimated 40% of the value of global ecosystem services is provided by wetlands (Zedler, 2003), yet it is estimated that more than half of peatlands, depressional (low-lying) wetlands, riparian zones, lake littoral zones, and floodplains have been lost, with little sign of abatement. (Verhoeven & Setter, 2010; Zedler & Kercher, 2005). Existing studies suggest that humans and human-related activities are the dominant agents of wetland loss and degradation primarily through conversion to agriculture and hydrological modification such as damming and pumping. These practices alter the natural timing of water fluctuations responsible for the diversity of vegetation communities and habitat types found in wetlands (Brock et al., 1999).

A critical step in effective management of declining wetland ecosystems is accurate detection. Wetlands are often located in remote and inaccessible locations, with borders that can be transitional in nature, so identifying and mapping these ecosystems presents a significant challenge. Compared to traditional methods of ground surveying, which require a substantial investment in labor and time, and aerial photography, which is costly and limited by spectral range and local weather conditions, imagery from satellite-borne sensors offers many advantages to the detection of wetlands. Repeat coverage over large spatial extents and across a variable range of the electromagnetic spectrum allows spaceborne remote sensors to provide a wealth of spatially explicit data for the detection of wetland ecosystems.

14.1.1 Challenges and Considerations

Despite the suitability of satellite data for studying wetlands, some characteristics of wetlands may confound a remote-sensing-based approach. The term *wetland* refers to multiple classes of aquatic-adapted ecosystems that vary broadly in topography, biotic communities, and hydrologic form and function. Broad classes include bogs, fens, swamps, marshes, and open shallow water. Yet this diversity of wetland forms does not always manifest as a unique spectral response. Some wetlands also tend to occupy transition zones between purely aquatic and purely terrestrial ecosystems, which presents a challenge when the goal is to define a logical boundary between wetland and upland. Wetlands themselves can also be transitional in nature (i.e., a fen may transition to a bog over time). To aid in the discrimination, plant species are commonly used as an indicator of the presence and class of wetland. The presence of acid-loving *Sphagnum* moss, grasses and sedges is often indicative of a bog or fen, whereas water lilies, reeds, and cattails indicate nutrient-rich marshes or swamps. However, key species used to distinguish between wetland classes may not

always manifest as a unique spectral response, nor will they always be present in homogenous stands large enough to be captured as pure pixels. Rapidly colonizing North American invasive species such as *Phragmites* spp. may also confound accurate wetland detection because they outcompete native plant species. Despite this challenge, it is important for remote-sensing-based detection of wetlands to identify the specific class not solely to achieve a higher classification accuracy, but also because failure to correctly categorize wetlands can have important downstream implications on quantifying their global impact. Bogs and fens, for example, have considerably greater carbon storage potential than marshes and swamps due to the accumulation of belowground peat, and have been found to be a stronger emitter of methane (Moore et al., 1994).

Wetlands can also exist in highly variable shapes and sizes, ranging from the vast Pantanal wetlands in Brazil, which cover more than 150,000 km^2, to the glaciated prairie pothole region of central Canada, where almost 88% of wetlands are less than 0.4 ha in area (Halabisky, 2011). In the context of remote sensing, delineating these small and ephemeral wetlands represents a key challenge (Ozesmi & Bauer, 2002), which is likely why most studies have centered on larger wetlands. However, pools as small as 0.2 ha represent important, often critical, habitat for wetland-dependent wildlife (Semlitsch & Bodie, 1998), and protection of small wetlands is vital for the maintenance of biodiversity (Gibbs, 1993; Semlitsch & Bodie, 1998). The variable size and shape of wetlands requires that a sensor with an appropriate spatial resolution is selected to capture the minimum mapping unit required. Medium- and coarse-resolution sensors such as AVHRR, MODIS, and Landsat may be appropriate for larger wetlands, while high spatial resolution sensors (<5 m) would be required to capture smaller wetlands. Depending on the size and heterogeneity of the target wetland feature relative to the spatial resolution, the mixed pixel problem can become a significant concern in wetland detection due to factors such as transitional borders, variable size, irregular shape, and mixed composition. Studies encompassing all wetland classes must include parameters that capture the range of characteristics found in all wetland classes in the study area (Table 14.1), while research questions that focus on a single wetland type must be able to uniquely identify that single class from among other similar wetland classes.

Another consideration rests on the timing of image acquisition over wetlands. Depending on their hydrologic characteristics, wetlands may be seasonally, episodically, or continuously inundated. Seasonal variation is not unique to wetland ecosystems, but is arguably more pronounced because this rapid variability in hydrologic regime is often the driver for the high biodiversity found in wetlands. Consequently, remote sensing data must be acquired during a narrower window than most terrestrial regions due to the rapid change in both water extent and vegetation cover. Delineation of boundaries often requires leaf-off data from early spring, when water levels are high after the spring melt and boundaries can be visualized. Mapping of aquatic vegetation communities requires imagery acquired later in the growing season.

Despite these challenges, wetlands have been studied using remote sensing data continuously since the availability of the first satellite data, and this exploration continues to cover novel aspects of wetland study.

TABLE 14.1
Land Cover Class Descriptions Adapted from Anderson et al. (1976) and the Canadian Wetland Classification System (NWW, 1997)

Class	Description
Barren Land	Land of limited ability to support life; less than one-third of the area has vegetation or other cover (e.g., sands, rocks, thin soil).
Forested Upland	Closed-canopy deciduous, coniferous, or mixed forests.
Herbaceous Upland	Land where vegetation is dominated by a mix of grasses, grasslike plants, forbs, shrubs, or bush; either naturally occurring or modified (e.g., old fields, roadside vegetation, meadows, mixed-composition short vegetation upland).
Water	All areas that are persistently water covered (e.g., lakes, reservoirs, streams, bays, estuaries)
Wetland	Bog, fen (or wet meadow), swamp, marsh, shallow open water.
Marsh	A shallow water wetland with water levels that can fluctuate daily, annually, or seasonally, resulting in highly variable hydrology. Receives water from the surrounding catchment as well as precipitation. Marsh vegetation is comprised of emergent aquatic macrophytes such as graminoids (e.g., rushes, reeds, sedges), floating-leaved species (e.g., lilies) and submergent species (e.g., water milfoil). Marsh plant communities are seasonal and dynamic, often shifting with water levels.
Swamp	Forested or wooded wetland, dominated by minerotrophic groundwater and a water table below the ground surface of the swamp for the majority of the year. Vegetation dominated by coniferous or deciduous trees or tall shrubs (generally over 30%).
Bog	A peat landform, raised or level with the surrounding terrain and isolated from runoff and groundwater, receiving water primarily from precipitation, fog, and snowmelt. Water table sits at or slightly below the bog surface. Treed or treeless, and usually covered with *Sphagnum* spp. and shrubs, or woody remains of shrubs.

14.1.2 Brief History of Wetland Remote Sensing Applications

Wetlands have been studied using all major satellite systems and a variety of classification methods (Ozesmi & Bauer, 2002). Frequently used satellite sensors for wetland detection include multispectral and hyperspectral imaging systems, radar systems, and LiDAR sensors. Active sensors have been used to map temporal food dynamics (Martinez & Le Toan, 2007), improve detection of below-canopy inundation (Lang & McCarty, 2009), estimate water storage (Grings et al., 2009), and assess soil moisture (Kasischke et al., 2009). Other sensors, such as microwave radiometers, have been used to estimate methane emissions in the Amazon basin (Melack et al., 2004). Notable studies using optical sensors include the large-scale mapping of aquatic vegetation and habitat features across the Great Lakes shoreline in Ontario, Canada, using IKONOS data and achieving an overall classification accuracy of 90% (Wei & Chow-Fraser, 2007); a multitemporal SPOT-5 classification tree approach to monitoring aquatic marsh vegetation in southern France, using a

variety of vegetation indices with accuracies greater than 80% (Davranche et al., 2010); and a Landsat Thematic Mapper (TM) approach combined with image texture and ancillary data to model palustrine wetland occurrence in Yellowstone National Park, with error rates less than 14% (Wright & Gallant, 2007). Often, a combination of sensors has proven advantageous. Prigent et al. (2001) described a global study to map wetland inundation using passive and active microwave sensors along with visible and infrared measurements. Major inundated wetlands, such as the Pantanal, riverine wetlands along the Amazon, and the Niger Delta, were well delineated; however, the accuracy assessment was hampered by the scarcity of quantitative observations across the globe. Others have mapped wetlands in the context of the Canadian Wetland Inventory (CWI) using a combination of RADARSAT-1 and Landsat enhanced TM (ETM) images in Quebec, Canada, with global accuracy values between 67% and 80% (Grenier et al., 2007). Similarly, Dingle-Robertson & King (2014) mapped wetlands in the context of the Ontario Wetland Evaluation System (OWES) using multitemporal WorldView-2, Landsat 5, and RADARSAT-2 data. Similarly, time series image data have been found useful for wetland boundary and change detection (e.g., Davranche et al., 2010; Johnston & Barson, 1993), while digital elevation models (DEMs) have proven useful to aid in delineation of wetlands by separating floodplains and inland valley wetlands from other landforms of higher elevation (e.g., Mwita et al., 2013).

Identifying and mapping wetland vegetation by species or type (e.g., floating, emergent, or submerged aquatic vegetation) represents another common objective. For species-level mapping, hyperspectral sensors have become a necessity and have been used successfully to discriminate between aquatic plant species (e.g., Becker et al., 2007; Hirano et al., 2003; Zomer et al., 2009). In mapping wetland species, many researchers have turned their focus toward identification of invasive species, such as *Phragmites australis*. The negative impacts of this species are related to its propensity to produce dense monospecific stands, outcompete native plant diversity, and subsequently reduce the overall richness of flora and associated faunal assemblages. However, the ability of satellite imagery to map to the species level requires further development of a universal library of spectral signatures because it is unclear whether all species or vegetation communities of interest possess unique signatures.

Several studies have developed successful methods for identifying *Phragmites* from remotely sensed imagery. Pengra et al. (2007) demonstrated the utility of EO-1 Hyperion data (220 unique spectral bands) for mapping *Phragmites*, using a reference spectrum of pure stands and a spectral correlation mapper algorithm to compare with the Hyperion raster image. Monodominant stands of the invasive plant were classified with a promising overall accuracy of 81.4%. The spatially varied *Phragmites* stands were arranged in linear clusters parallel to the shoreline, which was not optimal given the coarse spatial resolution (30 m) of Hyperion data, but provided a useful tool for wetland managers nonetheless. Using a multi-imagery technique, Walsh et al. (2008) leveraged the high spectral resolution of Hyperion data to identify unique plant spectral signatures, with high spatial resolution (2.44 m) QuickBird imagery to map invasive wetland plants in the Galapagos Islands using a spectra unmixing approach. Other studies have also found success by pairing hyperspectral ground or airborne radiometers with multispectral satellite imagery to successfully map wetland plant

species (e.g., Yuan & Zhang, 2006) or wetland degradation (e.g., Schmid et al., 2005). Collectively, these studies demonstrate that improved results can be obtained by integrating multiple sensors, which may become a necessity in order to obtain the appropriate spectral, spatial, and temporal resolution for species mapping.

Although satellite data from coarse spatial resolution satellites have been used in many studies, they are limited by their inability to identify wetlands that are small in size, long, or narrow in shape (Ozesmi & Bauer, 2002), or to distinguish between spatially complex features within wetlands. Coarser 30 m data require a minimum of 9 pure pixels (0.81 ha) to consistently identify a feature (Ozesmi & Bauer, 2002), resulting in many mixed pixels and small wetlands below this threshold being missed (Klemas, 2011; Powers et al., 2011). Popular Landsat imagery may be insufficient to provide the required detail and resolution for some applications of wetland use mapping (Mwita et al, 2013). The following section will discuss the use of high spatial resolution imagery in wetland detection.

14.2 HIGH SPATIAL RESOLUTION IMAGERY IN WETLAND REMOTE SENSING

Wetlands are often distinguished by high spatial complexity and temporal variability; therefore, they are best observed with sensors capable of providing a combination of high spatial, spectral, and temporal resolution. Advances in sensor technology continue to provide numerous options to match a research question or application with an appropriate sensor, though selection entails trade-offs between sensor benefits and cost. This section will focus on the advantages conferred by high spatial resolution sensors for wetland remote sensing detection at a nominal ground pixel resolution of <2 m.

While freely available satellite imagery such as Landsat are publicly available, the cost associated with obtaining high spatial resolution satellite data can be high. Despite this cost it is still significantly lower than the cost required to support field surveying or aerial photographs (see Wei & Chow-Fraser (2007) for a cost breakdown) while providing the advantage of repeat coverage for monitoring temporal trends and the addition of data outside of the optical range (e.g., in the near-infrared [NIR] region). Current work with high spatial resolution sensors has been used to successfully monitor the change in aquatic vegetation in coastal marshes (Wei & Chow-Fraser, 2007), to discriminate between submerged and emergent wetland vegetation (Davranche et al., 2010), and to estimate marshland composition and biomass in riparian marshes (Dillabaugh & King, 2008).

While high-resolution imagery provides the spatial resolution to capture smaller wetlands, it results in greater within-class spectral variance, making separation of mixed or similar land cover classes more difficult than with coarser-resolution imagery (Hu and Weng, 2011; Klemas, 2011). To address this increased variance, alternative classification approaches have been employed. In recent decades, object-based image analysis (OBIA), or geographic object-based image analysis (GEOBIA), has shown superior performance compared to traditional pixel-based methods. Packaging of pixels into discrete objects minimizes the variance (noise) inherent in high spatial resolution images, allowing the objects, rather than individual pixels, to be classified. The object-based approach has been successfully used in wetland

research for classifying macrophyte communities in coastal marsh habitat (Midwood & Chow-Fraser, 2010; Rokitnicki-Wojcik et al., 2011), evaluating the structure of patterned peatlands (Dissanska et al., 2009), and mapping multiple classes of wetlands according to the Canadian Wetland Inventory (Grenier et al., 2007). Fournier et al. (2007) reviewed wetland mapping methods to be applied to the Canadian Wetlands Inventory program and identified the object-based approach as most appropriate due to its flexibility and ability to address the spatial heterogeneity of wetlands.

14.3 CASE STUDY: DETECTING FRESHWATER WETLANDS IN ALGONQUIN PARK, ONTARIO, CANADA, USING GEOEYE-1 IMAGERY

In this section, a case study is presented that demonstrates an effective approach for detecting inland freshwater wetlands of varying size, shape, and composition. This method was evaluated on two freshwater wetlands that demonstrate significant change in open water extent and spatial distribution of macrophytic communities within a single growing season (Figure 14.1).

Multispectral GeoEye-1 satellite imagery was acquired during September 2012 and May 2013 over a study site in southern Ontario (44°00′ N, 80°00′ W). Multiseason images were processed and classified using a hierarchical image object segmentation scheme and a supervised *k*-nearest neighbor (*k*-NN) classifier. Ancillary spatial data, such as the normalized difference vegetation index (NDVI), elevation, and texture layers, were integrated into the model, while ground data and aerial imagery were used to refine, verify, and validate maps.

FIGURE 14.1 Seasonal change in inundation and vegetation extent in two freshwater marshes in southern and south-central Ontario. Algonquin Provincial Park in (a) May and (b) September. Oakland Swamp in rural Brantford, Canada, in (c) April, (d) June, and (e) August.

Approaches to wetland classification have encompassed traditional unsupervised (Jensen et al., 1995; Sawaya et al., 2003) and supervised algorithms (Wang et al., 2004; Yu et al., 2006), as well as fuzzy methods (Benz et al., 2004) and object-based approaches (Blaschke, 2010; Blaschke et al., 2014). They have also included more complex machine learning algorithms such as classification tree methods (Midwood & Chow-Fraser, 2010; Wright & Gallant, 2007), including random forest classification (Corcoran et al., 2013), with some complex models drawing from numerous data layers to discriminate among wetland types (Wright & Gallant, 2007). As a result, it is not surprising that many researchers have devoted entire studies to comparing the utility of different methods (Dingle-Robertson & King, 2011; Duro et al., 2012; Harken & Sugumaran, 2005; Shanmugam et al., 2006). However, no general consensus has been reached on a universal optimal method of identifying wetlands. The object-based approach has been used extensively in image analysis since the start of the twenty-first century, with hundreds of studies conducted on this topic (e.g., review by Blaschke (2010)). This case study demonstrates a GEOBIA supervised classification approach to identify and delineate inland freshwater wetlands in a managed provincial park in southern Ontario, using high spatial resolution (1.84 m) GeoEye-1 satellite imagery.

14.3.1 Study Area

The region encompassed by Algonquin Provincial Park in southern Ontario, Canada, represents 7630 km^2 of protected land, including approximately 340 ha of wetlands of all classes, as defined by the Canadian Wetland Classification System (NWWG, 1997). Established in 1893, it is the oldest provincial park in Canada, and contains a unique mixture of coniferous and deciduous forest types, allowing the park to support an uncommon diversity of native flora and fauna. The park also protects the headwaters of five major watersheds and encompasses more than 2,000 lakes, which make up about 10% of the total park area (Figure 14.2).

Logging activity occurs in the study area, as well as recreational use by park visitors, though the study site was situated in a less heavily visited section. Dominant wetland types in the area include freshwater inland and lacustrine (lake-associated) wetlands, swamps, and bogs. Several threatened species are known to exist in the park, including wetland-associated species such as the Blanding's turtle (*Emydoidea blandingii*). Due to the park's considerable ecological assets, knowledge of wetland class, locations, and boundaries are important for management and mitigation of negative effects from anthropogenic activities within the boundaries of the park and for regular monitoring of ecosystem health.

14.3.2 Satellite Data

Data were collected by the spaceborne GeoEye-1 sensor on May 23, 2013. and September 12, 2012. In the study area, full deciduous leaf-on conditions are typically reached by the end of May or beginning of June, while leaf-off conditions generally occur by late October or early November. All efforts were made to acquire imagery within the same year; however, this was not possible due to cloud cover over the

FIGURE 14.2 Map of study area located in Algonquin Provincial Park, Ontario, Canada.

site. The imagery had a pixel size of 1.85 m, across four spectral regions (blue [450–510 nm]; green [510–580 nm]; red [655–690 nm]; near-infrared [780–920 nm]), at a radiometric resolution of 16 bits. GeoEye-1 data were selected because they offered a balance between high spatial resolution imagery, spectral coverage, and cost. While a sensor with a spectral band encompassing the water-sensitive shortwave infrared region (approximately 10–100 μm) would have been useful for wetland detection, no high spatial resolution sensor with this criterion was available at the time of the study. Currently, most high spatial resolution sensors of <2 m pixel size operate in the visible and near-infrared (VNIR) regions only (e.g., GeoEye, SPOT 6/7, WorldView-1 and -2, Pleiades-1A and -1B, QuickBird). Image preprocessing and processing is demonstrated in the flowchart (Figure 14.3) and described subsequently.

14.3.3 Image Preprocessing

The images were atmospherically corrected using the broadband visible/NIR algorithm for 4-band VNIR sensors (Richter, 1996; Richter et al., 2006), including correction for thin cloud contamination across diverse surfaces (Zhang et al., 2002), which is a technique optimized for high spatial resolution imagery (Richter, 1996). This algorithm required three visible and one near-infrared band for correction

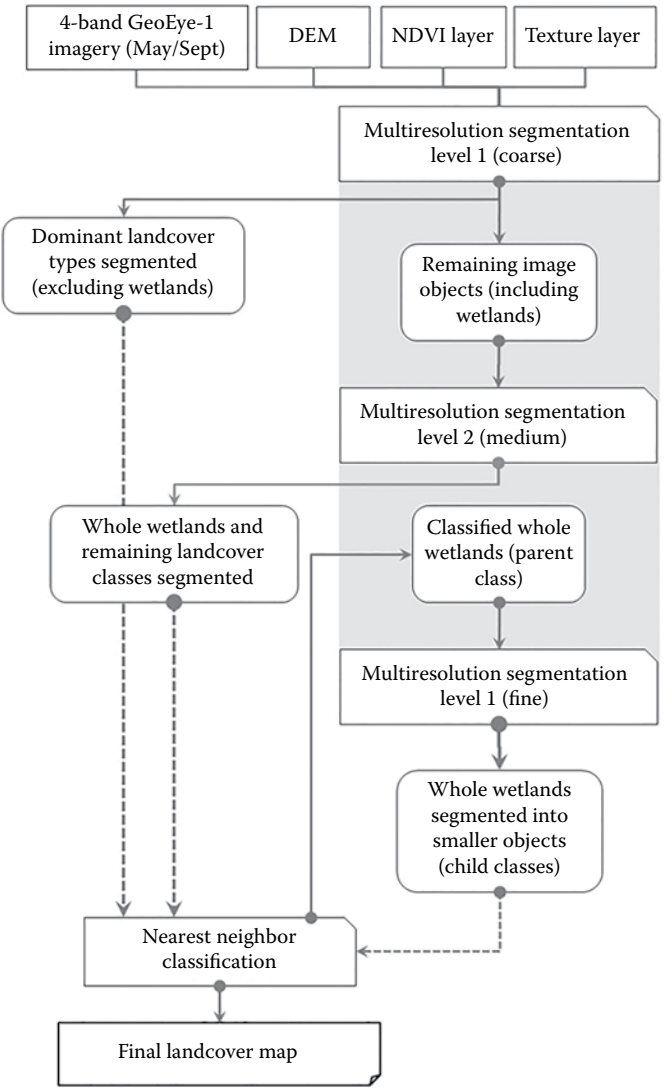


FIGURE 14.3 Multiscale segmentation process used to segment images at three levels, with a hierarchical parent-child relationship developed between wetlands and within wetland components at the medium (Level 2) and fine (Level 1) scale.

(Richter et al., 2006), which is a common band configuration for many high spatial resolution Earth-observing satellites (e.g., GeoEye-1, SPOT-6) and for economical (reduced-band) options of imagery, such as those from WorldView-2, which provides data at eight (full) or four (reduced) bands. Radiometric normalization was implemented in PCI Geomatica (ATCOR3 module, Geomatica version, 2014), though cloud cover and haze was minimal. After atmospheric correction, the images were projected to the Universal Transverse Mercator projection datum (NAD83, UTM Zone 17) and georeferenced to a root-mean-squared error (RMSE) of less than 2 pixels using a first-order polynomial transformation and a nearest neighbor resampling method, corresponding to less than 4 m ground error. The processed image was clipped to a 40 km^2 boundary using ArcGIS version 10.2 (Environmental Systems Research Institute, Redlands, California). The panchromatic layer was not used because it increased processing time to unrealistic lengths. This is a common limitation of high spatial resolution data.

14.3.4 Classification Method

14.3.4.1 Input Layers

Prior to classification the image was segmented into objects. Image objects were constructed from seven features related to image attributes such as image texture (localized neighborhood spectral variation), elevation, and NDVI. Layers included four multispectral bands from GeoEye-1 (blue, green, red, and near-infrared), a DEM layer, an NDVI layer, and a standard deviation texture layer (Figure 14.4).

A 10 m DEM was acquired from the Ontario Ministry of Natural Resources and Forestry, which was interpolated from a digital terrain model (DTM), a contour map, spot height data, and a water virtual flow map to a $\pm$10 m vertical precision. The DEM for each image scene was resampled to 2 m to match the resolution and alignment of the other input layers. Resampling the DEM did not provide any additional information but ensured continuity in pixel size across all input layers and avoided a coarser resolution affecting the boundaries of image objects. Elevation was included as an input layer because wetlands and water bodies are known to sit topographically low in the landscape due to their close association with ground water and surface runoff (Mitsch & Gosselink, 2000). Although depressions in the DEM layer are likely areas where water will pool, image texture has also been utilized for wetland delineation because wetland vegetation is often smaller and highly interconnected compared to terrestrial vegetation such as trees. Texture information refers to localized spatial variation in the spectral brightness of a digital image and has a high potential for revealing differences between classes in remotely sensed imagery (Berberoğlu et al., 2010). Texture measures were derived directly from GeoEye-1 satellite imagery by creating a first-order texture layer based on the standard deviation of the spectral layer(s) within a 3 by 3 pixel moving window. However, image texture does not correlate to the density of vegetation, whereas spectral vegetation indices are highly correlated with vegetation abundance. NDVI is a well-established indicator of live green vegetation (Rouse et al., 1974), and was created from the red and near-infrared bands of the multispectral data according to the following equation:

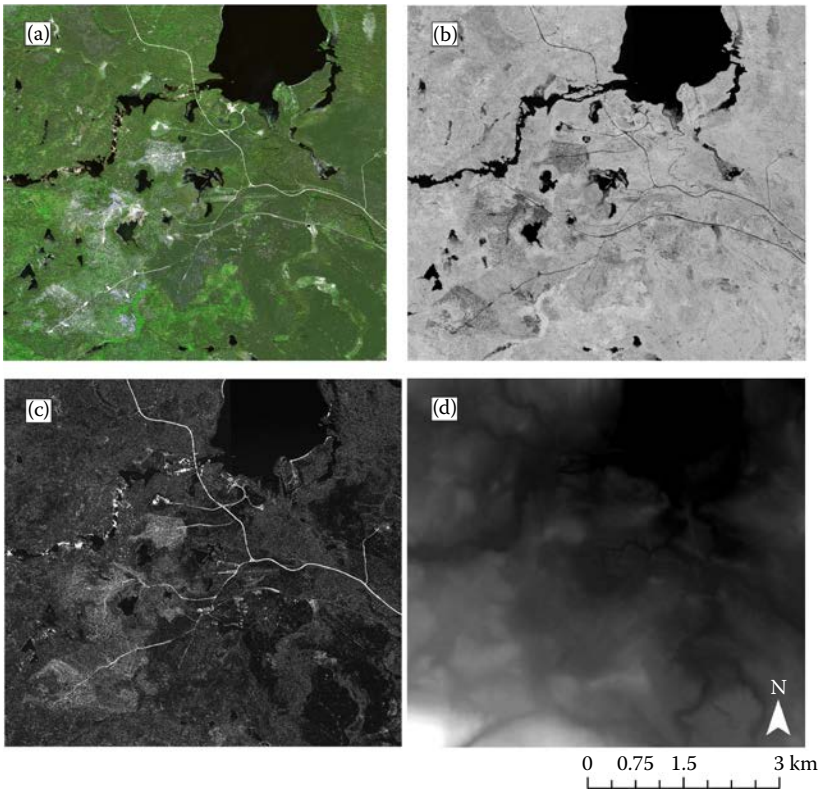


FIGURE 14.4 GeoEye-1 image over the study area. Input layers include (a) four multispectral bands (blue, green, red, near-infrared) shown in false color composite, (b) NDVI layer, (c) standard deviation texture layer, and (d) digital elevation layer.

$$\text{NDVI} = \frac{\rho\text{nir} - \rho\text{red}}{\rho\text{nir} + \rho\text{red}} \tag{14.1}$$

where ρ is the reflectance of the visible (red) and near-infrared (nir) bands of the electromagnetic spectrum. NDVI values range from −1 to 1; higher NDVI values indicate a greater coverage of live green vegetation, while values less than zero typically do not have any ecological meaning. NDVI has been used to separate water from dry land and for delineating wetland boundaries (Ozesmi & Bauer, 2002). All final image layers were weighted equally in the multiresolution segmentation process.

14.3.4.2 Multiresolution Segmentation

In the object-oriented approach, both spectral and spatial (or contextual) parameters are used to define an image object, whereas traditional per-pixel classifiers treat each individual pixel independently of its neighbors. This study employed the fractal net evolution approach (FNEA) to segmenting images (Baatz & Schäpe, 2000), which

was implemented through the multiresolution segmentation algorithm in Definiens Developer 7.0 (formerly eCognition) (Munich, Germany). Segmentation parameters of scale, color, and shape control the size, shape, and resultant spectral variation within segmented image objects. Weights of color and shape sum to 1, while shape is further divided into smoothness (relating to the smoothness of object edges) and compactness (relating to the closeness of an object shape with a circle), which sum to 1 (Definiens, 2008). The color parameter was set to 0.9 to place greater emphasis or weight on pixel values during the segmentation decision process. The 0.1 shape parameter was weighted equally to balance the compactness and smoothness of object boundaries equally. The most critical step is the selection of the scale parameter (unitless), which controls the size of the image objects by sequentially merging pixels pairwise, with the intent of minimizing the heterogeneity within (Blaschke & Hay, 2001; Mallinis et al., 2008). The scale parameter sets a threshold of homogeneity which determines how many neighboring pixels can be merged together to form an image object (Benz et al., 2004).

Three levels of scale parameterization were used to capture different land cover classes in a multiscaled segmentation approach. Segmentation scale designations of coarse, medium, and fine are relative terms relating to image composition and desired specificity of the target feature. Dominant land cover classes were segmented at the highest (coarse) level, while remaining classes were delineated at the medium level. Entire wetlands were segmented and defined as objects at this midrange scale, and further segmented at the finest scale level to delineate components within wetlands to classify them into as marsh, swamp, fen, or bog. These smaller (child) objects retain links to their larger (parent) class, which employs a true multiscale approach through applying vertical constraints in segmentation and classification. Classification for specific land cover classes was thus completed at each scale level, with remaining unclassified objects undergoing further segmentation, followed by classification. A thematic road was used to extract the road network for each scene and ensure parcels of intervening land were accurately delineated.

14.3.4.3 Note on Selection of an Appropriate Scale Parameter

There has been much debate surrounding the method of scale parameter selection, which determines the ultimate size and shape of the image objects created during the segmentation process. Many studies have employed a visual approach to selecting the best scale parameter, based on the premise that the human eye is best capable of interpreting and recognizing complex patterns in conjunction with neighborhood context (Benz et al., 2004; Myint et al., 2011). This approach is especially fitting for wetlands that can be highly variable in both size and shape, particularly when visual assessment is guided by field knowledge, thematic maps, and ancillary aerial imagery. However, opponents of this method point to the subjectivity of such an approach and the limitations for repeatability across different study areas. With the continuous advent of new computer-based algorithms, multiple quantitative and automated approaches to selecting the scale parameter have been developed, including automated parameterization using the potential of local variance to detect scale transitions (Drăguţ et al., 2010, 2014), supervised methods that use various indices to describe the discrepancy between reference polygons and corresponding

image objects (Clinton et al., 2010; Liu et al., 2012), and a comparison index using both topological and geometric object metrics (Moller et al., 2007). However, there is no perfect algorithm that is appropriate for all images and a certain element of trial, error, and repetition is inherent to the overall process of scale selection and evaluation. In this case study, scale selection was based on a combination of visual assessment using expert knowledge of the area as well as a semiautomated quantitative method (Modified Euclidean Distance 3; Yang et al., 2015) to select the optimal scale parameter. Additional details regarding this approach can be found in Mui et al. (2015) and Yang et al. (2015).

14.3.4.4 *k*-NN Classification

The nearest neighbor, or *k*-NN, approach is a simple yet efficient classification algorithm that has been shown to perform as well as more complicated methods, such as support vector machines (SVMs), when the same classification type (object or pixel based) and contextual features were used (Im et al., 2008). Classification was performed based on image object attributes. However, a considerable range of attributes can be extracted from image objects. In order to maintain a realistic processing time, a parsimonious model based on mean object value and standard deviation for each input layer was utilized. A spatially representative sample of training objects was hand selected to inform the classifier. This iterative process involved selecting training samples, comparing sample attributes, and refining training samples until a satisfactory result was achieved.

A minimum of 35 independently selected image objects per class were used for accuracy assessment. Validation image object selection was based on very-high-resolution (VHR) aerial photographs over each site, reference thematic maps, and ground truth data collected in the study site. Validation and training sample image objects did not overlap. Accuracy was assessed based on the error matrix and associated statistics of overall accuracy, kappa statistic, producer's accuracy (1−errors of omission), and user's accuracy (1−errors of commission). Object-based assessment was preferred over pixel-based methods because this study was most interested in determining if wetland boundaries and marsh vegetation communities were accurately classified. With this objective, accuracy was better assessed using individual objects, which have clearly defined boundaries.

14.4 RESULTS

14.4.1 Segmentation Results

The unitless scale parameter is a relative value with higher numbers resulting in larger image objects, and lower numbers producing smaller image objects. Final scale values varied at each level of segmentation. The dominant land covers of mixed forest and water classes were most accurately delineated at a scale of 125, which created image objects with boundaries that were clearly defined with minimal absorption of smaller classes. A scale value of 40 was used at the next scale level, which captured whole wetland boundaries, as well as less common classes of barren land and herbaceous upland. Whole wetlands were further segmented at the finest

level at a scale of 20 to further classify these objects into marsh, swamp, spruce bog, or water. This parent-child relationship maintained a hierarchical constraint, which limited classification of the five wetland classes to only those objects defined in the previous step as wetlands.

14.4.2 Classification Results

GeoeEye-1 data were classified into eight classes with an overall accuracy of 0.85–0.90 (kappa 0.82–0.88), while individual class accuracies varied (Table 14.2). Wetlands comprised a small percentage of the total mapped area, with the remainder consisting of open water and terrestrial land cover classes such as forested, nonforested, and barren (nonvegetated) land. Natural and managed forests were the dominant land cover type. Accuracy was higher in the earlier spring series image compared to the late summer series image.

The water class was mapped with the highest accuracy, which is not surprising given the distinct spectral characteristics of this feature. The herbaceous upland and barren land classes achieved the lowest accuracies. Out of the three wetland types, the swamp class received the lowest accuracy. The full classification map for each season is shown in Figure 14.5. Upland classes have been merged for better visualization of the target wetland classes.

Considerable change in vegetation and open water extent was experienced between seasons (Figure 14.6). Map accuracy was higher for the spring map compared to the late summer map. Classes with the lowest accuracy in the spring map were herbaceous land and swamps, while in the late summer map the lowest accuracy was achieved by the barren and herbaceous land classes.

TABLE 14.2
Summary of Error Matrix Statistics for the Spring and Summer Classification Maps, Including All Eight Land Cover Classes

	Spring (May)		Summer (September)	
Land Cover Class	**PA**	**UA**	**PA**	**UA**
Marsh	0.91	0.94	0.90	0.91
Swamp	0.83	0.81	0.81	0.79
Bog	0.87	0.84	0.82	0.83
Water	0.97	0.94	0.89	0.90
Forested Upland	0.94	0.91	0.86	0.84
Herbaceous Upland	0.80	0.82	0.78	0.77
Barren Land	0.92	0.95	0.79	0.76
Overall (kappa)	0.90 (0.88)		0.85 (0.82)	

PA, Producer's Accuracy; UA, User's Accuracy.

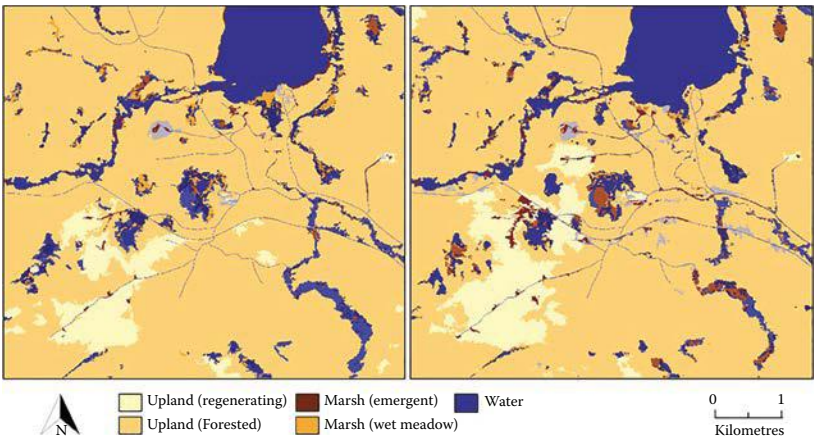


FIGURE 14.5 Final classification map over the Algonquin Park study region in the spring (left) and late summer (right). Upland classes have been merged to better visualize wetland classes. Upland (regenerating) class shows the extent of logging activity taking place within the boundaries of the park.

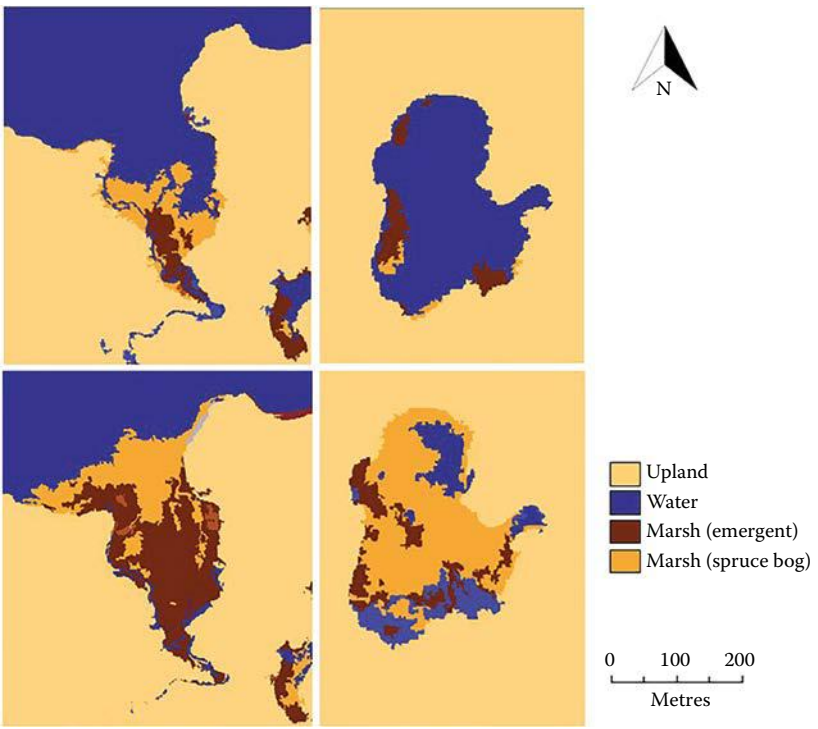


FIGURE 14.6 Comparison of wetland vegetation and open water extent in two freshwater marshes during the spring (top row) and late summer (bottom row).

14.5 DISCUSSION

The ability to identify a wetland from digital imagery is dependent on many factors, including the physical characteristics of the ground and the technical specifications of the sensor. In order to accurately detect wetlands, the spatially complex biotic and abiotic components such as open water and floating and emergent vegetation must be captured by the sensor. To delineate all wetlands, including those that are small (<2 ha) or linear in shape, the spatial resolution of the sensor must be sufficient to match the smallest feature on the ground. The focus of this study was to determine if single-source high spatial resolution imagery was sufficient to accomplish this task in a subset of Algonquin Provincial Park situated at the border of southern and central Ontario, Canada.

14.5.1 Satellite Considerations

GeoEye-1 data at 1.84 m pixel size proved satisfactory for capturing spatial detail representative of wetlands in the study area as well as dominant vegetation communities, and was used to determine wetland class to an overall accuracy of 90% in the spring and 85% in the late summer. Within-class variance was appropriately managed by applying the object-based approach to minimize spectral noise, and is supported by other studies evaluating wetland vegetation distribution (Dissanska et al., 2009; Midwood & Chow-Fraser, 2010; Rokitnicki-Wojcik et al., 2011).

Despite the greater range of spatial variation captured by GeoEye-1, high spatial resolution sensors are still subject to limitations of cloud cover and other atmospheric events. Cloud cover prevented acquisition of spring imagery during the most optimal time period (early May), while haze was present in a portion of the late summer (September) image, which may have resulted in the lower classification accuracy during this season. Additionally, the smaller swath width of high spatial resolution sensors (e.g., 15.2 km for GeoEye-1) reduces the probability of repeat coverage over a specific study area. Coarser spatial resolution satellites (e.g., Landsat 8 swath width of 185 km) that do not accept tasking orders cover larger regions of the globe continuously, and therefore can provide a more comprehensive archive of global imagery, albeit at the cost of less spatial detail. In this case study, we demonstrated the utility of ancillary spatial layers relevant to wetland ecosystem mapping. NDVI has a long history of use in vegetation mapping, while the DEM layer was incorporated to provide topographic information relevant to low-lying wetlands in the study area. The texture layer provided a measure of the spatial variation in pixel brightness values, and in wetland mapping has been found to limit oversegmentation and to produce image objects more representative of distinct vegetation communities (Mui et al., 2015).

14.5.2 Classification Approach

The multiscale object-based approach provided an effective method of partitioning wetlands and other land cover classes. The segmentation of whole wetland objects into smaller objects for within-wetland classification allowed this process to be constrained to its parent class, which minimized the potential for misclassification with terrestrial

objects. The scale parameters reported are for the purpose of comparison of land cover classes at each level (coarse, moderate, fine) within the study area, and not as a recommendation for optimal scale values to use for other images. Many factors, such as study area extent, composition, segmentation algorithm used, and specifications of the sensor, will affect the final scale value, and a universal scale parameter does not currently exist.

NDVI layers were found to improve boundary detection of classes at the coarse level, as well as for lacustrine wetlands which transition to open water (Mui et al., 2015). Elevation information improved segmentation of whole wetland boundaries, likely because of palustrine (inland) wetlands existing in depressional landforms. Texture contributed most at the finest scale level by limiting the amount of oversegmentation in an image (or production of more objects than necessary) (Mui et al., 2015). It was found that resultant image objects were larger than those segmented without textural information, and they also more accurately captured edges between different macrophyte communities. Previous work has shown that texture information can improve classification accuracy by reducing the confusion between permanent crops and perennial meadows (Peña-Barragán et al., 2011). Future work should focus on the integration of higher-order texture measures into classification schemes, such as those derived from the gray-level co-occurrence matrix (Haralick et al., 1973), which has shown success in discriminating between deciduous and evergreen tree species (Kim et al., 2009). This may help improve classification accuracy between treed uplands and swamps (treed wetlands), which represented the wetland class with the lowest classification accuracy in this study.

Wetlands may provide a challenge due to their relative rarity in the landscape. It is generally accepted that mapping error on less frequent classes will be higher than error on dominant classes (Cunningham, 2006), and the relative rarity of wetlands in the study area compared to dominant classes such as deciduous and conifer forests and lakes contributed to mapping error. Wright & Gallant (2007) documented a similar error for palustrine wetland mapping in Yellowstone National Park, for which wetlands comprised less than 6% of the total cover. Despite this, the recommended target of 85% overall accuracy (Foody, 2002; Thomlinson et al., 1999) was achieved for both the spring and late summer imagery, which suggests this approach was appropriate for the study area and objectives.

14.5.3 Seasonal Change

The higher map accuracy achieved by the spring map for all classes suggests that early season imagery, before vegetation has reached its peak growth, may provide better conditions for spectral-based mapping. Late season imagery showed a clear increase in vegetation extent and vigor across both aquatic and terrestrial regions, which dampen the below-canopy water signal. Because wetlands are characterized by soils inundated by water for all or part of the year, detection of the below-canopy signal is important for accurate delineation. Haze present in the late summer (September) image may have also limited classification accuracy by altering the spectral signal received by the sensor over the haze-covered region. Depending on the geographic location, season can have an obvious effect on wetland mapping, and other studies concerned with detecting

some aspect of change in wetlands must identify the optimal image acquisition period for their specific region. Munyati (2000) used dry season (September) images to map changes in a wetland floodplain in Zambia. The latter was chosen because at this time of year the wetland system stands out from the upland as the only area of dense green vegetation during a period of hot and dry climate. Similarly, to capture seasonal change in wetland inundation and vegetation in the central Amazon basin, Hess et al. (2003) used August–September and May–June synthetic-aperture radar (SAR) imagery to capture peak low and high water stages. Thus, the timing of image acquisition has been shown to be important for capturing the feature or process of interest and should be given due consideration before a final image is selected or tasked.

14.6 CONCLUSIONS

Presented in this chapter was a simple yet efficient methodology for mapping spatially, temporally, and spectrally complex wetlands using a single-source high spatial resolution optical sensor. This work demonstrates that wetland complexity can be captured using high spatial resolution satellite imagery, which includes wetlands of varying size and shape, as well as the aquatic vegetation communities found within. From a practical perspective, there are many areas where a previous wetland inventory does not exist or cannot be conducted for logistical reasons. The method of mapping wetlands presented here can provide starting information for better management of these ecosystems by delivering spatially explicit information on wetland location and extent. Other studies with the objective of capturing small, irregularly shaped, and isolated wetlands should consider the use of high spatial resolution data to meet these needs. Even studies concerned with larger wetlands can benefit from imagery with a finer pixel size to capture vegetation type and distribution of open water with greater accuracy and detail. With continuing advancements in sensor design and capability, we anticipate that high spatial resolution data will become more affordable and available for public use in due course, and thus researchers should continue to incorporate fine-resolution imagery into remote sensing studies to continue the advancement of our knowledge in this area.

REFERENCES

Anderson, J. R., Hardy, E. E., Roach, J. T., & Witmer, R. E. 1976. A land use and land cover classification system for use with remote sensor data. A revision of the land use classification system as presented in U.S. *Geological Survey Circular* 671, Professional Paper 964, USGS, Washington, DC.

Baatz, M., & Schäpe, A. 2000. Multiresolution segmentation: An optimization approach for high quality multi-scale image segmentation. *Angewandte Geographische Informationsverarbeitung* XII, 58, 12–23.

Becker, B. L., Lusch, D. P., & Qi, J. 2007. A classification-based assessment of the optimal spectral and spatial resolutions for Great Lakes coastal wetland imagery. *Remote Sensing of Environment*, 108, 111–120.

Benz, U. C., Hofmann, P., Willhauck, G., Lingenfelder, I., & Heynen, M. 2004. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry and Remote Sensing*, 58, 239–258.

Berberoğlu, S., Akin, A., Atkinson, P. M., & Curran, P. J. 2010. Utilizing image texture to detect land-cover change in Mediterranean coastal wetlands. *International Journal of Remote Sensing*, 31(11), 2793–2815.

Blaschke, T. 2010. Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(1), 2–16.

Blaschke, T., & Hay, G. J. 2001. Object-oriented image analysis and scale-space: Theory and methods for modelling and evaluating multiscale landscape structure. *International Archives of Photogrammetry and Remote Sensing*, 34(4), 22–29.

Blaschke, T., Hay, G. J., Kelly, M., Lang, S., Hofmann, P., Addink, E., Feitosa, R. Q., van der Meer, F., van der Werff, H., van Coillie, F., & Tiede, D. 2014. Geographic object-based image analysis—Towards a new paradigm. *ISPRS Journal of Photogrammetry and Remote Sensing*, 87(1), 180–191.

Brock, M. A., Smith, R. G. B., & Jarman, P. J. 1999. Drain it, dam it: Alteration of water regime in shallow wetlands on the New England Tableland of New South Wales, Australia. *Wetlands Ecology and Management*, 7, 37–46.

Clinton, N., Holt, A., Scarborough, J., Yan, L., & Gong, P. 2010. Accuracy assessment measures for object-based image segmentation goodness. *Photogrammetric Engineering & Remote Sensing*, 76(3), 289–299.

Corcoran, J. M., Knight, J. F., & Gallant, A. L. 2013. Influence of multi-source and multi-temporal remotely sensed and ancillary data on the accuracy of random forest classification of wetlands in northern Minnesota. *Remote Sensing*, 5, 3212–3238.

Cunningham, M. A. 2006. Accuracy assessment of digitized and classified land cover data for wildlife habitat. *Landscape and Urban Planning*, 78(3), 217–228.

Davranche, A., Lefebvre, G., & Poulin, B. 2010. Wetland monitoring using classification trees and SPOT-5 seasonal time series. *Remote Sensing of Environment*, 114, 552–562.

Definiens. 2008. *Definiens Developer 7 User Guide*. Munchen, Germany.

Dillabaugh, K. A., & King, D. J. 2008. Riparian marshland composition and biomass mapping using Ikonos imagery. *Canadian Journal of Remote Sensing*, 34(2), 143–158.

Dingle-Robertson, L., & King, D. J. 2011. Comparison of pixel- and object-based classification in land cover change mapping. *International Journal of Remote Sensing*, 32(6), 1505–1529.

Dingle-Robertson, L., & King, D. J. 2014. Evaluating spatial and seasonal variability of wetlands in eastern Ontario using remote sensing and GIS. PhD Thesis. University of Carleton.

Dissanska, M., Bernier, M., & Payette, S. 2009. Object-based classification of very high resolution panchromatic images for evaluating recent change in the structure of patterned peatlands. *Canadian Journal of Remote Sensing*, 35(2), 189–215.

Drăguţ, L., Csillik, O., Eisank, C., & Tiede, D. 2014. Automated parameterisation for multi-scale image segmentation on multiple layers. *ISPRS Journal of Photogrammetry and Remote Sensing*, 88, 119–127.

Drăguţ, L., Tiede, D., & Levick, S. R. 2010. ESP: A tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. *International Journal of Geographical Information Science*, 24(6), 859–871.

Duro, D. C., Franklin, S. E., & Dubé, M. G. 2012. A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. *Remote Sensing of Environment*, 118, 259–272.

Foody, G. M. 2002. Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, 80, 185–201.

Fournier, R. A., Grenier, M., Lavoie, A., & Hélie, R. 2007. Towards a strategy to implement the Canadian Wetland Inventory using satellite remote sensing. *Canadian Journal of Remote Sensing*, 33, 1–16.

Gibbs, J. P. 1993. Importance of small wetlands for the persistence of local populations of wetland-associated animals. *Wetlands*, 13(1), 25–31.

Grenier, M., Demers, A. M., Labrecque, S., Benoit, M., Fournier, R. A., & Drolet, B. 2007. An object-based method to map wetland using RADARSAT-1 and Landsat ETM images: Test case on two sites in Quebec, Canada. *Canadian Journal of Remote Sensing*, 33(1), S28–S45.

Grings, F., Salvia, M., Karszenbaum, H., Ferrazzoli, P., Kandus, P., & Perna, P. 2009. Exploring the capacity of radar remote sensing to estimate wetland marshes water storage. *Journal of Environmental Management*, 90(7), 2189–2198.

Halabisky, M. 2011. Object-based classification of semi-arid wetlands. *Journal of Applied Remote Sensing*, 5(1), 053511.

Haralick, R. M., Shanmugam, K., & Dinstein, I. 1973. Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, 3(6), 610–621.

Harken, J., & Sugumaran, R. 2005. Classification of Iowa wetlands using an airborne hyperspectral image: a comparison of the spectral angle mapper classifier and an object-oriented approach. *Canadian Journal of Remote Sensing*, 31(2), 167–174.

Hess, L. L., Melack, J. M., Novo, E. M., Barbosa, C. C., & Gastil, M. 2003. Dual-season mapping of wetland inundation and vegetation for the central Amazon basin. *Remote Sensing of Environment*, 87(4), 404–428.

Hirano, A., Madden, M., & Welch, R. 2003. Hyperspectral image data for mapping wetland vegetation. *Wetlands*, 23(2), 436–448.

Hu, X., & Weng, Q. 2011. Impervious surface area extraction from IKONOS imagery using an object-based fuzzy method. *Geocarto International*, 26(1), 3–20.

Im, J., Jensen, J.R., & Tullis, J.A., 2008. Object-based change detection using correlation image analysis and image segmentation. *International Journal of Remote Sensing*, 29(2), 399–423.

Jensen, J. R., Rutchey, K., Koch, M. S., & Narumalani, S. 1995. Inland wetland change detection in the Everglades Water Conservation Area 2A using a time series of normalized remotely sensed data. *Photogrammetric Engineering & Remote Sensing*, 61(2), 199–209.

Johnston, R., & Barson, M. 1993. Remote sensing of Australian wetlands: An evaluation of Landsat TM data for inventory and classification. *Australian Journal of Marine and Freshwater Research*, 44, 235–252.

Kasischke, E. S., Bourgeau-Chavez, L. L., Rober, A. R., Wyatt, K. H., Waddington, J. M., & Turetsky, M. R. 2009. Effects of soil moisture and water depth on ERS SAR backscatter measurements from an Alaskan wetland complex. *Remote Sensing of Environment*, 113(9), 1868–1873.

Kayranli, B., Scholz, M., Mustafa, A., & Hedmark, Å. 2010. Carbon storage and fluxes within freshwater wetlands: A critical review. *Wetlands*, 30(1), 111–124.

Kim, M., Madden, M., & Warner, T. A. 2009. Forest type mapping using object-specific texture measures from multispectral Ikonos imagery: Segmentation quality and image classification issues. *ISPRS Journal of Photogrammetry and Remote Sensing*, 75(7), 819–829.

Klemas, V. 2011. Remote sensing techniques for studying coastal ecosystems: An overview. *Journal of Coastal Research*, 27, 2–17.

Lang, M. W., & McCarty, G. W. 2009. LiDAR intensity for improved detection of inundation below the forest canopy. *Wetlands*, 29(4), 1166–1178.

Liu, Y., Bian, L., Meng, Y., Wang, H., Zhang, S., Yang, Y., Shao, X., & Wang, B. 2012. Discrepancy measures for selecting optimal combination of parameter values in object-based image analysis. *ISPRS Journal of Photogrammetry and Remote Sensing*, 68, 144–156.

Mallinis, G., Koutsias, N., Tsakiri-Strati, M., & Karteris, M. 2008. Object-based classification using Quickbird imagery for delineating forest vegetation polygons in a Mediterranean test site. *ISPRS Journal of Photogrammetry and Remote Sensing*, 63, 237–250.

Martinez, J. M., & Le Toan, T. 2007. Mapping of flood dynamics and spatial distribution of vegetation in the Amazon floodplain using multitemporal SAR data. *Remote Sensing of Environment*, 108(3), 209–223.

Melack, J. M., Hess, L. L., Gastil, M., Forsberg, B. R., Hamilton, S. K., Lima, I. B., & Novo, E. M. 2004. Regionalization of methane emissions in the Amazon basin with microwave remote sensing. *Global Change Biology*, 10(5), 530–544.

Midwood, J. D., & Chow-Fraser, P. 2010. Mapping floating and emergent aquatic vegetation in coastal wetlands of Eastern Georgian Bay, Lake Huron, Canada. *Wetlands*, 30, 1141–1152.

Mitsch, W. J., & Gosselink, J. G. 1993. *Wetlands*. Wiley & Sons, Hoboken, NJ.

Mitsch, W. J., & Gosselink, J. G. 2000. The value of wetlands: Importance of scale and landscape setting. *Ecological Economics*, 35(1), 25–33.

Möller, M., Lymburner, L., & Volk, M. 2007. The comparison index: A tool for assessing the accuracy of image segmentation. *International Journal of Applied Earth Observation and Geoinformation*, 9(3), 311–321.

Moore, T. R., Heyes, A., & Roulet, N. T. 1994. Methane emissions from wetlands, southern Hudson Bay lowland. *Journal of Geophysical Research*, 99(D1), 1455–1467, doi:10.1029/93JD02457.

Mui, A., He, Y., & Weng, Q. 2015. An object-based approach to delineate wetlands across landscapes of varied disturbance with high spatial resolution satellite imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 109, 30–46.

Munyati, C. 2000. Wetland change detection on the Kafue Flats, Zambia, by classification of a multitemporal remote sensing image dataset. *International Journal of Remote Sensing*, 21(9), 1787–1806.

Mwita, E., Menz, G., Misana, S., Becker, M., Kisanga, D., & Boehme, B. 2013. Mapping small wetlands of Kenya and Tanzania using remote sensing techniques. *International Journal of Applied Earth Observation and Geoinformation*, 21, 173–183.

Myint, S. W., Gober, P., Brazel, A., Grossman-Clarke, S., & Weng, Q. 2011. Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. *Remote Sensing of Environment*, 115(5), 1145–1161.

NWWG (National Wetlands Working Group). 1997. *The Canadian Wetland Classification System*, 2nd Edition. Warner, B. G., and Rubec, C. D. A. (eds.), Wetlands Research Centre, University of Waterloo, Waterloo, ON, Canada.

Ozesmi, S., & Bauer, M. 2002. Satellite remote sensing of wetlands. *Wetlands Ecology and Management*, 10, 381–402.

Peña-Barragán, J.M., Ngugi, M.K., Plant, R.E., Six, J., 2011. Object-based crop identification using multiple vegetation indices, textural features and crop phenology. *Remote Sensing of Environment*, 115, 1301–1316.

Pengra, B. W., Johnston, C. A., & Loveland, T. R. 2007. Mapping an invasive plant, *Phragmites australis*, in coastal wetlands using the EO-1 Hyperion hyperspectral sensor. *Remote Sensing of Environment*, 108(1), 74–81.

Powers, R. P., Hay, G. J., & Chen, G. 2011. How wetland type and area differ through scale: A GEOBIA case study in Alberta's Boreal Plains. *Remote Sensing of Environment*, 117, 135–145.

Prigent, C., Matthews, E., Aires, F., & Rossow, W. B. 2001. Remote sensing of global wetland dynamics with multiple satellite data sets. *Geophysical Research Letters*, 28(24), 4631–4634.

Richter, R. 1996. Atmospheric correction of satellite data with haze removal including a haze-clear transition region. *Computers & Geosciences*, 22(6), 675–681.

Richter, R., Schlapfer, D., & Muller, A. 2006. An automatic atmospheric correction algorithm for visible-NIR imagery. *International Journal of Remote Sensing*, 27(10), 2077–2085.

Rokitnicki-Wojcik, D., Wei, A., & Chow-Fraser, P. 2011. Transferability of object-based rule sets for mapping coastal high marsh habitat among different regions in Georgian Bay, Canada. *Wetlands Ecology and Management*, 19(3), 223–236.

Roulet, N. T. 2000. Peatlands, carbon storage, greenhouse gases, and the Kyoto protocol: Prospects and significance for Canada. *Wetlands*, 20(4), 605–615.

Rouse, J.W.J., Haas, R.H., Schell, J.A., Deering, D.W., 1974. Monitoringveg etation systems in the Great Plains with ERTS. In: *Third ERTS Symposium, NASA SP-351*, Washington, DC, pp. 309–317.

Sawaya, K., Olmanson, L., Heinert, N. J., Brezonik, P. L., & Bauer, M. E. 2003. Extending satellite remote sensing to local scales: Land and water resource monitoring using high-resolution imagery. *Remote Sensing of Environment*, 88, 144–156.

Schmid, T., Koch, M., & Gumuzzio, J. 2005. Multisensor approach to determine changes of wetland characteristics in semiarid environments (Central Spain). *IEEE Transactions on Geoscience and Remote Sensing*, 43(11), 2516–2525.

Semlitsch, R. D., & Bodie, J. R. 1998. Are small, isolated wetlands expendable? *Conservation Biology*, 12(5), 1129–1133.

Shanmugam, P., Ahn, Y. H., & Sanjeevi, S. 2006. A comparison of the classification of wetland characteristics by linear spectral mixture modelling and traditional hard classifiers on multispectral remotely sensed imagery in southern India. *Ecological Modelling*, 194, 379–394.

Thomlinson, J. R., Bolstad, P. V., & Cohen, W. B. 1999. Coordinating methodologies for scaling land cover classifications from site-specific to global: Steps toward validating global map products. *Remote Sensing of Environment*, 70, 16–28.

Verhoeven, J. T. A., & Setter, T. L. 2010. Agricultural use of wetlands: Opportunities and limitations. *Annals of Botany*, 105, 155–163.

Walsh, S. J., McCleary, A. L., Mena, C. F., Shao, Y., Tuttle, J. P., González, A., & Atkinson, R. 2008. QuickBird and Hyperion data analysis of an invasive plant species in the Galapagos Islands of Ecuador: Implications for control and land use management. *Remote Sensing of Environment*, 112(5), 1927–1941.

Wang, L., Sousa, W. P., & Gong, P. 2004. Integration of object-based and pixel- based classification for mapping mangroves with IKONOS imagery. *International Journal of Remote Sensing*, 25(24), 5655–5669.

Wei, A., & Chow-Fraser, P. 2007. Use of IKONOS imagery to map coastal wetlands of Georgian Bay. *Fisheries*, 32(4), 164–173.

Wright, C., & Gallant, A. 2007. Improved wetland remote sensing in Yellowstone National Park using classification trees to combine TM imagery and ancillary environmental data. *Remote Sensing of Environment*, 107, 582–605.

Yang, J., He, Y., & Weng, Q. 2015. An automated method to parameterize segmentation scale by enhancing intra-segment homogeneity and inter-segment heterogeneity. *IEEE Geoscience and Remote Sensing Letters*, 12(6), 1282–1286.

Yu, Q., Gong, P., Clinton, N., Biging, G., Kelly, M., & Schirokauer, D. 2006. Object-based detailed vegetation classification with airborne high spatial resolution remote sensing imagery. *Photogrammetric Engineering & Remote Sensing*, 72(7), 799–811.

Yuan, L., & Zhang, L. 2006. Identification of the spectral characteristics of submerged plant *Vallisneria spiralis*. *Acta Ecologica Sinica*, 26(4), 1005–1010.

Zedler, P. H. 2003. Vernal pools and the concept of "isolated wetlands." *Wetlands*, 23(3), 597–607.

Zedler, J. B., & Kercher, S. 2005. Wetland resources: Status, trends, ecosystem services, and restorability. *Annual Review of Environmental Resources*, 30, 39–74.

Zhang, Y., Guindon, B., & Cihlar, J. 2002. An image transform to characterize and compensate for spatial variations in thin cloud contamination of Landsat images. *Remote Sensing of Environment*, 82, 173–187.

Zomer, R. J., Trabucco, A., & Ustin, S. L. 2009. Building spectral libraries for wetlands land cover classification and hyperspectral remote sensing. *Journal of Environmental Management*, 90(7), 2170–2177.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

15 Geomorphic and Biophysical Characterization of Wetland Ecosystems with Airborne LiDAR

Concepts, Methods, and a Case Study

Murray Richardson and Koreen Millard

CONTENTS

15.1 INTRODUCTION

Airborne LiDAR is an active remote sensing technology that has been used for a wide variety of applications in wetland science and management. It allows for rapid collection of highly detailed topographic and vegetation structural characteristics over large, remote areas where ground-based observations are often limited due to poor accessibility. Unlike other sources of high-resolution elevation data such as satellite stereophotogrammetry, high-frequency (e.g., 25–150 kHz) laser pulses can, in many circumstances, penetrate vegetation canopies, providing subcanopy topographic information as well. Lower quality, coarser resolution topographic data sources do not capture subtle topographic and morphological gradients typical of low-gradient wetland environments with sufficient detail for scientific research and management purposes. Thus, airborne LiDAR and related technologies have led to new opportunities for studying wetlands in ways that were previously not possible with coarser resolution data sources. Indeed, since approximately the early 2000s, numerous studies have been published on the topic, demonstrating a wide variety of wetland-related applications for airborne LiDAR surveys, including classification and mapping, geomorphological analysis, vegetation analysis, hydrologic analysis, flood susceptibility modeling in coastal marshes, and wetland biogeochemistry. By all accounts, the use of airborne LiDAR in wetland science is growing rapidly, as are the development and application of newer technologies that can also generate three-dimensional point cloud information with similar or greater levels of detail, including terrestrial laser scanning and modern aerial photogrammetric techniques such as structure from motion (SfM).

Wetlands present unique challenges associated with the analysis of LiDAR point clouds and DEMs, however, and require different analytical considerations compared to applications in other environments such as urban areas, high-relief landscapes, and nonvegetated terrain. For example, microtopographic surface morphology can give rise to local-scale (e.g., distances on the order of 0.5–10 m) elevation gradients that are of similar magnitude to broader, mesoscale, topographic gradients (e.g., 10–1000 m distances) within the same wetland or wetland complex. Short wetland vegetation (e.g., shrub and graminoid species) can also be problematic in the point cloud classification step, resulting in topographic errors and inaccuracies associated with vegetation indices (Hopkinson et al. 2005). Finally, hydrologic derivatives and

digital elevation model (DEM) preprocessing techniques designed for nonwetland landscapes may produce unexpected or unwanted results when applied to low-gradient environments.

Despite some of these unique challenges, novel approaches to the analysis of airborne LiDAR surveys have resulted in significant advances in wetland science. The purpose of this chapter, therefore, is to provide an overview of analytical methods and issues associated with the analysis of airborne LiDAR for wetland research and management applications. We begin with an overview of typical LiDAR products within the context of three major wetland-related application areas, including (1) wetland vegetation analysis, (2) hydrologic and geomorphic analysis of wetlands, and (3) wetland detection and ecosystem classification. This is followed by a synthesis of accuracy issues related to ground surface topography and vegetation analysis in wetland environments. The chapter concludes with an illustrative case study using airborne LiDAR and field observations of topography and surface hydrology from a northern peatland complex in the James/Hudson Bay Lowlands, northern Ontario, Canada.

15.2 LiDAR DATA PRODUCTS, DERIVATIVES, AND THEIR APPLICATIONS IN WETLAND SCIENCE AND MANAGEMENT

15.2.1 Discrete Return Airborne LiDAR in Vegetated Environments

The vast majority of wetland studies involving airborne LiDAR data analysis have made use of discrete return surveys rather than full waveform surveys. A discrete return LiDAR sensor emits a pulse of laser energy (light), and the time required for the pulse to return to the system from the target is recorded (Lim et al. 2003). Through data processing, one to four discrete returns per pulse can be extracted, and for each pulse, X, Y, and Z values and intensity values (that correspond to a laser pulse's backscattered amplitude) can also be recorded. Since the speed of light is known, using the time for the pulse to travel from the sensor to the ground and back, the distance between the sensor and the target can be determined. Using an integrated differential global positioning system (GPS), scanning mirrors, and an inertial measurement unit (IMU), the precise location of each elevation point in a LiDAR swath can be determined (Wehr and Lohr 1999). Airborne LiDAR data are collected in flight lines (with overlap) that are merged together to make one point cloud, which, depending on the size of the area scanned and the density of points, could be hundreds of thousands or millions of points. As the laser pulse is emitted from the sensor, it diverges (Figure 15.1), resulting in a footprint on the ground, the size of which is dependent on wavelength and flying height (Hopkinson 2007). This means that as the beam travels through a vegetated canopy, it may encounter several objects that vary in height (i.e., elevation). It will first be reflected from the top of the canopy, sending a return back to the sensor. It will then be reflected from the intermediate canopy objects (e.g., branches, leaves), and should finally be reflected from the ground (Figure 15.1).

Most discrete return LiDAR systems used today are able to detect up to at least four returns (e.g., Optech ALTM 3100 or PEGASUS HA500). These returns are then

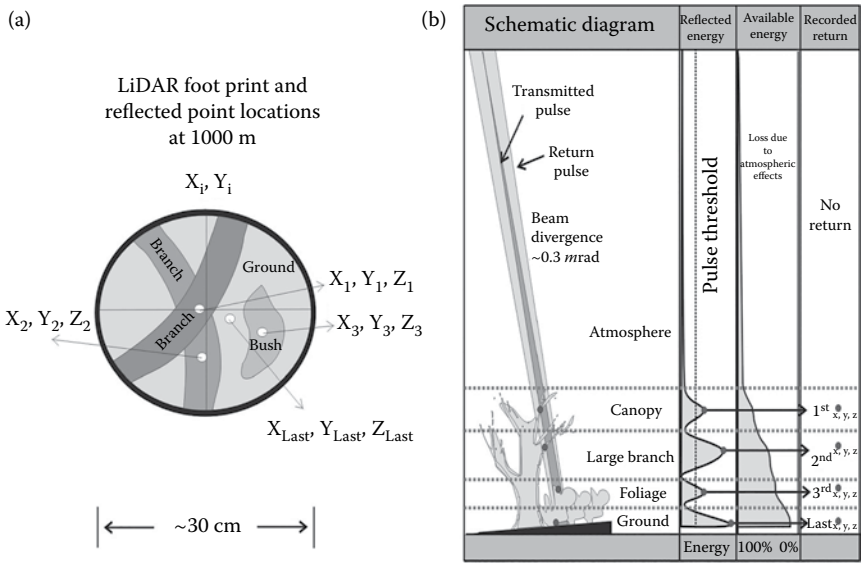


FIGURE 15.1 (a) Not all of the energy emitted from the pulse is reflected by a single object. Due to the divergence of the beam, the footprint of the pulse when it reaches the surface is larger than the pulse at the sensor. In this case, all objects within the 30 cm footprint reflect different amounts of transmitted pulse energy. The reflected energy is returned to the sensor by multiple objects within the footprint. (b) The wave is returned with varying amplitudes (reflected energy). By setting the pulse threshold, peak return amplitudes are output as point locations 1st, 2nd, 3rd, and Last. The Last return can be interpreted to be from the ground. [From Hopkinson, C. et al. 2005. *Canadian Journal of Remote Sensing* 31 (2):191–206; Millard, K. et al. 2009. *Geoarchaeology* 24 (5):576–588.]

classified as points that represent the surface of the earth (ground points) and those that represent other features such as vegetation (nonground, vegetation, buildings) (Wehr and Lohr 1999). Once the data have been classified, the points are interpolated to create various rasterized topographic, hydrologic, and vegetation derivatives as described in Sections 15.2.1.3 and 15.2.1.4.

A summary is provided in Table 15.1 showing typical LiDAR data products and derivatives that are used in wetland research applications. Many of the examples are not exclusive to wetland applications. We have grouped them into four broad categories, as described in the following subsections.

15.2.1.1 Original Data and Classified Point Clouds

LiDAR survey providers will commonly deliver classified point clouds and gridded products interpolated from ground- and/or vegetation-classified returns and laser return intensity. Methods for point cloud classifications are often proprietary, although it has become easier for researchers to implement classifications of raw point clouds using various software tools such as LAStools (https://rapidlasso.com/lastools/) and CloudCompare (http://www.danielgm.net/cc/). Typically, point clouds are interpolated or otherwise analyzed to produce gridded products that may

TABLE 15.1
Summary Table of Typical Airborne LiDAR Data Types and Derivatives Used for Wetland-Related Applications

	Data Format or Derivative	Description and Wetland-Related Application	References for Wetland-Related Applications
Original point cloud	All returns (x, y, z, i)	–	–
Classified point clouds	Bare earth returns (x, y, z, i)	Discrete LiDAR returns representing the ground surface after removing returns from any aboveground features	Richardson et al. (2009)
	Aboveground returns (x, y, z, i)	Discrete LiDAR returns from aboveground features only	–
Interpolated products	Interpolated digital surface model (DSM)	Interpolated grid of ground and aboveground feature elevations (typically 0.5–5 m resolution)	
	DEM or digital terrain model (DTM)	Interpolated grid of ground surface elevations based on bare earth returns (typically 0.5–5 m resolution); used to derive various geomorphometric and hydrologic derivatives	e.g., Rapinel et al. (2015)
	Interpolated intensity grid	Interpolated grid of laser pulse return intensity; sensitive to surface conditions including wetness	e.g., Chust et al. (2008)
Topographic, geomorphometric, and hydrologic derivatives	Residual analysis of grids	Relations of each grid cell to its neighbors (e.g., difference from mean elevation or standard deviation of values in a circular or rectangular window of specific size); used to provide contextual topographic information at a particular scale	e.g., Difebo et al. (2015)
	Terrain ruggedness	Local topographic heterogeneity within a user-defined grid neighborhood; can be used to describe microtopography and vegetation roughness on very high resolution ($\leq$ 1 m) DEMs or DSMs	e.g., Millard and Richardson (2013)
	Ratio of flow path length to gradient	Ratio of flow path length to gradient at any point in the landscape, relative to user-defined target (e.g., stream or watershed outlet); proxy indicator of spatial differences in water transit times to stream network and a descriptor of catchment morphology	e.g., Richardson et al. (2012)

(*Continued*)

TABLE 15.1 (*Continued*)
Summary Table of Typical Airborne LiDAR Data Types and Derivatives Used for Wetland-Related Applications

	Data Format or Derivative	Description and Wetland-Related Application	References for Wetland-Related Applications
	Topographic wetness index	Ratio of upslope contributing area to local slope; an index of the spatial distribution of terrain wetness and depth to water table; includes variants depending on method of slope calculation	e.g., Richardson et al. (2012), Lang et al. (2013), Rampi et al. (2014), Murphy et al. (2009)
	Downslope drainage index (Hjerdt et al. 2004)	Approximation of spatial distribution of local hydraulic gradient based on downslope drainage conditions	e.g., Richardson et al. (2009)
	Depth to water table index	Approximation of spatial distribution, local depth to water table based downslope topography from each pixel in the landscape to a target surface water pixel	e.g., Murphy et al. (2007)
	Lagg width and slope indices	The average lagg width for raised bog form and the average slope from the lagg to the top of the bog dome ("rand slope")	e.g., Richardson et al. (2010)
	Depression probability	Likelihood that an identified depression in a DEM is a naturally occurring feature; can be used to discriminate depressional wetlands	e.g., Lindsay et al. (2004), Creed and Sass (2011)
Vegetation derivatives	Vegetation height (maximum, minimum, mean)	Vegetation height measurements above ground, where height above ground is determined based on the height above ground classified LiDAR points	Hopkinson et al. (2005), Lefsky et al. (2002)
	Aboveground biomass	LiDAR point clouds provide a representation of forest structure that can be used in statistical models and compared with field measurements to estimate biomass	Englhart et al. (2013), Boehm et al. (2013), Riegel et al. (2013)

(*Continued*)

TABLE 15.1 (*Continued*)
Summary Table of Typical Airborne LiDAR Data Types and Derivatives Used for Wetland-Related Applications

Data Format or Derivative	Description and Wetland-Related Application	References for Wetland-Related Applications
Leaf area index (LAI)	LAI (the total one-sided area of green foliage per unit ground surface) will affect the penetration capabilities of LiDAR into the canopy and has been shown to be a function of vegetation height and other vegetation metrics	Luo et al. (2015)
Ratio of ground returns to all returns	If the ratio of ground to all hit points is high, vegetation density is high; conversely, if the ratio is low, few laser pulses were able to reach the ground, likely due to dense vegetation	Millard and Richardson (2013)
Standard deviation of all returns	In areas of taller vegetation, the standard deviation of the LiDAR points will be greater than in areas with shorter vegetation and therefore a multiplier of the standard deviation of all returns can be used to derive models based on field data	Millard and Richardson (2013), Hopkinson et al. (2005)
Canopy height model (CHM)	A surface representing the top of the canopy (often simply the first return)	Hopkinson et al. (2005)
Canopy cover or vegetation cover	The percentage of vegetation coverage within a pixel (usually calculated based on a group of neighboring pixels)	Millard and Richardson (2015)

represent ground or aboveground elevations or laser return intensity (interpolated products), or other statistical properties and compound derivatives for hydrologic and geomorphic analysis (topographic, geomorphometric, and hydrologic derivatives) and vegetation characterization (vegetation derivatives). In a few cases, researchers analyze the classified ground or canopy returns directly, rather than interpolating to an arbitrary grid size, which has the advantage of using the entire distribution of available points (e.g., Richardson et al. 2010).

15.2.1.2 Interpolated Products

The resolution of gridded products is often dependent on the particular application, as well as the original point cloud density. The latter is a function of surveying parameters, such as flight altitude and laser pulse frequency, and surface characteristics, including vegetation density. DEMs and digital surface models (DSMs) are interpolated from classified bare earth and all return point clouds, respectively. Typical LiDAR-derived DEM and DSM resolutions range from 0.5–5 m. The choice of resolution typically depends on point density of the LiDAR survey, but also on computational limitations. Laser return intensity (i) is sometimes interpolated to a gridded product and used to infer ground or vegetation surface properties. The analysis of DEMs, DSMs, and interpolated intensity data for wetland applications is common. However, other derivatives extracted from point clouds and/or DEMs and DSMs are more frequently used because they convey additional, important biophysical or hydrologic information about wetland types and characteristics.

15.2.1.3 Topographic, Geomorphometric, and Hydrologic Derivatives

These derivatives are typically produced through a moving-window (neighborhood) analysis of the interpolated bare earth DEM. The information content of these derivatives can vary substantially as a function of the input DEM resolution, and low-pass smoothing filters are typically applied to limit the effect of surface roughness and noise, particularly at higher (e.g., $<$1 m) pixel resolution (Richardson et al. 2009; MacMillan et al. 2003). For example, slope and aspect are commonly calculated from LiDAR DEMs and can provide information on local topographic gradient and solar insolation, although their utility in low-relief wetland environments is limited and generally not applicable due to the lack of larger scale relief that is required to make these localized metrics useful. Hydrologic derivatives typically involve use of topographically driven flow routing algorithms to predict water movement of flow concentration through landscapes and have been applied in wetland environments, along with drainage and connectivity metrics (see Section 15.2.2.1). Hydrologic derivatives are sensitive to DEM resolution, hydrologic preprocessing techniques, and algorithmic choices, as demonstrated in the case study.

15.2.1.4 Vegetation Derivatives

Vegetation derivatives are normally derived through analysis of either the all returns point cloud or the aboveground returns point clouds. Early adopters (e.g., Genc et al. 2005) of LiDAR data in vegetation research often simply considered the first return (i.e., the return with the shortest round-trip travel time) as the top of the canopy, and the last return (i.e., the return with the longest round-trip travel time) to be the

ground. In several circumstances, this method works well (e.g., bare ground or sparse or open canopy), but in areas with dense understory vegetation it has been shown that lack of penetration into the understory can cause significant error in capturing the true ground surface (Hopkinson et al. 2005). More sophisticated algorithms to separate ground from vegetation points have been created, including multiscale curvature classification (MCC) (Evans and Hudak 2007), and the Boise Centre for Aerospace Laboratory LiDAR algorithm (BCAL) (Streutker and Glenn 2006). These algorithms have been found to produce similar overall accuracy results, although BCAL performed better in areas of dense vegetation whereas MCC performed better where steep changes to topography were present (Tinkham et al. 2011). Other, simpler vegetation derivatives include gridded products that represent a statistical property of the point cloud (minimum, maximum, and standard deviation of return elevations), or ratios (e.g., ratio of ground return count to all-return count) within the areal unit defined by each pixel (Table 15.1).

15.2.2 Applications

15.2.2.1 Wetland Ecosystem Classification and Detection

Wetlands are characteristically low-gradient environments with distinct vegetation communities and hydrologic settings. Despite having very low relief, topographic structure is also a defining characteristic of different types of wetlands. For example, ombrotrophic peatlands develop characteristic surface morphologies that reflect the local topographic and hydrologic setting and the long-term process of ecological succession (Damman 1986). Vegetation communities similarly reflect hydrologic and nutrient conditions in wetlands and most classification systems use vegetation as a key discriminator of wetland type. High-resolution, LiDAR-derived topographic and vegetation information can therefore provide important information on wetland occurrence and type, and is now frequently used on its own or in conjunction with other imagery sources to classify and map wetlands over large areas. In this section, we have synthesized literature pertaining to wetland ecosystem classification (i.e., distinguishing different types of wetlands and/or their characteristics within a wetland complex or landscape), *and* wetland detection (i.e., binary landscape classifications showing distribution of wetlands). A tabulated summary is also provided (Table 15.2).

Numerous authors have demonstrated that LiDAR topographic and vegetation derivatives can be incorporated into image classification procedures to improve the thematic classification accuracy of wetland maps compared to the use of aerial photographs and satellite-based optical or radar imagery alone. In one of the earliest examples of this, Maxa and Bolstad (2009) combined elevation, slope, and terrain shape (curvature) grids derived from a 1 m resolution LiDAR DEM with 1 m resolution pan-sharpened IKONOS red/green/blue (RGB) imagery in a manual classification effort and reported better distinction of upland versus wetland classes in a forested landscape compared to the interpretations based on higher resolution aerial photographs. Thematic classification accuracies for maps created through automated, supervised, and unsupervised classifiers can also be improved through the incorporation of LiDAR topographic, hydrologic, and vegetation derivatives. LiDAR

TABLE 15.2
Tabulated Summary of Studies in Which LiDAR Imagery Was Used in Wetland Ecosystem Classification and Wetland Detection (Presence/Absence mapping)

	Reference	Wetland Type	Imagery Type(s)	Algorithms
Wetland ecosystem classification	Chadwick (2011)	Mangrove	LiDAR + IKONOS	Maximum likelihood
	Morris et al. (2005)	Salt marsh	LiDAR + multispectral aerial photographs	Artificial neural network
	Gilmore et al. (2008)	Salt marsh	LiDAR + QuickBird	Object-oriented classification
	Moffett et al. (2012)	Salt marsh	LiDAR + IKONOS	Object-oriented classification
	Brennan and Webster (2006)	Intertidal wetlands and non-intertidal submerged vegetation	LiDAR + aerial photographs	Object-oriented classification
	Chust et al. (2008)	Coastal wetland	LiDAR + aerial photographs + LiDAR intensity	Supervised classification (unspecified)
	Rapinel et al. (2015)	Floodplain marshes	LiDAR + multispectral aerial photographs	Object-oriented classification
	Onojeghuo and Blackburn (2011)	Coastal reed beds	LiDAR + LiDAR intensity + airborne hyperspectral	Maximum likelihood
	Korpela et al. (2009)	Peatland (mire)	LiDAR only	Random forest, support vector machine (SVM), *k*-nearest neighbor (*k*-NN)

(*Continued*)

TABLE 15.2 (*Continued*)
Tabulated Summary of Studies in Which LiDAR Imagery Was Used in Wetland Ecosystem Classification and Wetland Detection (Presence/Absence mapping)

	Reference	Wetland Type	Imagery Type(s)	Algorithms
	Anderson et al. (2010)	Peatland	LiDAR only	Maximum likelihood
	Torbick et al. (2012)	Peatland	LiDAR + ALOS PALSAR	Random forest, SVM, *k*-NN
	Millard and Richardson (2013)	Peatland complex	LiDAR + synthetic-aperture radar (SAR) polarimetry	Random forest
	Genc et al. (2005)	Freshwater marshes	LiDAR + Landsat + IKONOS	Unsupervised ISOdata
	Difebo et al. (2015)	Peatland complex	LiDAR + multispectral aerial photographs	Supervised ISOdata
	Suchenwirth et al. (2012)	Floodplain wetland	LiDAR + orthophotos + IKONOS	Object-oriented classification
	Franklin and Ahmed (2017)	Peatland complex	LiDAR + SAR	Object-oriented classification
	Millard and Richardson (2015)	Peatland complex	LiDAR only	Random forest
	Zhang (2014)	Tropical wetlands	LiDAR + hyperspectral	Random forest, SVM, *k*-NN
Wetland detection	Hogg and Holland (2008)	Freshwater wetlands	Lidar + multispectral aerial photographs + 20 m DEM	CART
	Richardson et al. (2009)	Forested wetlands	LiDAR only	CART/DTA thresholding
	Creed et al. (2003)	Forested wetlands	LiDAR only	DTA thresholding
	Murphy et al. (2007)	Freshwater wetlands	LiDAR only	DTA thresholding
	Lang et al. (2013)	Forested wetlands	LiDAR only	DTA thresholding

Note: ISODATA = The iterative self-organizing data analysis technique, CART = Classification and regression trees, DTA = Digital terrain analysis.

derivatives provide important contextual information about topographic position and hydrologic setting that improve statistical separability of classes, something that is particularly important for pixel-based image classifiers. For example, Difebo et al. (2015) created an ecosystem map with eight thematic classes for a northern peatland complex using supervised maximum likelihood classification with LiDAR and 4 m resolution IKONOS imagery. LiDAR information was incorporated by calculating an index of relative topographic position, difference from mean elevation (DME), of each focal pixel at multiple window sizes ranging from 15 × 15 to 300 × 300 pixels. The largest window size was found to improve the independently validated classification accuracy by approximately 10% compared to the use of IKONOS imagery alone. The characteristic dome morphologies of raised bogs in this landscape were clearly captured within the LiDAR DEM, and this information could be exploited by the image classifier to improve class separability, provided an appropriate scale was used to calculate the derivative.

Millard and Richardson (2013) combined a wide range of derivatives from both LiDAR and RADARSAT-2 synthetic-aperture radar (SAR) imagery to produce an ecosystem classification for Mer Bleue Bog in southeastern Ontario, Canada, using a random forest (RF) classifier. Only the LiDAR derivatives (including vegetation, hydrologic, and topographic) were found to contribute meaningful information to the mapped classes, as interpreted from the RF variable importance metrics. In a subsequent study, the same researchers were able to map a nearby peatland complex, Alfred Bog, into five thematic classes with an overall, independently validated classification accuracy of 73%, using only nine LiDAR derivatives, after removing redundant input variables (Millard and Richardson 2015). RF classification was also used in that study, and the variable importance measures indicated the top three most important LiDAR derivatives to be valley depth (a topographic derivative) and two vegetation derivatives including standard deviation of vegetation height and maximum vegetation height.

Chust et al. (2008) combined various LiDAR derivatives, including basic topographic metrics and laser return intensity, with four-band multispectral [RGB near-infrared (NIR)] aerial photography for classification of estuarine, coastal habitats. LiDAR intensity data required a smoothing filter to reduce speckle, but were found to improve supervised classification of wetland habitats by ∼8% compared to use of RGB-NIR imagery alone. In a similar coastal mapping study, Gilmore et al. (2008) successfully exploited differences in vegetation community structure by incorporating LiDAR-derived canopy heights into an object-oriented classification routine. For thematic mapping of an herbaceous, valley-bottom wetland into nine classes using a multisensor approach, Rapinel et al. (2015) achieved the highest accuracies (85.5%) only after inclusion of LiDAR-derived elevation, vegetation height, and laser return intensity. Thus, there is ample support in the literature for using airborne LiDAR derivatives to improve wetland mapping products. Moreover, when LiDAR topographic and vegetation derivatives are used as input into image classifiers, their distributions can be visualized within each mapped class and potentially used as ecological indicators of wetland function since they reflect spatial variations in geomorphic and biophysical characteristics and their underlying ecohydrological drivers (Millard and Richardson 2013).

Numerous studies have also used LiDAR DEMs for detection and mapping of small wetlands in forested landscapes. The analytical approaches used tend to differ from the ecosystem classification approaches described previously. Specifically, these approaches typically involve deriving contextual hydrologic information related to topographic setting, local drainage, and wetness conditions. An early example is by Creed et al. (2003), who mapped small subcanopy or cryptic wetlands in a boreal shield landscape in Ontario, Canada, using a depression detection algorithm. A related method based on a stochastic approach to depression filling (Lindsay and Creed 2006) has been subsequently applied for detection of small wetland features and vernal pools in forested and glaciated landscapes of North America (Creed et al. 2008; Wu et al. 2014). In a similar landscape setting, Richardson et al. (2009) calculated the downslope drainage index (Hjerdt et al. 2004) on a 1 m resolution LiDAR DEM and applied an edge detection algorithm to segment the landscape according to drainage conditions. The segments were thresholded with a decision tree model using field-based observations as training data to map small forested wetlands and vernal pools. Hogg and Holland (2008) used an estimate of groundwater loading potential derived from a DEM and surficial geology information (Baker et al. 2003), combined with elevation and profile curvature information. These derivatives were calculated using a coarser resolution DEM, followed by a LIDAR-derived DEM, and the latter improved accuracy of a decision tree classifier from 76% to 84%.

The topographic wetness index (TWI) (Beven and Kirkby 1979) is commonly used for mapping saturated areas and wetlands. TWI is a compound hydrologic derivative derived from the ratio of upslope contributing area (α) to local slope (β). There are many different methods for calculating TWI, and most variants differ on the basis of how α is calculated, with some allowing for more distributed flow than others (Rampi et al. 2014). With high-resolution DEMs (e.g., < 5 m), this variable can be particularly important since nondistributed algorithms can produce artificially linear flow concentrations through low-relief landscapes. Several studies have shown that when computing TWI from LiDAR-derived DEMs for mapping forested wetlands, methods that produce more distributed flow (e.g., multiple flow direction versus single flow direction algorithms) are preferable (Rampi et al. 2014; Lang et al. 2013). The case study in Section 15.4 of this chapter also presents a similar analysis for the purpose of upscaling point measurements of near-surface soil moisture in a northern peatland complex.

15.2.2.2 Hydrologic and Geomorphic Analysis of Wetlands

In addition to their utility for wetland detection and classification, LiDAR terrain derivatives can also be used to infer hydrological, ecological, and biogeochemical processes occurring within wetlands with remote sensing. Topographic structure imparts a strong control on hydrologic flow paths and spatial patterns of inundation in wetlands, nutrient availability, and vegetation communities. Moreover, structural characteristics of wetlands, including topographic gradients, evolve over time through ecohydrological feedback mechanisms, most notably in peatlands (Bridgham et al. 1996). LiDAR imagery and derivatives can be used in this process, allowing researchers to link high-resolution spatial *patterns* to underlying ecological *processes*,

thus improving understanding of wetland hydrological and biogeochemical functions. In this section, we discuss examples of this from the scientific literature.

In forested and ombrotrophic peatlands, LiDAR topography data have been successfully used to quantify geomorphic form, including the marginal lagg zone that serves as an ecologically important transitional ecotone between upland peatland ecosystems. The lagg area is typically nutrient enriched and wetter than the central areas of the peatland due to localized inputs from the surrounding upland area. These abiotic factors result in distinct differences in vegetation with the lagg zone. Richardson et al. (2010) found that LiDAR ground returns could be analyzed to quantify the average lagg width of northern forested peatlands in northern Minnesota and southcentral Ontario, Canada. They found LiDAR-derived lagg widths to increase as a function of the ratio of upslope contributing area to hydraulic gradient (as estimated from LiDAR). Moreover, they found that porewater chemistry corresponded spatially with the LiDAR-defined lagg zone, including higher concentrations of sulfate, dissolved organic carbon, and mercury. Langlois et al. (2015) analyzed the topographic structure of six different ombrotrophic peatlands in eastern Canada and classified lagg features as either confined or unconfined transitional areas. Confined laggs were topographically depressed relative to the adjacent upland slope and bog areas, whereas unconfined laggs were adjacent to flatter or receding (sloping away from the lagg) upland areas. Water levels and nutrient conditions were distinctly different between these two types of lagg areas due to the abiotic control of topography, with corresponding influences on vegetation. More recently, Langlois et al. (2017) used a moving split-window analysis of LiDAR-based topographic and vegetation derivatives to detect ecological gradients and boundaries occurring with confined and unconfined laggs of these same study sites. The study showed that ecological gradients and boundaries were reflected in the LiDAR derivatives, but boundary detection varied as a function of scale, namely, the size of the moving split-window applied.

LiDAR topography has been used for hydrologic analyses and simulation in a wide variety of wetland environments including tidal salt marshes, deltaic wetlands, northern peatlands, forested wetlands, and prairie wetlands. Due to the exceptionally low relief conditions of such environments, these types of analyses were generally not possible prior to the availability of airborne LiDAR. One of the earliest examples was by Töyrä et al. (2003), who assessed the accuracy of LiDAR DEMs for hydrologic applications in the 3900 km^2 Peace–Athabasca Delta in northwestern Canada. Despite elevation biases and random errors associated with different vegetation classes, they concluded that LiDAR ground surface topography could be useful for many hydrologic applications in this low-gradient environment.

Various studies demonstrate use of LiDAR DEMs to improve understanding of hydrologic processes and functions of wetlands such as flooding, water storage, flow path connectivity, and runoff generation. In coastal salt marshes, airborne LiDAR has helped researchers map and model tidal inundation zones and associated vegetation communities (Morris et al. 2005). Knowledge gained from such analyses can be used to guide restoration efforts in disturbed coastal marshes (Millard et al. 2013) or to predict ecosystem impacts of sea level rise due to climate warming (Moeslund et al. 2011). Lane and D'Amico (2010) demonstrated potential to estimate storage volumes

in isolated wetlands of Florida for quantifying and modeling hydrologic ecosystem services at a landscape scale. LiDAR DEMs have been used to estimate water storage capacity of depressional wetlands in Prairie Pothole regions of North America and to improve parameterization and output of hydrologic simulation models (Huang et al. 2013, 2011; Wu and Lane 2016). Richardson et al. (2012) used quantitative analysis of LiDAR to characterize the near-stream zone of different size streams in a northern peatland complex in the James/Hudson Bay Lowlands in Ontario, Canada, and identified a scale effect that helped interpret differences in runoff-generating potential during high flow periods as a function of watershed size. In this low-gradient, peatland-dominated landscape, approaches for characterizing the dynamic storage capacity of natural ponds and depressions, as described by Wu and Lane (2016), hold potential for improving the representation of the dynamic hydrologic connectivity of discrete landscape units and their effects on watershed-scale runoff response (Quinton and Roulet 1998). Moreover, there is a need to test effectiveness of different topographic flow routing algorithms in large-wetland complexes, including their sensitivities to DEM preprocessing, DEM resolution, and algorithm types since most work of this nature has focused on wetland detection in landscapes with larger scale relief (e.g., Lang et al. 2013, 2012; Wu et al. 2014).

15.2.2.3 Wetland Vegetation Analysis

Vegetation type, height, biomass, and structure often covary with many other ecological and/or hydrological gradients in wetland environments. Also, many biophysical and climate models (e.g., energy flux) have been shown to be sensitive to errors in vegetation parameterizations and require accurate measurements of vegetation height (Hopkinson et al. 2005). In particular, wetlands have not been well represented in these models, often with data being extrapolated from a few point-scale measurements across large areas. Therefore, there is great interest in spatially measuring wetland vegetation characteristics, and many authors have demonstrated opportunities and challenges associated with the use of LiDAR for such purposes.

The ability of LiDAR to provide measurements for the derivation of vegetation structural information, including height, canopy closure, and vertical foliage distribution, has been assessed in a variety of environments including various wetland environments. The first successful measurement of vegetation height in a wetland environment using LiDAR showed a statistical separation between heights of wetland classes as measured through a profile (nonscanning) LiDAR (Jensen et al. 1987). LiDAR technology has advanced significantly from its early states, and now scanning LiDAR can be used to measure these parameters with near-survey-grade levels of precision (better than 0.01 m) and accuracy (better than 5 cm). Outside of wetland environments and especially in the forestry sector, LiDAR is now regularly used for measuring forest and tree heights, and these measurements are highly correlated with field observations of height, diameter at breast height, and canopy characteristics (White et al. 2016). In wetland environments, there are fewer published results. Most analyses have focused on relating field measurements to raw LiDAR points (Millard et al. 2008) or LiDAR derivatives such as canopy height models or canopy cover (Lefsky et al. 2002). Since originally reported by

Bradbury et al. (2005), many authors have found relationships between the standard deviation of the LiDAR all-hits data set with field measurements of vegetation height, including in wetlands. When the standard deviation of LiDAR return elevations is high, there is likely taller vegetation due to the variability in multiple returns reflected from the ground, top, and within the canopy. Hopkinson et al. (2005) discussed the limitations of using a multiplier of the standard deviation of the LiDAR pulse returns to estimate canopy height where standard deviation is low. Due to inherent noise in the LiDAR data, the standard deviation will always be positive (nonzero), and therefore height values will also be positive, even where there is no vegetation. Additionally, the standard deviation of pulses will increase with slope regardless of vegetation presence. However, they noted that adjustments can be made if these vegetation-specific relationships are understood. Hopkinson et al. (2005) also produced frequency distributions of the first and last pulse in each of the vegetation classes, which can be used to assess the differences in vertical distribution of pulses within the canopy. Chadwick (2011) compared field measurements to LiDAR-derived canopy height model (CHM) in different mangrove classes and found that LiDAR systematically underestimates canopy height by up to 2 m for the tallest trees. Similar to Hopkinson et al. (2005), they attributed this to LiDAR pulses missing the tree apexes, as well as an overestimation of ground height due to lack of full pulse penetration.

There are still many limiting factors and uncertainties related to LiDAR system settings and methods used to measure vegetation parameters (Hopkinson et al. 2005), and this information is not often published in studies and sometimes not provided by the data provider. In addition to limited penetration into dense vegetation, wetland environments present a unique challenge in the measurement of vegetation parameters with LiDAR: the wetness of the environment may influence the sensor's ability to measure pulses (Hopkinson et al. 2005). Infrared energy is absorbed by water, and therefore, in wet areas, weak laser backscatter from the saturated ground conditions often occurs. While these areas can be sometimes identified using the intensity of the return (Zhang 2008), laser backscatter returned from a wet ground surface will not be strong, and therefore laser returns will be biased to the foliage because they have a higher intensity of reflection (Hopkinson et al. 2005). LiDAR has been used to obtain regional estimates of aboveground biomass, with a variety of different LiDAR derivatives used to derive statistical correlations with field measurements. Ballhorn et al. (2012) estimated aboveground biomass in tropical peat swamp forests from LiDAR using comparison of height histograms with field plots.

15.3 ACCURACY OF LiDAR GROUND SURFACE TOPOGRAPHY AND VEGETATION INDICES IN WETLAND ENVIRONMENTS

Although airborne LiDAR surveys can produce dense three-dimensional point clouds of ground surface elevations and vegetation, it is important to appreciate the inherent positional (x, y) and elevation (z) inaccuracies, and how these errors are influenced by surface properties, such as surface roughness and vegetation conditions. This is particularly relevant for wetland applications because LiDAR elevation errors can represent high relative percent error given the very low relief

that is typical of wetland landscapes. This section provides a synthesis of knowledge related to accuracy issues in wetlands with respect to both ground elevation and vegetation information.

15.3.1 LiDAR Ground Surface Elevation Error in Wetland Environments

Various studies have been published on the topic of LiDAR-derived ground surface elevation accuracies in wetland environments (Table 15.1). Most studies report elevation errors as the standard deviation of the difference between LiDAR-derived versus field surveyed elevations, or the root-mean-squared error (RMSE) of predicted (LiDAR-derived) versus observed (surveyed) elevations.

Field surveyed elevations are typically collected using electronic theodolites or differential GPS [e.g., real-time kinematic (RTK) GPS] systems with millimeter- to centimeter-level horizontal and vertical precisions. Errors associated with these field survey instruments are normally considered negligible compared to the errors associated with the LiDAR system, provided the manual surveying was conducted correctly. Under ideal conditions (e.g., nonvegetated, hard ground surface such as roads or bedrock outcrops), a typical RMSE of discrete LiDAR ground return elevations is around 0.08 m. A large source of this error is the result of x, y positional inaccuracies, which are typically on the order of 0.5 m under ideal conditions (Wehr and Lohr 1999). These (x, y, z) positional errors propagate from errors in three-dimensional positioning of the LiDAR sensor on the aircraft.

LiDAR accuracy assessments are commonly conducted by comparing a survey point elevation to the elevation of a LiDAR ground return within some maximum separation distance ($\sim$0.5 m) to account for positional errors. Others are conducted on the basis of point-to-pixel or pixel-pixel comparisons, whereby pixel values represent interpolated elevations from two or more point measurements (either LiDAR points, survey points, or both). Areal interpolation improves accuracy of LiDAR elevations provided they are compared to manual survey measurements that have been interpolated at a similar scale (Richardson et al. 2012). This improvement in accuracy results from the averaging process, which results in smoothing of random error and natural variations in elevation surfaces, thus facilitating detection of subtle elevation gradients in wetland environments (Richardson et al. 2010). In Table 15.1, only point-to-point comparisons are shown for consistency. In those studies that reported both point-to-point and pixel-to-pixel comparisons, errors were lower for the interpolated (pixel-to-pixel) scenarios. From Table 15.1, the average standard deviation and RMSE of LiDAR ground surface elevations for studies reporting point-to-point comparisons is 0.13 m (0.09–0.22 m) and 0.16 m (0.07–0.37 m), respectively. In wetland environments, these errors are generally caused by (1) large local-scale spatial variability in ground surface elevations, and (2) vegetation conditions. Another potential source of ground surface elevation error results from ambiguities in identifying the true ground surface due to the presence of compressible surface material such as moss, peat, or floating vegetation mats. This latter source of error in not discussed further here but should be considered when designing topographic surveying protocols for validation of LiDAR-derived elevation products. For example, some researchers report using base plates with approximately 10 cm diameter to

present the surveying range pole from sinking too far into the ground surface (Schmid et al. 2011; Richardson et al. 2010).

Large local-scale variability in surface elevations on the order of $\sim 10^{-2}$ m to 10^{-1} m are typical in many noninundated wetlands such as peatlands due to the presence of microforms that result from ecohydrological feedbacks affecting vegetation growth and peat accumulation (Belyea and Baird 2006; Richardson et al. 2010). This surface roughness results in elevation errors due to the aforementioned errors in the x and y positioning of ground points since it is impossible to precisely colocate a discrete ground-return measurement with a survey point measurement. The error associated with this source of error should be proportional to the scale of surface roughness within a radius of each survey point that is dictated by the (x, y) position of the LiDAR system, described previously.

Elevation errors are strongly influenced by vegetation conditions in wetlands and tend to increase as a function of vegetation height and density. Vegetation may cause systematic error (bias) or random error (as indicated by RMSE or standard deviation). Ground surface elevation errors can also increase as a result of short, dense vegetation, which may either limit the number of bare earth returns due to poor laser penetration or cause difficulties in separating ground versus vegetation returns during point cloud classification. Elevation bias due to vegetation is normally positive, and often increases with vegetation height within a given study. Hopkinson et al. (2005) reported positive elevation biases in a boreal wetland complex on the order of 0.06 m–0.15 m, and random error ranging from 0.03 m to 0.22 m standard deviations. The highest bias and random error observations (0.15 m and 0.22 m, respectively) were for aquatic vegetation and this was attributed to the effect of standing water below the vegetation. Schmid et al. (2011) showed significant, positive correlation between the top of canopy height and the vertical error in LiDAR ground surface elevations in a coastal marsh environment, and error was exacerbated by vegetation density. They used a height-dependent correction factor to improve accuracy of LiDAR ground elevations for their study site.

15.3.2 Accuracy of LiDAR-Derived Vegetation Parameters

LiDAR is now commonly used to extract vegetation properties such as height, canopy cover, and biomass. Most available research has focused on forestry applications, with fewer examples of wetland-related applications. While this is likely due to the aforementioned issues with error in returns and/or issues with point cloud classification in short dense vegetation, several examples of measurement of vegetation height and vegetation biomass have been reported in a variety of different wetland classes. Hopkinson et al. (2005) reported a -0.24 m mean difference in aquatic canopy height (derived from the nearest canopy point) as compared with field measurements. Additionally, they found no relationship between canopy height derived from rasterized canopy data with field measurements in aquatic vegetation. Conversely, for a tropical wetland environment, Genc et al. (2005) used the difference between the canopy height surface and the DEM surface to produce a vegetation height raster and found a strong relationship between the two ($R^2 = 0.92$, $n = 16$).

In the case of extracting biophysical information in areas with short, dense wetland vegetation, one commonly used method has been to derive a relationship between the standard deviation of the all returns point cloud heights and field measurements. The standard deviation of vegetation heights is useful when considering the distribution of LiDAR vegetation data. LiDAR ground points and upper canopy points fall at the extremes of the distribution. Therefore, short vegetation should have a small standard deviation in height, whereas taller vegetation will have a larger standard deviation. In a salt marsh environment, Millard et al. (2008) improved models of vegetation height from LiDAR using this method by an order of magnitude (the relationship improved from an underestimation of 0.4–0.036 m). Hopkinson et al. (2005) found similar results in an aquatic marsh environment. More recently, Luo et al. (2015) used the same method to measure vegetation height and found a strong relationship (cross-validated $R^2 = 0.84$ and an RMSE of 0.14 m). Similar to Hopkinson et al. (2005), they reported that the relationship between a rasterized canopy height model and vegetation height was poor in short wetland vegetation. Luo et al. (2015) also investigated the ability to derive the leaf area index (LAI) from LiDAR and found that using predicted vegetation height resulted in a stronger relationship ($R^2 = 0.79$) than using a laser penetration index ($R^2 = 0.70$).

In boreal and tropical peatlands, several results have been reported on attempts to use LiDAR to estimate aboveground biomass. Lefsky et al. (2002) assessed the relationships between several different LiDAR vegetation derivatives and aboveground biomass in boreal forest, including muskeg and peatland areas. Canopy cover multiplied by mean canopy profile height resulted in the strongest relationship ($R^2 = 0.88$), but several other derivatives also resulted in similar relationships (e.g., canopy cover $R^2 = 0.84$). Englhart et al. (2013) and Boehm et al. (2013) both found strong relationships between LiDAR point clouds and aboveground biomass in tropical peatlands, but both Kronseder et al. (2012) and Riegel et al. (2013) found poor results using LiDAR point clouds alone ($R^2 = 0.41$ and 0.18, respectively). Riegel et al. (2013) improved the relationship by adding optical data into their regression, but relationships were generally still poor ($R^2 = 0.37$). In mangroves, LiDAR has been used to assess a variety of vegetation and canopy characteristics with high accuracy. For example, Wannasiri et al. (2013) were able to estimate crown diameter ($R^2 = 0.75$) and estimate individual tree heights ($R^2 = 0.80$). Zhang (2008) was able to detect small gaps in mangrove canopy due to lightning strikes with an average size of 0.59 m^2.

15.4 CASE STUDY: LiDAR ACCURACY ASSESSMENT AND DIGITAL TERRAIN ANALYSIS IN A NORTHERN PEATLAND COMPLEX, JAMES BAY LOWLANDS, ONTARIO, CANADA

15.4.1 Introduction, Study Site, and Objectives

This case study focuses on analysis of surface topography and hydrologic conditions in a large peatland complex in northern Ontario, Canada, using discrete return airborne LiDAR. The case study highlights some of the common issues associated with LiDAR applications in wetlands discussed in Sections 15.1 to 15.3, focusing

on LiDAR-derived ground elevation resolution and accuracy and the application of high-resolution ($\leq$ 5 m pixel dimensions) topographic wetness indices in low-gradient peatland environments.

The study site is located in the vicinity of the De Beers Victor Diamond Mine approximately 500 km northeast of Timmins, Ontario and 90 km west of Attawapiskat, Ontario, in the James Bay Lowland ecoregion of the Hudson Bay Lowlands ecozone (52.83° N, 83.93° W). The James/Hudson Bay Lowlands falls within a zone of sporadic discontinuous permafrost and represents the world's largest contiguous expanse of peatlands (Riley 2011). The construction of the De Beers Victor Diamond Mine, which initiated in 2006, facilitated access to this remote landscape for peatland scientists and other researchers. Regional soils consist of thick deposits of marine clay and clay till that are overlain by peat deposits, averaging approximately 2 m in thickness, and situated upon a locally karstic Silurian limestone aquifer known as the Attawapiskat Formation. The groundwater table is near or above the surface in most areas and is associated with development of a patterned peatland complex with an array of bogs and fens. Minerotrophic fens (e.g., ribbed, riparian, ladder) are topographically low lying, and typically portray directional seepage and/or convey water (Mitsch and Gosselink 2008; Quinton et al. 2003). Ombrotrophic bogs (domed, mound, flat) are marginally raised in elevation above the fens, and thus receive precipitation as their sole source of water and act as important water storage and release features (Sjörs 1959). The development of raised bogs in this region was initiated by the creation of interfluvial divides between drainage channels, and interfluve width imposes a constraint on bog growth and maximum height of domes above channel base heights (Glaser et al. 2004). Drainage networks in headwater systems are predominantly channel fens or fen water tracks. Fen water tracks run down the sloping sides of bogs at approximately right angles to the bog edges (Sjörs 1959). Channels along these drainages are intermittent to nonexistent, but during periods of high flow or during the freshet when ground ice is prevalent, overland flow can be observed between small, flowing pools (Richardson et al. 2012).

A discrete return airborne LiDAR survey was conducted over the Victor Diamond Mine site and surrounding area to provide baseline information on topographic conditions prior to dewatering of the regional aquifer system surrounding the mine pit. Several published articles have subsequently made use of this data set for peatland ecosystem classification (Difebo et al. 2015; Franklin and Ahmed 2017) and hydrologic analysis (Richardson et al. 2012). Here, we present some additional analyses of these data to demonstrate vertical resolution and accuracy of LiDAR-derived ground surface elevations in this northern peatland complex, and to evaluate the extent to which spatial patterns of surface drainage and wetness can be modeled using a LiDAR DEM. The specific objectives of these analyses were (1) to quantify the vertical accuracy of LiDAR-derived surface elevations relative to field surveyed elevations derived with an RTK differential GPS system along a transect that is representative of different peatland types in this region; (2) to demonstrate capabilities for LiDAR-based analysis of peatland landforms using topographic profiles; and (3) to test topographic flow-routing and wetness

index algorithms for their ability to predict spatial variations in peatland surface moisture, including the effects of hydrologic preprocessing methods and DEM resolution.

15.4.2 Methods

15.4.2.1 LiDAR Acquisition

A discrete return airborne LiDAR survey was conducted in July 2007 by Terrapoint Canada Inc. over a 462 km^2 area encompassing the 37 km^2 watershed shown in Figure 15.2. The study area was reduced to this watershed area to reduce computational demands for all subsequent steps. Laser pulse returns were classified into bare earth and vegetation classes by Terrapoint and delivered as tiled, *xyzi* ASCII files. The nominal density of bare earth returns was 0.5 points per square meter of ground surface. DEMs were interpolated at three different resolutions (1, 2.5, and 5 m) from the classified bare earth returns, using an inverse distance weighted interpolated with a low weighting exponent (0.5) and a maximum of four neighboring points. These DEMs were used for the calculation of hydrologic derivatives as described in Section 15.4.2.5.

15.4.2.2 Topographic Surveying and Accuracy Assessment

Approximately 350 ground survey elevation points were collected along a 1.6 km research transect (Figure 15.2) using a Topcon HiPer GL RTK GPS system. The range pole used for the roving receiver was equipped with a 10 m diameter base plate to prevent the pointed tip from sinking into the peat and causing a downward bias in elevation measurements. The nominal point spacing of these measurements was 5 m, although measurements could not be collected in a few localized areas due to either the presence of standing water or forest canopy cover. An offset of +6 cm was observed for survey points relative to LiDAR elevations, including over barren limestone outcrop features. This offset was attributed to a possible shift in the local benchmark used for the base station, or in the setup of the base station receiver overtop of it. A correcting offset of −6 cm was therefore applied to the survey elevations.

For each survey point, the nearest bare earth LiDAR return was identified within a maximum 2.5 m radius and used for accuracy assessment. Any points without a LiDAR bare earth return within a maximum ground distance of 2.5 m were discarded from subsequent analyses. For the remaining 223 observations, RMSE was calculated by comparing survey point elevations to the nearest LiDAR ground return elevation. The standard deviation of the difference in observed versus LiDAR-derived elevations was also calculated. Finally, we repeated this analysis using smaller values for the maximum ground distance (1.0, 0.5, 0.25, and 0.1 m). This had the effect of reducing the number of paired observations available for the analysis, but decreased the ground distance between observation pairs. The purpose of this was to first produce a conservative accuracy assessment using a wider search radius and larger sample size, followed by more optimistic estimates based on smaller search radii and sample sizes. The assumption was that smaller search radii would produce lower RMSE and

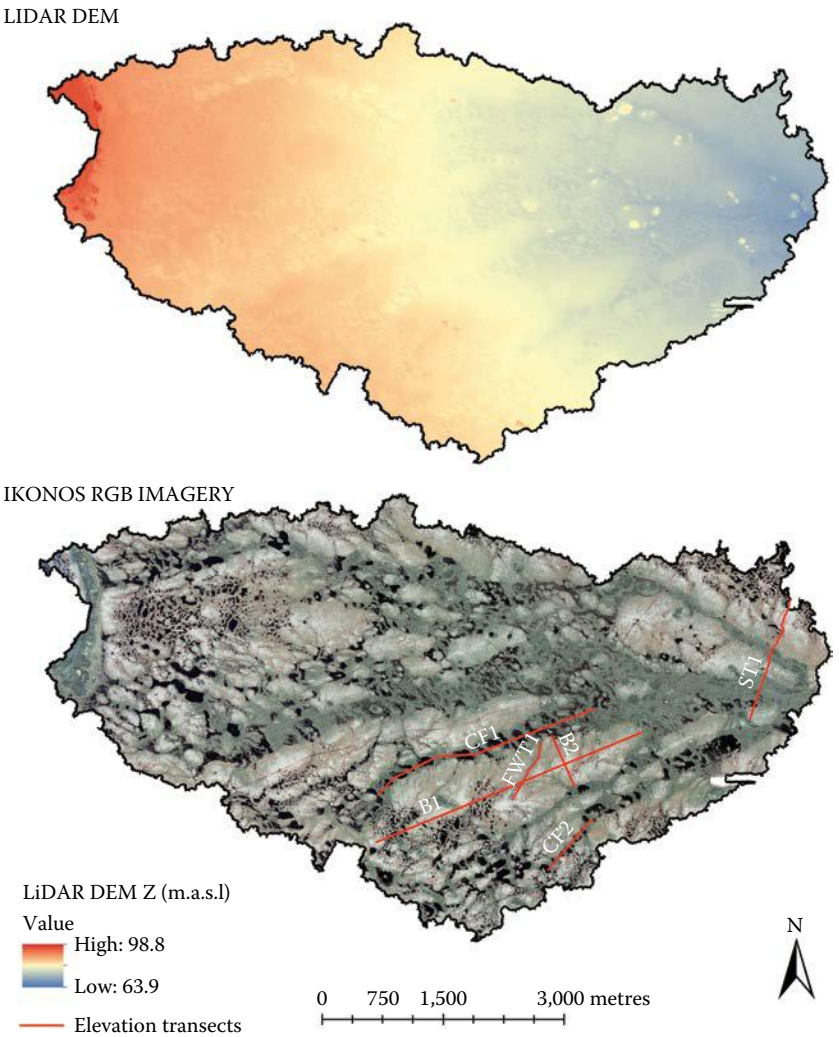


FIGURE 15.2 One-meter resolution LIDAR-derived DEM (top) and satellite-derived IKONOS (bottom) imagery of the study site, showing the location of the topographic survey transect (ST1) and five additional elevation profile transects (B1 = longitudinal cross section of a bog, B2 = transverse cross section of a bog, FWT1 = fen water track draining from bog, CF1 and CF2 = longitudinal cross sections of channel fens).

standard deviations compared to larger search radii and might be more representative of the accuracy of individual ground return elevations. The need for this analysis reflects the difficulty of precisely colocating manual survey points with individual LiDAR ground returns with inherently low horizontal positional accuracy (on the order of $\pm$ 60 cm).

15.4.2.3 Topographic Profiles

ArcGIS Desktop 10.5.1 (ESRI, 2016) was used to digitize six topographic profiles, including one along the 1.6 km survey transect (ST1) described previously. The locations of these profiles are indicated in Figure 15.2, and were chosen to represent different peatland features, including longitudinal and transverse profiles of a domed bog (B1 at 3000 m and B2 at 800 m), and downslope profiles of two channel fens (CF1 at 3000 m and CF2 at 800 m) and a fen water track (FWT1 at 800 m) or ladder fen draining radially off a domed bog toward a channel fen. The average topographic gradient was calculated from the slope term of a fitted linear regression model of elevation versus distance for each transect. For the transverse bog dome profile, B2, average gradient was calculated separately for each side of the dome.

15.4.2.4 Soil Moisture Surveys

On three dates during the summer of 2012 (July 1, August 1, and August 28), near-surface volumetric soil moisture (VSM) measurements were collected along the survey transect (ST1) using a HydroSense handheld time domain reflectometer (TDR) equipped with 12 cm length probes (Figure 15.3). In total, 14 clusters of soil moisture observations were visited on each date. The cluster locations were chosen at equally spaced distances along ST1 and shifted at least 30 m to the side to avoid having the influence of the transect in subsequent image analysis steps. At each cluster a center flag was set up, followed by at least four additional flags at 8 m distances from the center in each of the four cardinal directions. At each flag, three TDR measurements were collected in a hummock feature, and three were collected in a hollow feature, provided that both features were present. The average of three measurements was calculated for each hummock and hollow at each flag. The TDR probes were inserted at a 45° angle for an approximately 7.5 cm depth integrated measurement. At some cluster sites, additional flags were inserted for TDR readings to achieve representative sampling of hummocks and hollows for that site (Figure 15.3).

15.4.2.5 Digital Terrain Analysis and Soil Moisture Models

All digital terrain analyses were conducted using the System for Automated Geoscientific Analysis (SAGA) GIS program (Conrad et al. 2015). TWI was computed using three different algorithms at three different DEM resolutions (1, 2.5, and 5 m) following hydrologic preprocessing with two different methods, for a total of 18 different realizations. The TWI methods differed primarily in the way in which water is distributed downslope through the DEM grid cells. We tested (1) TWI based on deterministic-8 (D8) flow routing (TWI-D8); (2) TWI based on a maximum downslope gradient (TWI-MDG) as described by Qin et al. (2011), which distributes flow from one pixel to multiple neighboring pixels as a function of the maximum local slope; and (3) the SAGA wetness index (SWI), which allows for more significant redistribution of water over flat areas of the landscape, as described by Böhner and Selige (2006). For hydrologic preprocessing prior to calculation of the TWI surfaces, we used the Sink Removal module after the Sink Routes module with (1) the Fill Sinks option, and (2) Deepen Drainage Routes option. Both methods cause

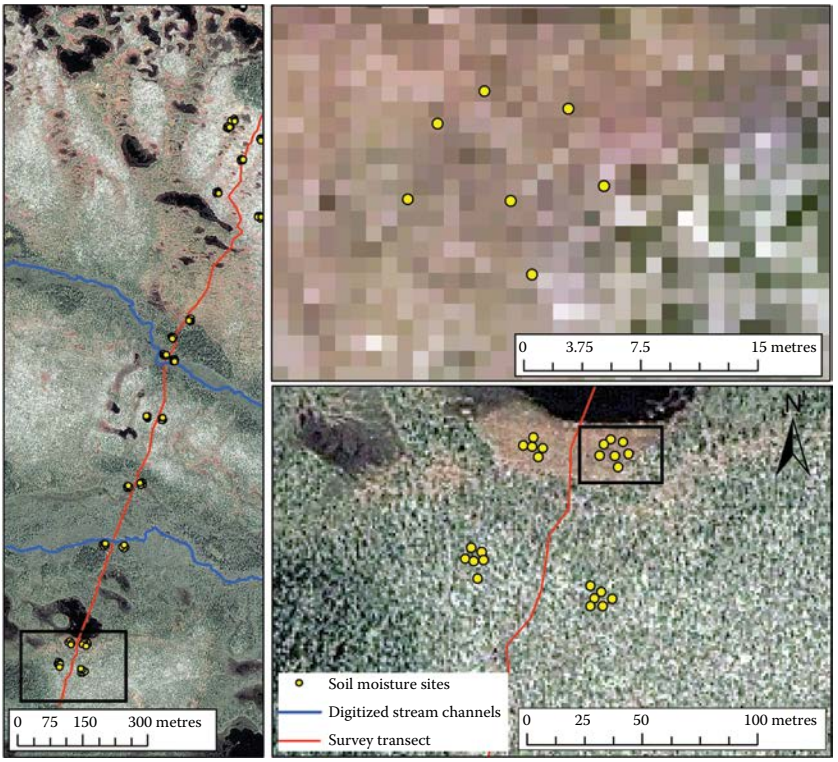


FIGURE 15.3 Map of survey transect (left) showing soil moisture survey sites with inset maps (bottom and top right) for detail.

artificial modifications to the DEM, which can strongly impact subsequent analyses and observed spatial patterns (Lindsay and Creed 2005). The fill method raises a sink pixel elevation to the minimum of its surrounding eight neighbors, whereas the deepen method, or breaching, lowers the elevation of one of the surrounding eight pixels to the elevation of the sink pixel.

Regression analysis was used to model the relation between the various TWI realizations and near-surface soil moisture at two different scales. The point-scale models were based on average VSM measurements taken at each flag site ($n = 120$), and the aggregated models were based on an average of all flag sites at each cluster ($n = 20$). Separate models were produced for different wetland microforms: (1) hummocks only, (2) hollows only, and (3) hummocks *and* hollows. VSM observations were averaged across all three sampling dates to simplify the total number of models required.

Finally, we tested for the relation between VSM and a derivative representing local topographic position derivative, the DME, computed at multiple scales from a nonhydrologically preprocessed, 1 m resolution DEM. DME is calculated from a moving window of specified radius and calculates a value for the central, target pixel, which is the elevation difference of that pixel relative to the mean elevation

of all pixels in its neighborhood. Thus, DME represents local topographic position within an area defined by the moving-window radius and can represent relief at multiple scales. One advantage of this derivative is that it does not require hydrologic preprocessing or additional algorithmic choices related to flow routing, as is the case with TWI. Moreover, it can preserve high-resolution elevation information while increasing the neighborhood size in order to test for scale effects. Difebo et al. (2015) used DME to improve an ecosystem classification map for the same study region and found the largest window sizes to be most effective (up to 150 m radius). For the present study, DME was calculated with sequentially larger neighborhood radii: 3, 7, 15, 30, and 70 m. As with the TWI models, regression models were fit for different microform types, but only at the point scale.

15.4.3 Results and Discussion

15.4.3.1 LiDAR Ground Return Elevation Accuracy Assessment

LiDAR-derived surface elevations along the 1.6 km transect S1 cross two stream channel basins, with a drainage divide located at approximately 700 m (Figure 15.4). These features are clearly captured within the data, despite the very low range in elevation (2.0 m total relief). The LiDAR elevation transect shows high local variability, on the order of approximately 50 cm over distances of ~5–50 m. A similar scale of local topographic variability is also observed in the manually surveyed elevations, although there is not always a clear correspondence between LiDAR-derived versus surveyed elevations over short distances. Nevertheless, the surveyed elevations clearly reflect the mesoscale topographic relief along this transect, including two stream

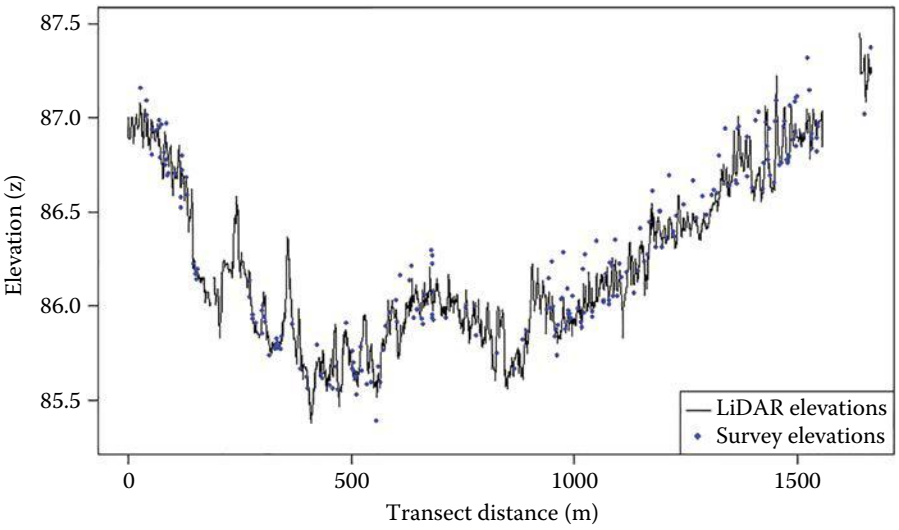


FIGURE 15.4 LiDAR-derived elevations interpolated at 1 m spacing along the 1.6 km survey transect, and 224 manual survey elevations used for the point-to-point accuracy assessment.

valley bottoms and a corresponding drainage divide, as captured by the higher density LiDAR surface elevations.

The overall RMSE for all 224 surveyed points was found to be 0.10 m. This is among the lowest elevation error magnitudes reported in Table 15.2 for point-to-point comparisons and may reflect the relatively low density of vegetation conditions at the study site compared to many of the cited studies. Our error assessment was based on a point-to-point comparison, whereby the LiDAR ground return that was closest to a survey elevation point was used paired with that surveyed point for the accuracy assessment. A problem with this approach is the inherent ambiguity in the actual ground surface distance between a surveyed point with high positional accuracy and a single LiDAR ground return with lower positional accuracy. Moreover, as the ground return point density decreases (e.g., in areas of higher vegetation density), the likelihood of a ground return point being found within close proximity (e.g., less than 1 m) to a survey point decreases. To help address this issue, we repeated the accuracy assessment several times while sequentially reducing the maximum ground distance separation between a surveyed elevation point and associated ground return. Although this has the effect of reducing the number of observations for the accuracy assessment, it also reduces the observed error, as seen in Figure 15.5. Notably, with a maximum separation distance of 2.5 cm, only four observations were available for comparison, but the RMSE was reduced from 10.0 to 4.6 cm. This analysis demonstrates that individual ground return accuracy may in fact be better than what is typically reported due to the challenge of colocating a manual survey point with a LiDAR return. As the ground return point density decreases (e.g., in areas of higher vegetation density), this problem is exacerbated because the

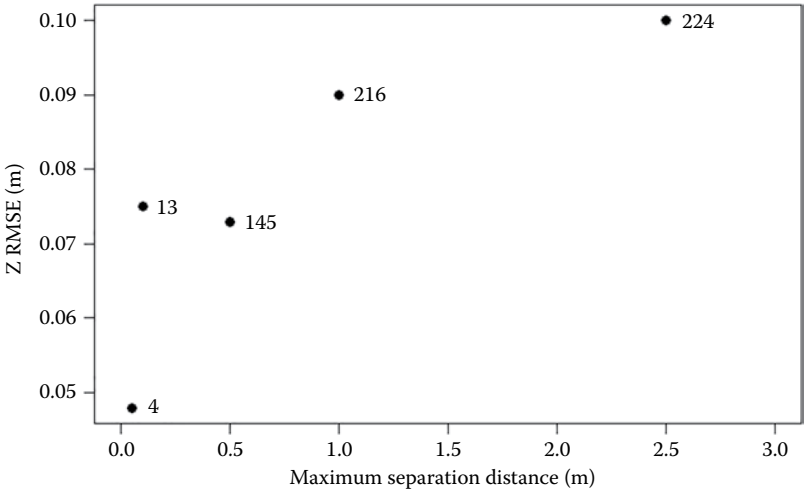


FIGURE 15.5 Observed point-to-point RMSEs calculated from surveyed elevations as a function of increasing maximum ground separation distance. The numbers beside each point indicate the sample size available for the error calculation, which decreases as the separation distance decreases.

likelihood of colocating survey points with ground returns within a minimum ground separation distance decreases. Thus, while various researchers have found vegetation to introduce positive biases in ground surface elevations, our analysis also suggests that an increased random error component may also be introduced due to lower ground return densities, hence the larger ground separation distance between manual versus LiDAR ground returns. Figure 15.4 demonstrates that, despite the ambiguity of individual ground return accuracy of short length scales and lack of large-scale relief, the mesoscale relief on the order of 1–2 m, including drainage structures and watershed divides, is still captured within the LiDAR DEM.

15.4.3.2 Peatland Morphology Examples

Although typically flat and devoid of large-scale topographical relief, peatlands exhibit characteristic geomorphologies that can be quantified with high-resolution topographic data (Richardson et al. 2010; Difebo et al. 2015). Figure 15.6 shows five examples of topographic profiles derived from a 1 m LiDAR-derived bare earth DEM for the study site, including bog and fen features, which together occupy more than 90% of this landscape. Transects B1 and CF1 are equal-length, longitudinal transects for a bog and channel fen, respectively. While they exhibit similar average gradients, the profiles demonstrate differences in surface morphology, with a clear pattern of

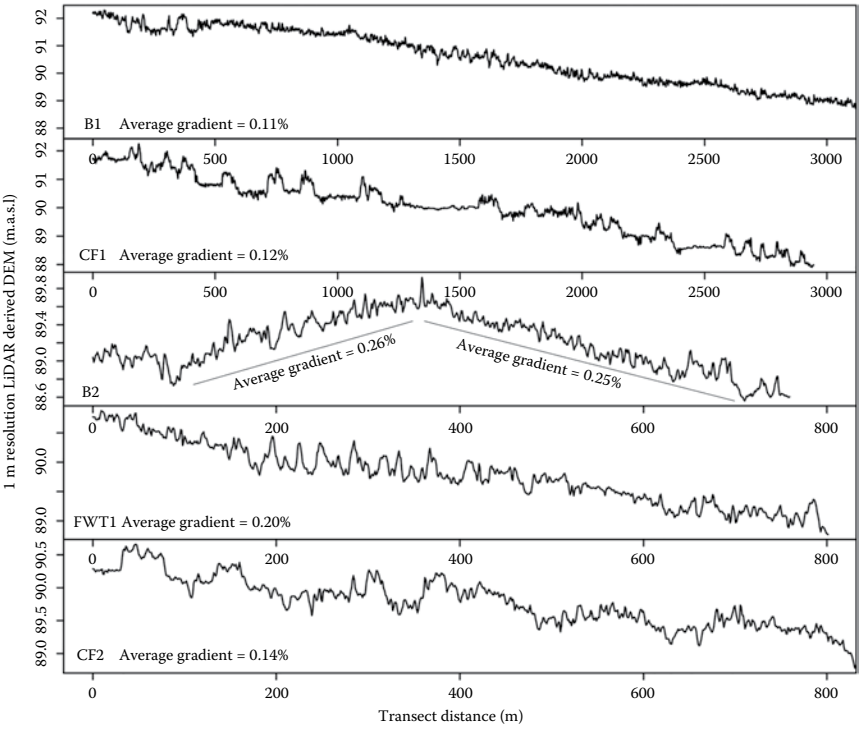


FIGURE 15.6 Elevation profiles for transects shown in Figure 15.2. Note different transect lengths for the top two transects (3000 m) compared to the bottom three transects (800 m).

ridges and pools for the channel fen transect (CF1) giving rise to a lower frequency periodicity of surface undulations compared to the bog transect (B1). The shorter channel fen transect (CF2) also exhibits a similar ridge-pool periodicity. The ridge-pool succession observed for the fen water track (FWT1) transect shows a higher periodicity indicative of smaller ridge-pool features drainage off the larger bog dome structure. CF2 and FWT1 transects both exhibit a steeper average gradient (0.14% and 0.2%, respectively) than the larger channel fen transect, CF1, and longitudinal bog transect (B1). The convex, raised dome structure seen for transcet B2 shows a dome height of approximately 0.8 m above the confining stream channels, as well as a slight asymmetry, most likely due to differences in base heights of the stream channels on either end of the transect. The rand slope (i.e., the sloping margin of a raised bog) exhibits higher average gradients (0.25%–0.26%) compared to any of the other transects shown.

In summary, despite some ambiguities in the accuracy of individual LiDAR ground return elevations as discussed in Section 15.4.3.1, Figures 15.5 and 15.6 demonstrate that LiDAR-derived surface topography can be used to quantitatively describe subtle, mesoscale geomorphic characteristics of some representative bog and fen forms in the James/Hudson Bay Lowlands. Given the synoptic nature of LiDAR surveys, this presents new possibilities for advancing current understanding of how peatland ecosystems evolve over timescales of centuries to millennia. Moreover, efforts to mathematically model peatlands as complex adaptive systems would benefit from LiDAR-derived surface topography for calibration and validation (Table 15.3).

15.4.3.3 Digital Terrain Analyses and Soil Moisture Model Results

Given the large number of algorithms available for computing TWI, as well as DEM resolution and hydrologic preprocessing options, it is unclear from previous research what options would be most suitable for application in low-gradient, wetland-dominated environments such as the James/Hudson Bay Lowlands. The results presented in this section demonstrate that such choices have a strong impact on simulated patterns of wetness at our study site. An example of the six different methods for calculating TWI is shown in Figure 15.7 at 2.5 m resolution. Qualitatively, the hydrologic preprocessing method (breach versus fill) has the largest impact on the spatial pattern of TWI. Realizations based on the fill method exhibit a more diffuse pattern of wetness compared to breach-based realizations, which exhibit more concentrated patterns of moisture distribution. Moreover, fill preprocessing appears to result in stronger differences among the three different TWI algorithms. Conversely, there appears to be little difference among TWI-D8 and TWI-MDG for the scenarios based on breach preprocessing. The SWI algorithm produces more distributed spatial patterns of wetness compared to the TWI-MDG algorithm, particularly with fill preprocessing.

As seen in Table 15.4, regression model results for VSM vary among the different TWI algorithms and preprocessing methods, corroborating the qualitative differences seen in Figure 15.7. DEM resolution also strongly influences the explanatory power of TWI. Overall, the point-scale models demonstrate lower predictive power of TWI compared to the aggregated models. For simplicity, therefore, the subsequent discussion focuses on the aggregated model scenarios.

TABLE 15.3
Some Studies Reporting Point-to-Point LiDAR Ground Elevation Accuracy Assessments in Wetland Environments

	Vertical Error		
Wetland or Vegetation Type	Bias (m)	Standard Deviation (and/or RMSE if in Brackets) (m)	Authors
Grass/herbs	0.02	0.10	Hopkinson et al. (2005)
Low shrubs	0.06	0.12	Hopkinson et al. (2005)
Tall shrubs	0.06	0.03	Hopkinson et al. (2005)
Aquatic vegetation	0.15	0.22	Hopkinson et al. (2005)
Deltaic wetland—Bare	0.00*	(0.07)	Töyrä et al. (2003)
Deltaic wetland—Graminoid	0.07*	(0.15)	Töyrä et al. (2003)
Deltaic wetland—Willow	0.15*	(0.26)	Töyrä et al. (2003)
Deltaic wetland—Dead willow	0.14*	(0.17)	Töyrä et al. (2003)
Salt marsh	0.1	0.12	Hladik and Alber (2012)
Salt marsh	−0.01**	0.09	Hladik and Alber (2012)
Forested wetland	0.00*	(0.10)	Richardson et al. (2010)
Salt marsh—all	0.15	0.18 (23.3)	Schmid et al. (2011)
Salt marsh—*Spartina alterniflora*	0.11	0.11 (0.16)	Schmid et al. (2011)
Salt marsh—*Juncus roemerianus*	0.30	0.22 (0.37)	Schmid et al. (2011)
Salt marsh—*Borrichia frutescens*	0.11	0.10 (0.15)	Schmid et al. (2011)
Salt marsh—*Salicornia virginica*	0.02	0.12 (0.12)	Schmid et al. (2011)
Salt marsh	0.13	0.12 (0.07)	Morris et al. (2005)
Northern peatland	0.00*	0.10 (0.10)	This study

* Bias after offset correction; ** Bias after block adjustment correction.

The aggregated models were based on a sample size of 20, with each observation representing the average VSM for four to six hummock and hollow measurements within a cluster with an average radius of approximately 16 m. Models based on hummocks *only* were mostly insignificant, whereas soil moisture in hollow features were almost always statistically significant, with many models explaining more than 50% of the spatial variation in VSM. Compared to hollows, hummocks tend to have very low bulk density due to the presence of live mosses and other vegetation causing very low VSM readings and consequently low overall variability. Moreover, hummocks represent fine-scale topographic highs, often over very short length scales, and may not be adequately captured given the point density of the LiDAR ground returns in this study ($< 1\ m^{-2}$).

For hollows *only*, the best overall model was achieved using the SWI algorithm at 2.5 m resolution. However, SWI performed poorly at 1, 5, and 10 m resolutions. Conversely, TWI-MDG resulted in relatively strong models under both preprocessing scenarios and at both 2.5 and 5 m resolutions, but not 1 and 10 m. Models for hummocks *and* hollows largely followed the same patterns as for hummocks only.

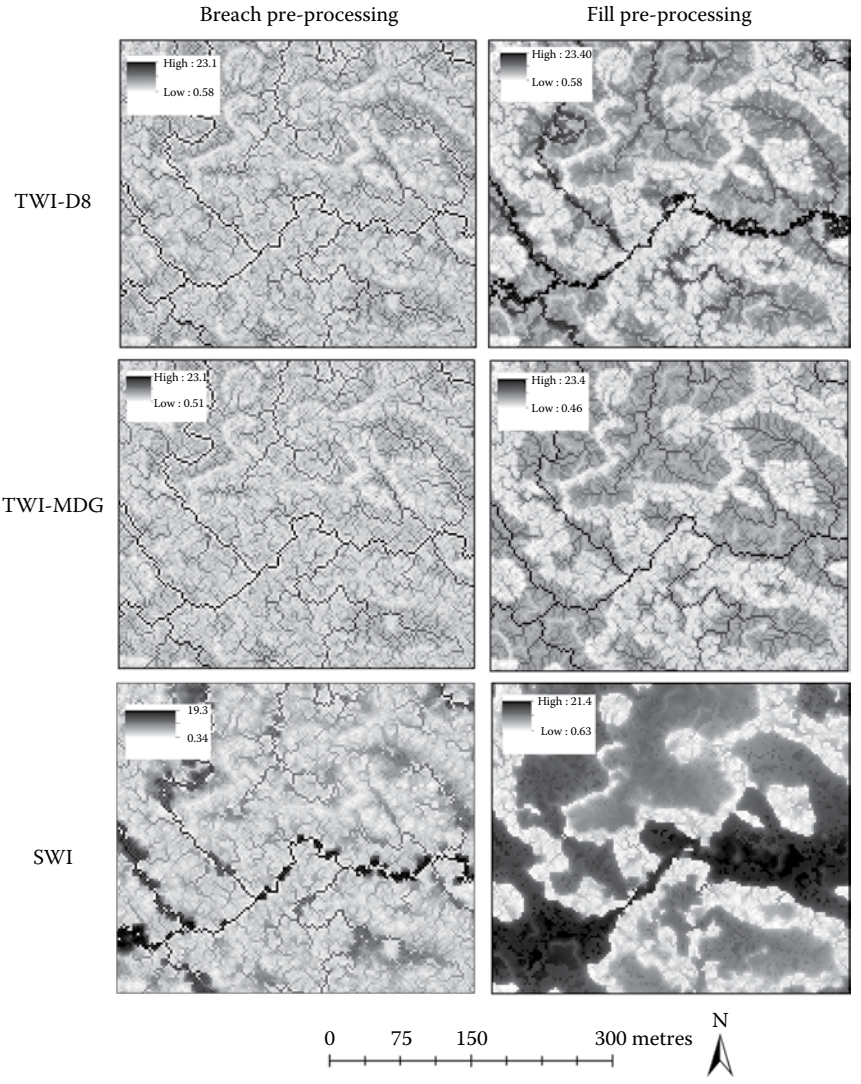


FIGURE 15.7 Examples of TWI surfaces calculated using three different algorithms (top, middle, and bottom panels) after breach (left) versus fill (right) hydrologic preprocessing, as described in Section 15.4.2.5.

This likely reflects the contribution of hollow features to the overall variability in terrain wetness and associated relations with TWI (Figure 15.8).

For point models ($n = 120$), virtually none were significant at 1 m resolution for breach preprocessing scenarios or for hummock models. At 2.5, 5, and 10 m resolutions, most models were weakly significant ($R^2 \leq 0.2$). For fill preprocessing scenarios, the best point models were achieved for TWI realization based on the more distributed flow routing algorithms (TWI-MDG and SWI). TWI-MDG was

TABLE 15.4
Summary of Regression Model Results (R^2 and Significance Level)

			Breach Preprocessing				Flood Preprocessing			
			DEM Resolution				DEM Resolution			
Scenarios			**1 m**	**2.5 m**	**5 m**	**10 m**	**1 m**	**2.5 m**	**5 m**	**10 m**
Hummocks and hollows	Points	TWI-D8	n.s.	0.12***	0.10***	0.14***	0.14*	0.30***	0.24***	0.15***
		TWI-MDG	n.s.	0.19***	0.14***	0.13***	0.31***	0.37***	0.35***	0.27***
		SWI	0.06**	0.32***	0.10***	n.s.	0.38***	0.33***	0.16***	n.s.
	Aggregated	TWI-D8	n.s.	0.39**	0.42***	0.21	0.13*	**0.64*****	**0.56*****	0.19*
		TWI-MDG	n.s.	**0.56*****	**0.58*****	0.18*	0.35***	**0.51*****	**0.66*****	0.36**
		SWI	0.23*	**0.65*****	0.22*	n.s.	0.51***	0.46***	0.19*	n.s.
Hollows only	Points	TWI-D8	n.s.	0.10***	0.12***	0.12***	0.13***	0.29***	0.26***	0.12***
		TWI-MDG	n.s.	0.20***	0.16***	0.12***	0.33***	0.39***	0.41***	0.23***
		SWI	0.05**	0.36***	0.09***	n.s.	0.42***	0.38***	0.10***	0.03*
	Aggregated	TWI-D8	n.s.	0.41**	0.41**	0.22*	0.38**	**0.68*****	**0.54*****	0.21*
		TWI-MDG	n.s.	**0.63*****	**0.57*****	0.21*	0.37***	**0.54*****	**0.68*****	0.35***
		SWI	0.26*	**0.72*****	0.16*	n.s.	**0.57*****	0.48***	0 n.s.	n.s.
Hummocks only	Points	TWI-D8	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
		TWI-MDG	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
		SWI	n.s.	n.s.	0.02*	n.s.	n.s.	n.s.	0.10***	n.s.
	Aggregated	TWI-D8	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
		TWI-MDG	n.s.	n.s.	n.s.	n.s	n.s.	n.s.	n.s.	n.s.
		SWI	n.s.	n.s.	0.13*	n.s.	n.s.	n.s.	0	n.s.

Note: n.s. = not statistically significant. All slopes were positive and are not indicated for brevity. Models with $R^2 > 0.5$ are indicated in bold. Note that, for multiple hypothesis testing, a Bonferroni correction should be applied. With 144 test scenarios, only tests with a significance level of $p < 0.0001$ should be considered significant after Bonferonni correction at the 90% significance level.

*$p < 0.05$; ** $p < 0.001$; *** $p < 0.0001$.

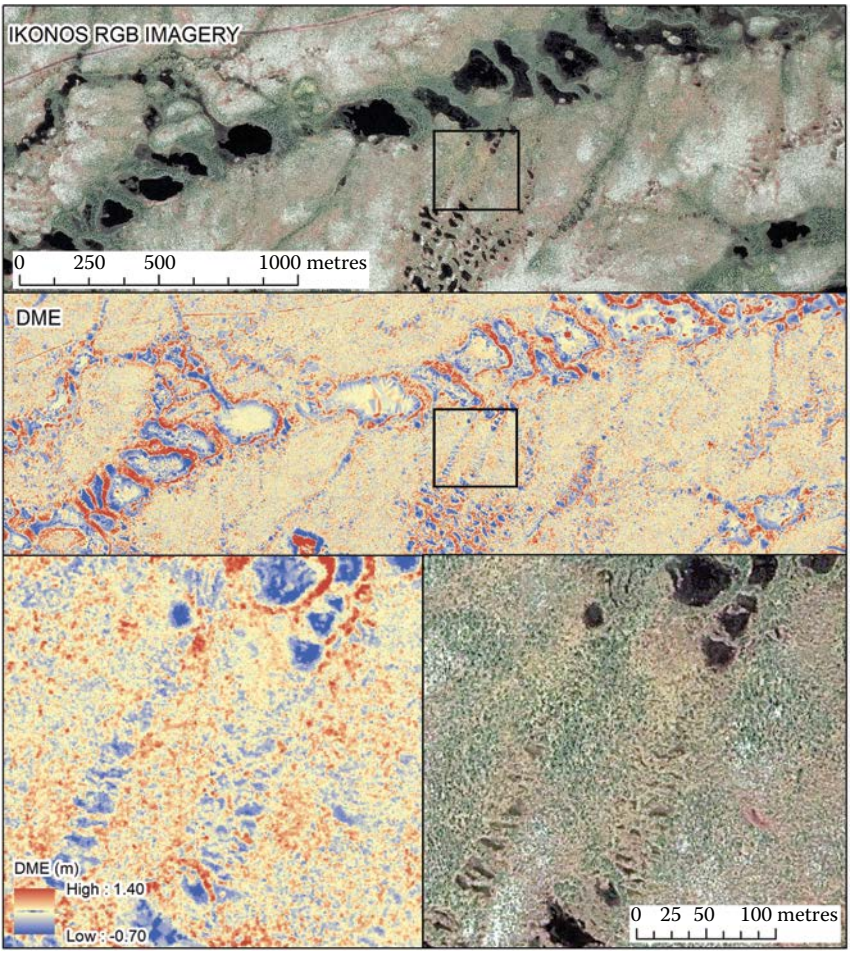


FIGURE 15.8 Comparison of IKONOS imagery versus DME computed from 1 m resolution LiDAR DEM using a 30 m radius moving-window neighborhood. DME captures local topographic forms that correspond well to different types of peatland formations, including two ladder fens shown in bottom panels.

a more robust predictor of VSM for 1–5 m resolutions, compared to SWI, although SWI models generally performed the best of all point models, and at a resolution of 1 m ($R^2 = 0.42$ for hummock models, $R^2 = 0.38$ for hummock and hollow models). At 10 m resolution, model goodness of fit generally decreased compared to the higher resolution scenarios. Otherwise, there was no consistent effect of resolution on the goodness of fit for models based on the more distributed flow routing algorithm.

Overall, these findings demonstrate strong scale dependence of TWI-based VSM models in this peatland landscape, as indicated by the effect of DEM resolution and aggregation of point scale models. Our results also demonstrate an important

effect of flow routing and hydrologic preprocessing algorithms on digital terrain analysis outputs and associated models. Although VSM models varied widely across the different model scenarios, it was shown that distributed flow routing algorithms tended to work more effectively than the commonly used D8 algorithm. This corroborates the findings of Lang et al. (2013), who demonstrated that TWI indices based on more distributed flow routing algorithms were more suitable for wetland mapping in a humid, temperate region watershed in the U.S. Coastal Plain region. We also found that moderate DEM resolution models (2.5 and 5 m resolution) generally performed better than higher (1 m) and lower (10 m) resolution for the aggregated models. For point-scale models, which generally underperformed compared to aggregated models, the best models were achieved at higher (1–5 m) resolutions. In all cases, soil moisture in hummocks could not be well predicted by TWI at any scales. Thus, plot-scale soil moisture is driven largely by wetness conditions of hummocks in this landscape.

Finally, for point-scale models, the DME derivative was found to outperform all TWI models for the prediction of VSM (Table 15.5). The maximum coefficient of determination for hummock and hollow models combined was 0.47, compared to 0.39 for the highest performing TWI model. The optimal scale for the DME calculation was determined to be a radius of either 30 or 70 m, with performance dropping substantially for both smaller and larger window sizes. Since the scale of analysis is defined by the radius of the moving window, the DEM resolution was not varied for these scenarios. An advantage of this derivative is the ability to test the effect of different analysis scales while preserving a high DEM resolution. Moreover, DME does not require hydrologic preprocessing or testing of different types of algorithms, as in the case of TWI. DME outperformed TWI, which implies that the upslope contributing area does not help in the prediction of surface soil

TABLE 15.5
Point-Scale VSM Models Based on DME Derivative Computed at Multiple Resolutions from an Unprocessed 1 m Resolution Digital Elevation Model

	Analysis Neighborhood (Radius)					
Scenario	**3 m**	**7 m**	**15 m**	**30 m**	**70 m**	**150 m**
Hummocks + hollows	n.s.	0.10**	0.31***	0.47***	0.46***	0.21**
Hollows only	n.s.	0.12***	0.38***	0.52***	0.47***	0.25**
Hummocks only	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.

Note: n.s. = not statistically significant. All slopes were negative and are not indicated for brevity. After Bonferroni correction for 18 tests, only models with $p < 0.01$ should be considered significant at the 90% significance level. For comparison with TWI, these results should be compared to the point-based scenarios reported in Table 15.4.

* $p < 0.05$; ** $p < 0.001$; *** $p < 0.0001$.

moisture in this landscape. Rather, local topographic variability within a radius defined by 30–70 m appears to most strongly correlate with soil moisture in this peatland complex.

In summary, the results of this study allow us to recommend that, for the purpose of surface moisture mapping with the TWI in a northern peatland complex, researchers should use higher resolution DEMs (1–5 m) that have been preprocessed using a sink-fill algorithm. Distributed flow routing methods should be preferred over deterministic methods, and at higher spatial resolutions. However, due to the large number of algorithms available for calculating TWI, results may vary widely depending on the specific choices made. Instead, we recommend the simpler metric, DME, which requires fewer algorithmic decisions and which outperformed all realizations of the TWI in the prediction of surface soil moisture patterns.

15.5 CONCLUSION

Remote sensing with airborne LiDAR has become an important and active area of wetland science research in the past two decades. We have presented a review of this expanding field with examples from a wide range of wetland ecosystem types, including forested wetlands, northern peatlands, freshwater marshes, salt marshes, mangroves, and floodplain environments. LiDAR remote sensing allows for three-dimensional characterization of ground and vegetation features, which provide valuable information on wetland form and function. This information can be used for detection and classification of wetland ecosystems, hydrologic and geomorphic analysis, and wetland restoration.

Issues and challenges arise in the application of digital vegetation and terrain analysis in wetland environments with airborne LiDAR, requiring careful consideration on behalf of the researcher. Wetlands are frequently characterized by short and dense vegetation, which can be challenging to detect and analyze compared to taller vegetation typical of forested environments. Low-lying, dense vegetation also introduces error in ground return elevations to reduced laser penetration and related challenges associated with classifying laser pulse returns into ground versus aboveground features. These errors may be exacerbated by local-scale topographic variability and may be high relative to the surrounding topography, given that wetland environments lack large-scale relief. Nevertheless, mesoscale topographic gradients can be well characterized in wetland environments, even in areas with slope gradients on the order of only 0.1%. Moreover, biophysical characteristics of wetland vegetation inferred from aboveground returns can be linked to wetland ecosystem classes, flood periodicity, and water sources.

Through a case study of a northern peatland complex in the James/Hudson Bay Lowlands, we demonstrated concepts related to the accuracy of LiDAR-derived ground surface elevations, as well as geomorphic and hydrologic analysis. Ground return elevations had an average RMSE of 0.10 m based on a point-to-point comparison, although we demonstrated this estimate of error to be conservative due to ambiguities associated with horizontal positioning of LiDAR ground returns. As a result, *relative* vertical accuracy may be considerably better than 0.10 m. Even under this conservative estimate of ground elevation errors, however, LiDAR-derived

topography was found to accurately capture mesoscale topography, and can be used to analyze hydrologic gradients as well as different types of peatland landforms such as raised bogs, channel fens, and ladder fens.

The well-known and commonly used TWI was found to correlate moderately well with local surface soil moisture patterns, although results varied widely as a function of DEM resolution, hydrologic preprocessing technique, and flow routing algorithms. VSM of hummocks (locally raised microforms) was not adequately modeled by TWI, whereas VSM of local depressions (hollows) could be moderately well explained by the index. We found the results of this analysis to be scale sensitive, with improvements being realized through aggregation of field observations up to the plot scale. Given some of these challenges associated with the TWI, we tested a simple, scalable derivative, DME, which describes local topography. At the point scale, DME computed at 1 m resolution, using a local neighborhood defined by a radius of either 30 or 70 m, outperformed all TWI realizations. The DME derivative requires no hydrologic preprocessing or other algorithmic choices and can be easily adapted to different scales of relief while preserving the original input resolution. These case study results corroborate other studies in demonstrating the outstanding capabilities of airborne LiDAR surveying to provide highly detailed, synoptic information on wetland biophysical and hydrologic information over large, remote areas.

REFERENCES

Anderson, K, JJ Bennie, EJ Milton, PDM Hughes, R Lindsay, and R Meade. 2010. "Combining LiDAR and IKONOS data for eco-hydrological classification of an ombrotrophic peatland." *Journal of Environmental Quality* 39 (1):260–273.

Baker, ME, MJ Wiley, PW Seelbach, and ML Carlson. 2003. "A GIS model of subsurface water potential for aquatic resource inventory, assessment, and environmental management." *Environmental Management* 32 (6):706–719.

Ballhorn, U, J Jubanski, K Kronseder, and F Siegert. 2012. "Airborne LiDAR measurements to estimate tropical peat swamp forest above ground biomass." *2012 IEEE International Geoscience and Remote Sensing Symposium (IGARSS).*

Belyea, LR and AJ Baird. 2006. "Beyond 'the limits to peat bog growth': Cross-scale feedback in peatland development." *Ecological Monographs* 76 (3):299–322.

Beven, KJ and MJ Kirkby. 1979. "A physically based, variable contributing area model of basin hydrology/Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant." *Hydrological Sciences Journal* 24 (1):43–69.

Boehm, H-DV, V Liesenberg, and SH Limin. 2013. "Multi-temporal airborne LiDAR-survey and field measurements of tropical peat swamp forest to monitor changes." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 6 (3):1524–1530.

Böhner, J and T Selige. 2006. "Spatial prediction of soil attributes using terrain analysis and climate regionalisation." *Gottinger Geographische Abhandlungen* 115:13–28.

Bradbury, RB, RA Hill, DC Mason, SA Hinsley, JD Wilson, H Balzter, GQA Anderson et al. 2005. "Modelling relationships between birds and vegetation structure using airborne LiDAR data: a review with case studies from agricultural and woodland environments." *Ibis* 147 (3):443–452.

Brennan, R and TL Webster. 2006. "Object-oriented land cover classification of lidar-derived surfaces." *Canadian Journal of Remote Sensing* 32 (2):162–172.

Bridgham, SD, J Pastor, JA Janssens, C Chapin, and TJ Malterer. 1996. "Multiple limiting gradients in peatlands: A call for a new paradigm." *Wetlands* 16 (1):45–65.

Chadwick, J. 2011. "Integrated LiDAR and IKONOS multispectral imagery for mapping mangrove distribution and physical properties." *International Journal of Remote Sensing* 32 (21):6765–6781.

Chust, G, I Galparsoro, A Borja, J Franco, and A Uriarte. 2008. "Coastal and estuarine habitat mapping, using LIDAR height and intensity and multi-spectral imagery." *Estuarine, Coastal and Shelf Science* 78 (4):633–643.

Conrad, O, B Bechtel, M Bock, H Dietrich, E Fischer, L Gerlitz, J. Wehberg, V Wichmann, and J Böhner. 2015. "System for Automated Geoscientific Analyses (SAGA) v. 2.1.4." *Geoscience Model Development* 8, 1991–2007. DOI:10.5194/gmd-8-1991-2015.

Creed, IF, FD Beall, TA Clair, PJ Dillon, and RH Hesslein. 2008. "Predicting export of dissolved organic carbon from forested catchments in glaciated landscapes with shallow soils." *Global Biogeochemical Cycles* 22 (4):3629–3648. DOI: 10.1002/hyp.1357.

Creed, IF SE Sanford, FD Beall, LA Molot, and PJ Dillon. 2003. "Cryptic wetlands: integrating hidden wetlands in regression models of the export of dissolved organic carbon from forested landscapes." *Hydrological Processes* 17 (18):3629–3648.

Creed, IF, and GZ Sass. 2011. "Digital terrain analysis approaches for tracking hydrological and biogeochemical pathways and processes in forested landscapes." In *Forest Hydrology and Biogeochemistry*, 69–100. Springer.

Damman, AWH. 1986. "Hydrology, development, and biogeochemistry of ombrogenous peat bogs with special reference to nutrient relocation in a western Newfoundland bog." *Canadian Journal of Botany* 64 (2):384–394.

Difebo, A, M Richardson, and J Price. 2015. "Fusion of multispectral imagery and LiDAR digital terrain derivatives for ecosystem mapping and morphological characterization of a northern peatland complex." In *Remote Sensing of Wetlands: Applications and Advances*, 399–412. CRC Press.

Englhart, S, J Jubanski, and F Siegert. 2013. "Quantifying dynamics in tropical peat swamp forest biomass with multi-temporal LiDAR datasets." *Remote Sensing* 5 (5):2368–2388.

Evans, JS and AT Hudak. 2007. "A multiscale curvature algorithm for classifying discrete return LiDAR in forested environments." *IEEE Transactions on Geoscience and Remote Sensing* 45 (4):1029–1038.

Esri Inc. 2016. ArcMap 10.5.1 [Computer software]. Esri Inc., Redlands, CA.

Franklin, SE and OS Ahmed. 2017. "Object-based wetland characterization using Radarsat-2 quad-polarimetric SAR data, Landsat-8 OLI imagery, and airborne lidar-derived geomorphometric variables." *Photogrammetric Engineering and Remote Sensing* 83 (1):27–36.

Genc, L, SE Smith, and BA Dewitt. 2005. "Using satellite imagery and LIDAR data to corroborate an adjudicated ordinary high water line." *International Journal of Remote Sensing* 26 (17):3683–3693.

Gilmore, MS, EH Wilson, N Barrett, DL Civco, S Prisloe, JD Hurd, and C Chadwick. 2008. "Integrating multi-temporal spectral and structural information to map wetland vegetation in a lower Connecticut River tidal marsh." *Remote Sensing of Environment* 112 (11):4048–4060.

Glaser, PH, B Hansen, DI Siegel, AS Reeve, and PJ Morin. 2004. "Rates, pathways and drivers for peatland development in the Hudson Bay lowlands, northern Ontario, Canada." *Journal of Ecology* 92 (6):1036–1053.

Hjerdt, KN, JJ McDonnell, J Seibert, and A Rodhe. 2004. "A new topographic index to quantify downslope controls on local drainage." *Water Resources Research* 40 (5):W05602. DOI:10.1029/2004WR003130.

Hladik, C and M Alber. 2012. "Accuracy assessment and correction of a LIDAR-derived salt marsh digital elevation model." *Remote Sensing of Environment* 121:224–235.

Hogg, A and J Holland. 2008. "An evaluation of DEMs derived from LiDAR and photogrammetry for wetland mapping." *The Forestry Chronicle* 84 (6):840–849.

Hopkinson, C. 2007. "The influence of flying altitude, beam divergence, and pulse repetition frequency on laser pulse return intensity and canopy frequency distribution." *Canadian Journal of Remote Sensing* 33 (4):312–324.

Hopkinson, C, LE Chasmer, G Sass, IF Creed, M Sitar, W Kalbfleisch, and P Treitz. 2005. "Vegetation class dependent errors in lidar ground elevation and canopy height estimates in a boreal wetland environment." *Canadian Journal of Remote Sensing* 31 (2):191–206.

Huang, S, C Young, OI Abdul-Aziz, D Dahal, M Feng, and S Liu. 2013. "Simulating the water budget of a Prairie Potholes complex from LiDAR and hydrological models in North Dakota, USA." *Hydrological Sciences Journal* 58 (7):1434–1444.

Huang, S, C Young, M Feng, K Heidemann, M Cushing, DM Mushet, and S Liu. 2011. "Demonstration of a conceptual model for using LiDAR to improve the estimation of floodwater mitigation potential of Prairie Pothole Region wetlands." *Journal of Hydrology* 405 (3):417–426.

Jensen, JR, ME Hodgson, HE Mackey Jr, and W Krabill. 1987. "Correlation between aircraft MSS and LIDAR remotely sensed data on a forested wetland." *Geocarto International* 2 (4):39–54.

Korpela, I, M Koskinen, H Vasander, M Holopainen, and K Minkkinen. 2009. "Airborne small-footprint discrete-return LiDAR data in the assessment of boreal mire surface patterns, vegetation, and habitats." *Forest Ecology and Management* 258 (7):1549–1566.

Kronseder, K, U Ballhorn, V Böhm, and F Siegert. 2012. "Above ground biomass estimation across forest types at different degradation levels in Central Kalimantan using LiDAR data." *International Journal of Applied Earth Observation and Geoinformation* 18:37–48.

Lane, CR and E D'Amico. 2010. "Calculating the ecosystem service of water storage in isolated wetlands using LiDAR in North Central Florida, USA." *Wetlands* 30 (5):967–977.

Lang, M, G McCarty, R Oesterling, and I-Y Yeo. 2013. "Topographic metrics for improved mapping of forested wetlands." *Wetlands* 33 (1):141–155.

Lang, M, O McDonough, G McCarty, R Oesterling, and B Wilen. 2012. "Enhanced detection of wetland-stream connectivity using LiDAR." *Wetlands* 32 (3):461–473.

Langlois, MN, JS Price, and L Rochefort. 2015. "Landscape analysis of nutrient-enriched margins (lagg) in ombrotrophic peatlands." *Science of the Total Environment* 505:573–586.

Langlois, MN, M Richardson, and J Price. 2017. "Delineation of peatland lagg boundaries from airborne LiDAR." *Journal of Geophysical Research: Biogeosciences* 122 (9):2191–2205.

Lefsky, MA, WB Cohen, DJ Harding, GG Parker, SA Acker, and ST Gower. 2002. "Lidar remote sensing of above-ground biomass in three biomes." *Global Ecology and Biogeography* 11 (5):393–399.

Lim, K, P Treitz, M Wulder, B St-Onge, and M Flood. 2003. "Lidar remote sensing of forest canopy and stand structure." *Progress in Physical Geography* 27 (1):88–106.

Lindsay, JB and IF Creed. 2005. "Removal of artifact depressions from digital elevation models: Towards a minimum impact approach." *Hydrological Processes* 19 (16):3113–3126.

Lindsay, JB and IF Creed. 2006. "Distinguishing actual and artefact depressions in digital elevation data." *Computers and Geosciences* 32 (8):1192–1204.

Lindsay, JB, IF Creed, and FD Beall. 2004. "Drainage basin morphometrics for depressional landscapes." *Water Resources Research* 40 (9):W09307. DOI::10.1029/2004WR003322.

Luo, S, C Wang, F Pan, X Xi, G Li, S Nie, and S Xia. 2015. "Estimation of wetland vegetation height and leaf area index using airborne laser scanning data." *Ecological Indicators* 48:550–559.

MacMillan, RA, TC Martin, TJ Earle, and DH McNabb. 2003. "Automated analysis and classification of landforms using high-resolution digital elevation data: Applications and issues." *Canadian Journal of Remote Sensing* 29 (5):592–606.

Maxa, M and P Bolstad. 2009. "Mapping northern wetlands with high resolution satellite images and LiDAR." *Wetlands* 29 (1):248–260.

Millard, K, C Burke, D Stiff, and A Redden. 2009. "Detection of a low-relief 18th-century British siege trench using LiDAR vegetation penetration capabilities at Fort Beauséjour–Fort Cumberland National Historic Site, Canada." *Geoarchaeology* 24 (5):576–588.

Millard, K, C Hopkinson, A Redden, T Webster, and H Stewart, 2008. "Mapping vegetation friction indicators in a tidal salt marsh environment." In Hopkinson, C., Petroniro, A., Pomeroy, J., (Eds). *Hydroscan: Airborne Laser Mapping of Hydrological Features and Resources*, 167–190. Canadian Water Resources Agency, Saskatoon, Saskatchewan.

Millard, K, AM Redden, T Webster, and H Stewart. 2013. "Use of GIS and high resolution LiDAR in salt marsh restoration site suitability assessments in the upper Bay of Fundy, Canada." *Wetlands Ecology and Management* 21 (4):243–262.

Millard, K and M Richardson. 2013. "Wetland mapping with LiDAR derivatives, SAR polarimetric decompositions, and LiDAR–SAR fusion using a random forest classifier." *Canadian Journal of Remote Sensing* 39 (4):290–307.

Millard, K and M Richardson. 2015. "On the importance of training data sample selection in random forest image classification: A case study in peatland ecosystem mapping." *Remote Sensing* 7 (7):8489–8515.

Mitsch, WJ and JG Gosselink. 2008. *Wetlands*, 4th edn. John Wiley & Sons, Inc., Hoboken, NJ.

Moeslund, JE, L Arge, PK Bøcher, B Nygaard, and J-C Svenning. 2011. "Geographically comprehensive assessment of salt-meadow vegetation-elevation relations using LiDAR." *Wetlands* 31 (3):471.

Moffett, KB, SM Gorelick, RG McLaren, and EA Sudicky. 2012. "Salt marsh ecohydrological zonation due to heterogeneous vegetation–groundwater–surface water interactions." *Water Resources Research* 48 (2):W02516. DOI:10.1029/2011WR010874.

Morris, JT, D Porter, M Neet, PA Noble, L Schmidt, LA Lapine, and JR Jensen. 2005. "Integrating LIDAR elevation data, multi-spectral imagery and neural network modelling for marsh characterization." *International Journal of Remote Sensing* 26 (23):5221–5234.

Murphy, PNC, J Ogilvie, and P Arp. 2009. "Topographic modelling of soil moisture conditions: A comparison and verification of two models." *European Journal of Soil Science* 60 (1):94–109.

Murphy, PNC, J Ogilvie, K Connor, and PA Arp. 2007. "Mapping wetlands: A comparison of two different approaches for New Brunswick, Canada." *Wetlands* 27 (4):846–854.

Onojeghuo, A and G Blackburn. 2011. "Optimising the use of hyperspectral and LiDAR data for mapping reedbed habitats." *Remote Sensing of Environment* 8 (115):2025–2034

Qin, C-Z, A-X Zhu, T Pei, B-L Li, T Scholten, T Behrens, and C-H Zhou. 2011. "An approach to computing topographic wetness index based on maximum downslope gradient." *Precision Agriculture* 12 (1):32–43.

Quinton, WL, M Hayashi, and A Pietroniro. 2003. "Connectivity and storage functions of channel fens and flat bogs in northern basins." *Hydrological Processes* 17 (18):3665–3684.

Quinton, WL and NT Roulet. 1998. "Spring and summer runoff hydrology of a subarctic patterned wetland." *Arctic and Alpine Research* 30 (3):285–294.

Rampi, LP, JF Knight, and CF Lenhart. 2014. "Comparison of flow direction algorithms in the application of the CTI for mapping wetlands in Minnesota." *Wetlands* 34 (3):513–525.

Rapinel, S, L Hubert-Moy, and B Clément. 2015. "Combined use of LiDAR data and multispectral earth observation imagery for wetland habitat mapping." *International Journal of Applied Earth Observation and Geoinformation* 37:56–64.

Richardson, MC, M-J Fortin, and BA Branfireun. 2009. "Hydrogeomorphic edge detection and delineation of landscape functional units from lidar digital elevation models." *Water Resources Research* 45 (10) DOI: 10.1002/hyp.9322.

Richardson, M, S Ketcheson, P Whittington, and J Price. 2012. "The influences of catchment geomorphology and scale on runoff generation in a northern peatland complex." *Hydrological Processes* 26 (12):1805–1817.

Richardson, MC, CPJ Mitchell, BA Branfireun, and RK Kolka. 2010. "Analysis of airborne LiDAR surveys to quantify the characteristic morphologies of northern forested wetlands." *Journal of Geophysical Research: Biogeosciences* 115 (G3):G03005. DOI:10.1029/2009JG000972.

Riegel, JB, E Bernhardt, and J Swenson. 2013. "Estimating above-ground carbon biomass in a newly restored coastal plain wetland using remote sensing." *PLoS ONE* 8 (6):e68251.

Riley, JL. 2011. *Wetlands of the Ontario Hudson Bay Lowland: A Regional Overview*. Nature Conservancy of Canada, Toronto, Ontario, Canada, 156 pp.

Schmid, KA, BC Hadley, and N Wijekoon. 2011. "Vertical accuracy and use of topographic LIDAR data in coastal marshes." *Journal of Coastal Research* 27 (6A):116–132.

Sjörs, H. 1959. "Bogs and fens in the Hudson Bay lowlands." *Arctic* 12 (1):2–19.

Streutker, DR and NF Glenn. 2006. "LiDAR measurement of sagebrush steppe vegetation heights." *Remote Sensing of Environment* 102 (1):135–145.

Suchenwirth, L, M Förster, A Cierjacks, F Lang, and B Kleinschmit. 2012. "Knowledge-based classification of remote sensing data for the estimation of below-and above-ground organic carbon stocks in riparian forests." *Wetlands Ecology and Management* 20 (2):151–163.

Tinkham, WT, H Huang, A Smith, R Shrestha, MJ Falkowski, AT Hudak, TE Link, NF Glenn, and DG Marks. 2011. "A comparison of two open source LiDAR surface classification algorithms." *Remote Sensing* 3 (3):638–649.

Torbick, N, A Persson, D Olefeldt, S Frolking, W Salas, S Hagen, P Crill, and C Li. 2012. "High resolution mapping of peatland hydroperiod at a high-latitude Swedish mire." *Remote Sensing* 4 (7):1974–1994.

Töyrä, J, A Pietroniro, C Hopkinson, and W Kalbfleisch. 2003. "Assessment of airborne scanning laser altimetry (lidar) in a deltaic wetland environment." *Canadian Journal of Remote Sensing* 29 (6):718–728.

Wannasiri, W, M Nagai, K Honda, P Santitamnont, and P Miphokasap, 2013. Extraction of mangrove biophysical parameters using airborne LiDAR. *Remote Sensing* 5(4), 1787–1808.

Wehr, A and U Lohr. 1999. "Airborne laser scanning—An introduction and overview." *ISPRS Journal of Photogrammetry and Remote Sensing* 54 (2):68–82.

White, JC, NC Coops, MA Wulder, M Vastaranta, T Hilker, and P Tompalski. 2016. "Remote sensing technologies for enhancing forest inventories: A review." *Canadian Journal of Remote Sensing* 42 (5):619–641.

Wu, Q, C Lane, and H Liu. 2014. "An effective method for detecting potential woodland vernal pools using high-resolution LiDAR data and aerial imagery." *Remote Sensing* 6 (11):11444–11467.

Wu, Q and CR Lane. 2016. "Delineation and quantification of wetland depressions in the prairie pothole region of North Dakota." *Wetlands* 36 (2):215–227.

Zhang, C. 2014. "Combining hyperspectral and LiDAR data for vegetation mapping in the Florida Everglades." *Photogrammetric Engineering and Remote Sensing* 80 (8):733–743.

Zhang, K. 2008. "Identification of gaps in mangrove forests with airborne LIDAR." *Remote Sensing of Environment* 112 (5):2309–2325.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

16 Fraction Vegetation Cover Extraction Using High Spatial Resolution Imagery in Karst Areas

Xiangkun Qi, Chunhua Zhang, Yuhong He, and Kelin Wang

CONTENTS

16.1 INTRODUCTION

Karst is a type of landscape formed above carbonate bedrock. About 15% of global land has some form of karst landscape, either aboveground or belowground. Southwestern China contains one of the largest karst landscapes in the world, with an area of about 540,000 km^2 (Su, 2002). Karst is one of the most fragile landscapes worldwide (Legrand, 1973; Parise and Gunn, 2007). Meanwhile, the large population of southwestern China has frequently overexploited the sloped lands with inappropriate agricultural practices over the last half of the twentieth century, which has caused the rapid decline of ecological conditions in the karst area (Wen et al., 2011). Therefore, environmental problems, such as karst rocky desertification (KRD) (the loss of soil and vegetation cover and exposure of bedrock outcrop), water shortage, and land subsidence in karst regions, are becoming of increasing concern.

Vegetation, which serves as an indicator of environmental change and a primary producer of the global ecosystem, plays an important role in the matter as well as in energy cycles (Pregitzer and Euskirchen, 2004). It also promotes soil development, regulates the regional and local climate, and provides habitats for wildlife. In karst areas, vegetation is important for keeping soil from eroding, reducing the occurrence of KRD, and protecting regional water resources. Traditional vegetation studies based on field measurements are labor intensive, time consuming, and costly.

Furthermore, only sample point data are available from field surveys. Remote sensing technology has been widely applied to monitor environmental change and recovery in southwestern China (Jiang et al., 2014). Satellite sensors could acquire data for large areas in a short time and have a relatively low cost. Moreover, more than 40 years' worth of satellite image collections (i.e., Landsat imagery) and various spatial and spectral resolutions make remote sensing technology one of the most powerful tools for land surface monitoring. Another advantage of applying remote sensing techniques in vegetation monitoring is their ability to integrate with different sources for a better understanding of spatial and temporal changes in ecological dynamics of karst regions.

There has been increasing application of remote sensing technology to study karst vegetation dynamics and the progress of recovery, starting in the 21st century (Tong et al., 2016). Medium- and coarse-resolution satellite data have been applied to vegetation monitoring in karst regions. Commonly used medium- and coarse-resolution satellite imagery include the Landsat Thematic Mapper (TM) (e.g., Yang et al., 2011; Bai et al., 2013), ASTER (e.g., He et al., 2008; Xiong et al., 2013), MODIS, and SPOT-VGT images (e.g., Zhang et al., 2014; Wang et al., 2015; Tong et al., 2016). However, few researchers have discussed the rationale behind imagery resolution in karst areas. High spatial resolution images (e.g., SPOT-5 and IKONOS) are commonly used to extract ground truth data for accuracy assessment, but they have not been used for vegetation monitoring in this region.

The karst regions have high heterogeneity within the landscape. Some common landforms include towers, valleys, sinkholes, poljes, and cockpits (Zhang et al., 2011). The typical subtropical monsoon climate and carbonate bedrock environment create a unique surface-underground dual hydrology system. Although there is an ample amount of precipitation (average precipitation = 1000–1800 mm), surface water can flow into the underground water system through sinkholes and crevices. Consequently, it is common to see droughts in the rainy season in the karst area (Fan et al., 2011). One of the influences of this inflow is that surface soil moves into the underground water system. This makes the surface soil in karst regions shallow and patchy. These features—drought, lack of soil, and high intensity of anthropogenic disturbances—cause the strong heterogeneity within vegetation communities in the karst area (Li et al., 2003). High-resolution imagery could provide more detailed vegetation information that would be helpful in exploring the mechanisms of vegetation dynamics. From this perspective, high spatial resolution imagery is a promising prospect in karst vegetation monitoring.

Vegetation indices are commonly used to extract vegetation information, particularly the normalized difference vegetation index (NDVI), which has been widely applied to extract fraction vegetation cover (FVC) information (Wang et al., 2014; Xie et al., 2015). FVC is an important parameter for modeling carbon, water, and energy exchanges (Xiao and Moody, 2005). Considering the high heterogeneity of karst land surfaces, linear spectral mixture analysis (LSMA) is a practical method for extracting subpixel FVC and other fraction information from images. This approach assumes that the reflectance of one pixel is a mixture of the reflectance of features on the ground. The aim of LSMA is to separate the mixed pixel into a set of end-member spectra and estimate the percentage of end-members present in that pixel (fraction

images). LSMA has proven to be a useful method for identifying vegetation cover at the subpixel level for highly heterogeneous regions where vegetation is widely dispersed (Xiao and Moody, 2005).

A critical step for LSMA is end-member selection. Generally, LSMA hypothesizes that reflectance for one pixel is a linear mixture of end-members. The number of end-members must account for all land feature types in the pixel (Theseira et al., 2002). LSMA allows the same number of land feature types for each end-member, and each material has a solid spectral response. However, the same material could have different spectral curves; this method does not incorporate the variability in natural conditions. Multiple end-member spectral mixture analysis (MESMA) is widely used to account for within-class spectral variability (Roberts et al., 1998). MESMA allows each pixel to use different combinations of end-members and overcomes the LSMA limitation of using the same number of end-members to model all pixels. The dominant FVC extraction in the karst is the dimidiate pixel model, which is commonly based on NDVI. To the best of our knowledge, there has been no attempt to use the MESMA approach to extract FVC in karst areas.

Strong terrain relief is one of the most important features of karst landforms. The rough landscape in southwestern China has taken shape through alternating stages of relative stability and uplift in Cenozoic carbonate rock (Wang et al., 2004). Terrain relief produces substantial shadowing in satellite images (Figure 16.1). Variations in terrain conditions and solar illumination in mountainous regions can affect the extraction of vegetation information from satellite imagery (Gao and Zhang, 2009). Topography can cause bidirectional reflectance and shadow effects, consequently altering surface reflectance (Song and Woodcock, 2003). The same vegetation species may have different spectral curves (Yue et al., 2011). Likewise, different vegetation types could have similar spectral responses. Weak reflectance from shaded areas complicates the extraction of vegetation communities. Therefore, these topographical effects are a limiting factor when using remote sensing technology in karst regions. Band ratio is commonly used in the southwest of China to minimize the change in solar illumination caused by topography (Tong et al., 2014; Xie et al., 2015). The use of additional digital elevation model (DEM) data may also benefit vegetation monitoring (Zhang et al., 2014).

Topographic correction models, such as the C correction model and the Sun-Canopy-Sensor (SCS) model, have also been applied to Landsat imagery to minimize topographic effects (Qi et al., 2013). The results of topographic correction on Landsat imagery show very limited benefits for vegetation classification due to the large relief in karst peak and depression regions. Fine details in high spatial resolution imagery are often accompanied by noise, such as shadow effects. The topographic correction for high spatial resolution requires additional high-resolution DEM data, which are either expensive or hard to obtain. Therefore, reducing the shadow effects in high spatial resolution imagery is one of the most important preprocessing steps.

Our study explores the potential of the MESMA method to extract FVC in the karst area of southwestern China based on multispectral high spatial resolution ALOS data (10 m resolution). The aim is to verify the applicability of the spectral mixture analysis (SMA) approach in heterogeneous karst areas, and to analyze the impact of shadow effects on FVC extraction.

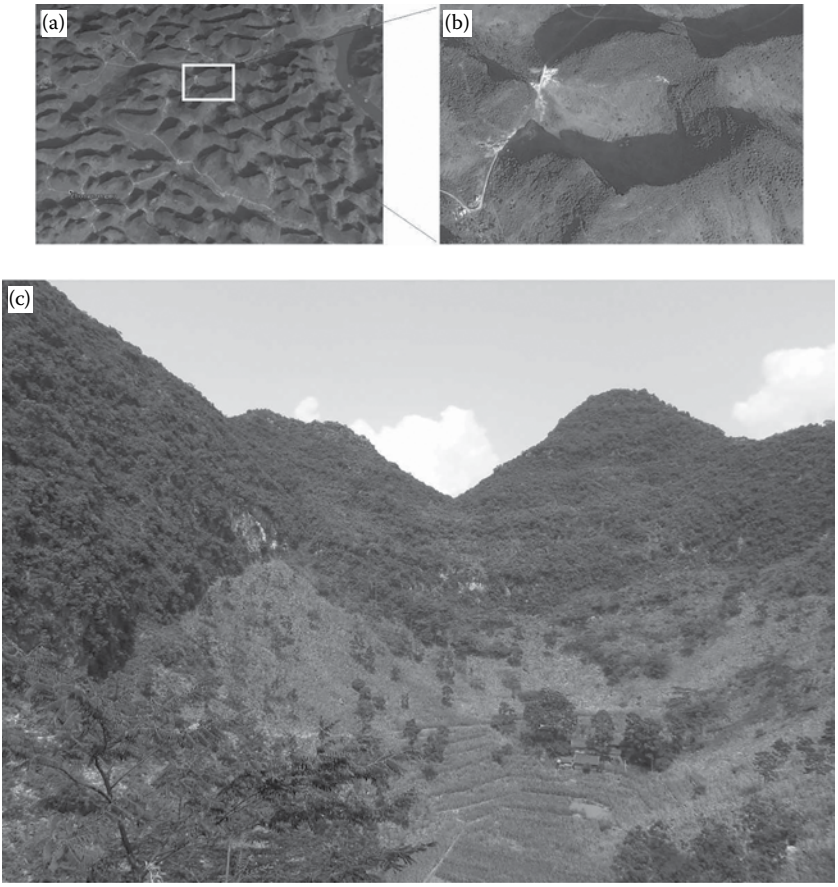


FIGURE 16.1 Shadow effects in the karst area. Satellite images (a) and (b) included a mass of shadow areas. (a) A section of a Landsat 8 image taken on December 30, 2014, in Dahua County, Guangxi Province. (b) A detailed view of a depression on a SPOT-5 image, taken on December 15, 2014, and extracted from Google Earth. (c) A photo, corresponding to (b), taken on July 22, 2011, during the field trip.

16.2 STUDY AREA

Our study area is located in southwestern China (Duan County, Guangxi Province). The topography is rugged, with elevation ranging between 230 and 870 m. The climate is warm moist subtropical, with mean annual precipitation of 1090–1920 mm and a mean annual temperature of 20°C. The maximum precipitation, recorded in the summer months, accounted for 45%–60% of the annual precipitation. Typical landforms in this region are tower karsts and depressions, with 60% of the study area having slope angles steeper than 25°. Although the dominant vegetation community in the area is subtropical evergreen forest, a mass of grass and shrub exists, and the landscape is fragmented because of severe human disturbances and geological limitations (Figure 16.2). This area has a relatively high population density and more

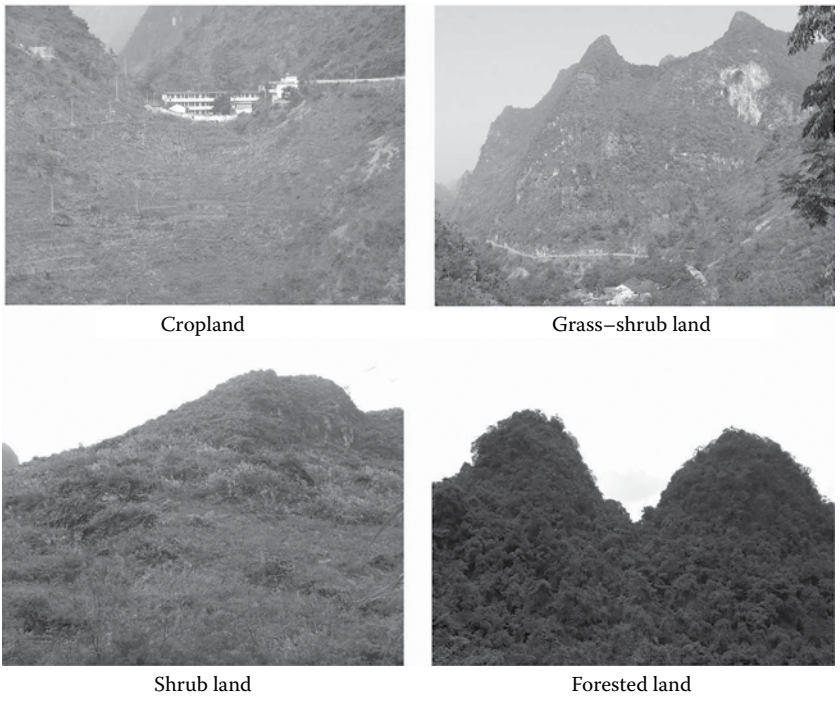


FIGURE 16.2 Vegetation landscapes in the karst area.

than 1000 years of agricultural development. Tree cutting for timber and excessive agricultural practices on the sloped land have led to the disappearance of the forested areas in the karst regions from the 1950s through the 1980s (Wen et al., 2011). To restore the degraded karst vegetation, some ecological restoration projects, such as the Grain for Green program and the banning of logging, were implemented in this fragile karst region.

16.3 DATA AND METHODS

The ALOS imagery (Figure 16.3) was acquired on June 4, 2009. The data contained four multispectral bands (blue, green, red, and near-infrared [NIR]) with 10 m spatial resolution, and a panchromatic band with 2.5 m spatial resolution. The data have a relatively large sun elevation angle (72.5°), which should reduce topographic effects on the images (Song and Woodcock, 2003). However, there was an obvious shadowed area on the image because of the rugged terrain relief, of which the average slope is 45°. All images went through L1G processing (systematic correction) and georeferencing, and then were projected to a Universal Transverse Mercator (UTM) map projection. The original multispectral and panchromatic imagery were fused by the Gram–Schmidt procedure in the ENVI software package (Laben and Brower, 2000) to produce four-band pan-sharpened multispectral ALOS imagery with 2.5 m spatial resolution for the quantitative assessment of linear spectral unmixing results.

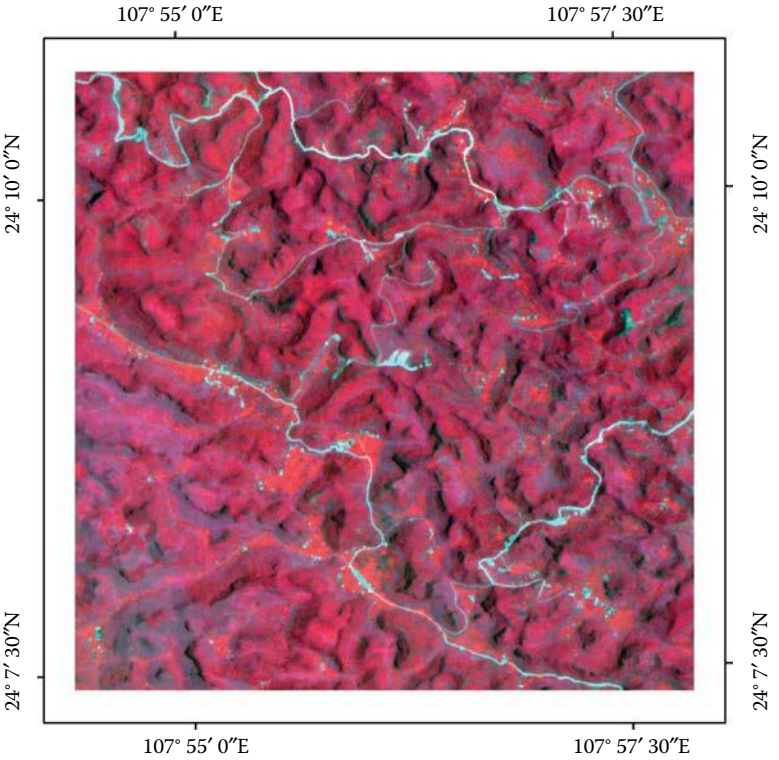


FIGURE 16.3 Multispectral ALOS imagery of the karst area (NIR, red, and green bands as R, G, B).

Traditionally, the end-members for one pixel in a karst region could be photosynthetic vegetation, nonphotosynthetic vegetation, shadow, water, bare soil, or bare rock (Wan and Cai, 2003; Yue et al., 2011). Nonphotosynthetic vegetation and bare soil are commonly found on land surfaces in the winter season. The preceding two classes were recorded in the growing season, when the nonphotosynthetic vegetation was hard to find and the bare soil was covered by crops. Therefore, only four types of end-members were available in the study area. We chose the vegetation-high albedo-shadow (vegetation-rock-shadow) model for spectral mixture analysis, after masking the water area.

MESMA was applied to extract fraction vegetation cover. This approach assumes that the spectrum measured by a sensor is a linear combination of the spectra of all components within the pixel (Adams et al., 1993). The mathematical model can be expressed as

$$R_{\mathrm{i}} = \sum_{k=1}^{n} f_k R_{ik} + \varepsilon_i \tag{16.1}$$

where i is the number of spectral bands used; k is 1, … , n (number of end-members); R_i is the spectral reflectance of band i of a pixel, which contains one or more

end-members; f_k is the proportion of end-member k within the pixel; R_{ik} is the spectral reflectance of end-member k within the pixel on band i; and ε_i is the error for band i. To solve for f_k, the following conditions must be satisfied: (1) selected end-members should be independent of each other, (2) the number of end-members should be less than or equal to the number of spectral bands used, and (3) selected spectral bands should not be highly correlated. Therefore, there were a maximum of three end-members in four-band ALOS data, theoretically. A common approach for obtaining f_k is to use a least-squares solution by minimizing the residual error. The sum of f_k of all optical end-members equals 1.

The spectral mixture analysis procedure presented here consists of two main steps: (1) selection of the optimal end-members to form the definitive spectral library, and (2) decomposing the mixed pixels to calculate the fraction images. Identifying a high-quality set of image end-members has been defined as a critical stage of spectral mixture modeling (Tompkins et al., 1997).

For the MESMA method, the end-member average root-mean-squared error (RMSE) (EAR) approach was used to select the most appropriate end-members. The end-members are selected by producing the lowest RMSE within a class (Dennison and Roberts, 2003). EAR can be expressed as shown in Equation 16.2:

$$\mathrm{EAR}_i = \frac{\sum_{i=1}^{N} \mathrm{RMSE}_{i,j}}{n-1} \tag{16.2}$$

where i is an end-member; j is the modeled spectrum; N is the number of end-members; and n is the number of modeled spectra. The -1 term corrects for the zero error resulting from an end-member modeling itself.

According to the most appropriate end-members, trials using just two optical end-members (vegetation and rock) showed that three end-members (vegetation, rock, and shadow) were needed to accurately model the ALOS images in the karst region. To reduce the shadow effects, we performed a shade normalization of the fraction images, obtained by dividing each end-member by the total percent cover of all nonshade end-members (1 − shade fraction) in each pixel. This suppresses the shade fraction so that we obtain more information on the relative abundance of nonshade end-members (Rogan and Franklin, 2001).

Accuracy assessment was conducted using error matrices. Kappa variance was used to measure the accuracy of the achieved percentage estimate of vegetation cover (Congalton and Green, 2009). Overall accuracy (OA) for each class was also calculated. The percentage of vegetation cover was classified into five levels (0%–30%, 31%–50%, 51%–70%, 71%–90%, and 91%–100%). Field validation sites were collected from 2007 to 2011. Percentage of vegetation cover was measured through visual estimation of cross walking and picture analysis. Most of the validation plots were located near the road or path because large relief mountain areas are hard to access. For measuring the vegetation cover in the shaded area, part of the reference points for accuracy assessments were determined by visual interpretation of pan-sharpened multispectral ALOS (2.5 m) imagery, as well as a historical land use and land cover maps. Spatial resolution imagery of 2.5 m from Google Earth, in which

the materials in the shaded area could be seen clearly, was also used to identify the reference points. Finally, a total of 368 reference points (which included 116 points in the shaded area) were collected to verify the accuracy of the fraction vegetation cover from the ALOS data.

16.4 RESULTS

As mentioned, reference points were selected in both south-facing and shaded areas. To decrease the error of the visual estimation in the field, we classified vegetation cover into five intervals, each representing 20%. Whenever a predicted value fell within this interval, it was verified as correct. Classification results illustrated that vegetation cover could be successfully identified using MESMA for ALOS imagery (Table 16.1). The classification of vegetation cover in the south-facing area had better results than in the shaded area. Overall, accuracies in the south-facing area and the shaded area were 85.7% and 58.6%, respectively. The overall accuracy reached 77.2%. Kappa coefficients were 0.82 and 0.48, respectively, and the total kappa value reached 0.71. Although the accuracy in the shaded area was lower than in the south-facing area, the mispredicted points (values) were close to the correct value. For example, when the referenced value was 31%–50% in the shaded area, the predicted value of 11 was correct. Two mispredicted percentages of vegetation cover were 0%–30%, and seven mispredicted percentages were 51%–70% (Tables 16.1 and 16.2).

Comparing the predicted values in the south-facing area and the shaded area, there is an obvious consistency. The largest difference is 7.6%, in the interval of 31%–50%. In total, the value in the shaded area is a bit larger than that in the south-facing area, except for the high vegetation cover area (91%–100%) (Table 16.2).

Compared with the fraction vegetation cover map predicted using MESMA and NDVI data (Figure 16.4), the values from MESMA visually reduced the shadow effects. There is a slight difference in values between the south-facing area and the shaded area, while these two areas had a similar vegetation cover. NDVI data were more impacted by terrain relief. The value in the shaded area is lower than in the south-facing area.

16.5 DISCUSSION

We selected the vegetation-rock-shadow model to unmix the ALOS imagery in the highly heterogeneous karst region. Though the ALOS data were acquired with a large sun elevation angle, there is a visible shadow area in the imagery. The spectral scatterplot of the imagery formed a visible triangle in this study and the three vertices of the triangle represent vegetation, rock, and shadow end-members, respectively. Some researchers found that there was nonphotosynthetic vegetation and bare soil that were easily confused with bedrock (Wan and Cai, 2003; Yue et al., 2011). However, these researchers often used satellite imagery from the winter. One of the reasons for this is that the growing season is also the rainy season in the karst region, and it can be challenging to acquire cloud-free satellite imagery in the summer. Unmanned aerial systems (UASs) make it possible to acquire timely high spatial resolution images whenever weather conditions are suitable (Zhang et al., 2017). The high

TABLE 16.1
Error Matrix of Vegetation Cover Classification

		Classified Data					Reference Totals
		0%–30%	31%–50%	51%–70%	71%–90%	91%–100%	
South-facing area, based on MESMA							
Reference data	0%–30%	45	6	0	0	0	51
	31%–50%	7	42	0	0	0	49
	51%–70%	0	2	44	3	0	49
	71%–90%	0	0	8	50	0	58
	91%–100%	0	0	0	10	35	45
	Classified totals	52	50	52	63	35	252
Shaded area, based on MESMA							
Reference data	0%–30%	13	5	0	0	0	18
	31%–50%	2	11	7	0	0	20
	51%–70%	0	5	15	5	0	25
	71%–90%	0	0	5	20	4	29
	91%–100%	0	0	0	15	9	24
	Classified totals	15	21	27	40	13	116
Total for study area, based on MESMA							
Reference data	0%–30%	58	11	0	0	0	69
	31%–50%	9	53	7	0	0	69
	51%–70%	0	7	59	8	0	74
	71%–90%	0	0	13	70	4	87
	91%–100%	0	0	0	25	44	69
	Classified totals	67	71	79	103	48	368

TABLE 16.2
Predicted Vegetation Cover (%) in the South-Facing Area and the Shaded Area

Vegetation Cover	0%–30%	31%–50%	51%–70%	71%–90%	91%–100%
South-facing area	18.0	38.1	59.9	75.3	92.0
Shaded area	18.3	45.7	60.5	80.2	88.8
Difference	−0.3	−7.6	−0.6	−4.9	3.2

spatial resolution UAS images would allow for much more detailed information to be extracted. Therefore, remote sensing experiments using UAS should be attempted in further studies in the karst region.

Fraction shadow cover was extracted using MESMA. We divided fraction vegetation cover and rock cover by the total percent cover of all nonshade end-members (1 − shade fraction) in each pixel, respectively. This shade normalization suppressed the shade fraction and obtained more information for nonshade end-members. Compared with NDVI data, this shade normalization assumed that vegetation and rock in shaded areas had similar spectral curves, and thus made a linear stretch to reduce shadow effects. The experiment results proved that this method surpassed nonlinear stretching,

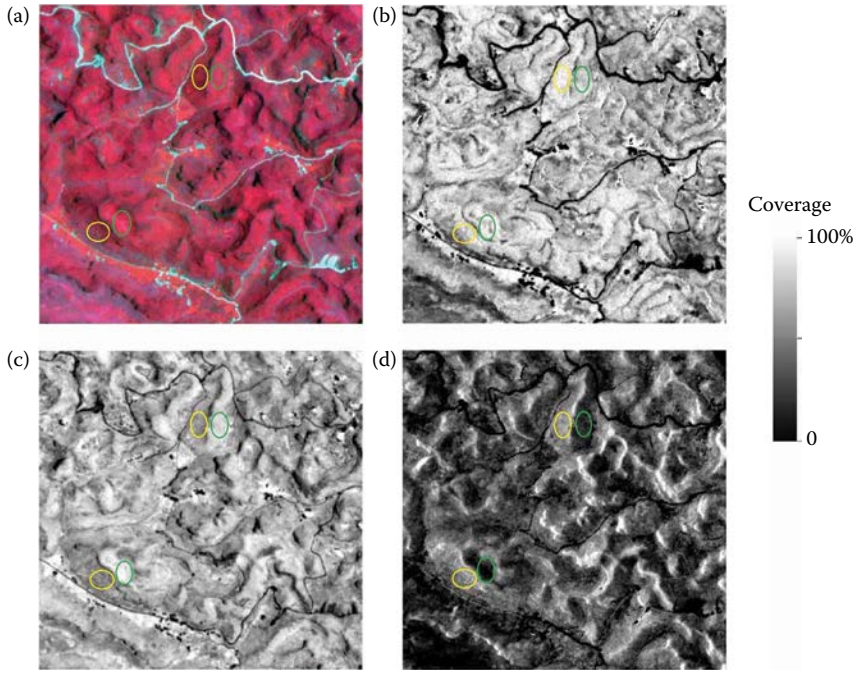


FIGURE 16.4 Contrast of vegetation cover in the south-facing area and the shaded area. (a), (b), (c), and (d) are ALOS imagery, with NIR, red, and green bands as R, G, B; fraction vegetation cover predicted using MESMA, NDVI, and fraction shadow cover, respectively. The yellow circle is the shaded area and the green circle is the south-facing area.

such as NDVI. However, some studies showed that there was a difference in spectral responses in the shaded areas, according to the different materials (Fitzgerald et al., 2005; Yang and He 2017). Fitzgerald et al. found that using two end-members (two spectral responses of different materials) unmixed better results than using only one end-member in the shaded area, based on MESMA. Yang and He (2017) recognized that the spectral scatterplots in the shaded area had a similar triangle shape to those in the south-facing area, and thus separated the nonshadow area and shadow area before using the SMA approach. However, Yang's (2017) research was mainly aimed at urban areas, where the shaded areas are regular and could be separated. In the mountainous region, the shadow showed gradient change. The shadow percentage within each pixel varied, causing this to not be a suitable method. In further studies, imagery with more bands (such as WorldView-3, which contains eight bands) should be used to enable the selection of more end-members and to improve the accuracy of vegetation extraction in highly heterogeneous regions.

16.6 CONCLUSION

Prediction of vegetation cover in the highly heterogeneous karst regions in southwestern China, using the MESMA approach based on high spatial resolution ALOS imagery, successfully extracted fraction vegetation cover. The overall accuracy for the study area is 77.2%. However, the accuracy of vegetation cover estimation in the shaded area is lower than in the south-facing area, in spite of the fact that the difference in vegetation cover is marginal in the shaded and south-facing areas. The lower estimation accuracy in the shaded area is likely caused by shadows. The MESMA method is capable of reducing shadow effects in the terrain relief region. Future studies should explore suitable methods for vegetation cover estimations in shaded areas and reduce topographic impact in the karst region.

REFERENCES

Adams, J. B., Smith, M. O., and Johnson, P. E. 1993. Imaging spectroscopy: Interpretation based on spectral mixture analysis. In: *Remote geochemical analysis: Elemental and mineralogical composition* (pp. 145–166). Cambridge, England: Cambridge University Press.

Bai, X. Y., Wang, S. J., and Xiong, K. N. 2013. Assessing spatial-temporal evolution processes of karst rocky desertification land: Indications for restoration strategies. *Land Degradation and Development*, 24(1), 47–56.

Congalton, R. G., and Green, K. 2009. *Assessing the accuracy of remotely sensed data: Principles and practices* (2nd ed.). Boca Raton: CRC Press.

Dennison, P. E., and Roberts, D. A. 2003. Endmember selection for mapping chaparral species and fraction using multiple endmember spectral mixture analysis. *Remote Sensing of Environment*, 41, 123–135.

Fan, F., Wang, K., Xiong, Y., Xuan, Y., Zhang, W., and Yue, Y. 2011. Assessment and spatial distribution of water and soil loss in karst regions, southwest China. *Acta Ecologica Sinica*, 31, 6353–6362 (in Chinese with English abstract).

Fitzgerald, G. J., Pinter Jr, P. J., Hunsaker, D. J., and Clarke, T. R. 2005. Multiple shadow fractions in spectral mixture analysis of a cotton canopy. *Remote Sensing of Environment*, 97(4), 526–539.

Gao, Y., and Zhang, W. 2009. A simple empirical topographic correction method for ETM+ imagery. *International Journal of Remote Sensing*, 30(9), 2259–2275.

He, X.Y., Wang, K.L., Zhang, W., Chen, Z.H., Zhu, Y.G., Chen, H.S. 2008. Positive correlation between soil bacterial metabolic and plant species diversity and bacterial and fungal diversity in a vegetation succession on Karst. *Plant Soil*, 307, 123–134.

Jiang, Z., Lian, Y., and Qin, X. 2014. Rocky desertification in Southwest China: Impacts, causes, and restoration. *Earth-Science Reviews*, 132, 1–12.

Laben, C. A., and Brower, B. V. 2000. Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening. U.S. Patent No. 6,011,875, Eastman Kodak Company.

Legrand, H.E. 1973. Hydrological and ecological problems of karst regions. *Science*, 179, 859–864.

Li, X., Su, Z., Lu, S., Ou, Z., Xiang, W., Ou, Z., and Lu, S. 2003. The spatial pattern of natural vegetation in the karst regions of Guangxi and the ecological signality for ecosystem rehabilitation and reconstruction. *Journal of Mountain Science*, 21, 129–139 (in Chinese with English abstract).

Parise, M., and Gunn, J. 2007. Natural and anthropogenic hazards in karst areas: Recognition, analysis and mitigation. *Geological Society of London*, 279, 1–3.

Pregitzer, K. S., and Euskirchen, E. S. 2004. Carbon cycling and storage in world forests: Biome patterns related to forest age. *Global Change Biology*, 10 (12), 2052–2077.

Qi, X., Wang, K., and Zhang, C. 2013. Effectiveness of ecological restoration projects in a karst region of southwest China assessed using vegetation succession mapping. *Ecological Engineering*, 54, 245–253.

Roberts, D. A., Gardner, M., Church, R., Ustin, S., Scheer, G., and Green, R. O. 1998. Mapping chaparral in the Santa Monica Mountains using multiple endmember spectral mixture models. *Remote Sensing of Environment*, 65, 267–279.

Rogan, J., and Franklin, J. 2001. Mapping wildfire burn severity in southern California forests and shrublands using enhanced thematic mapper imagery. *Geocarto International*, 16, 91–101.

Song, C. H., and Woodcock, C. E. 2003. Monitoring forest succession with multitemporal Landsat images: Factors of uncertainty. *IEEE Transactions on Geoscience and Remote Sensing*, 41, 2557–2567.

Su, W. C. 2002. Controlling model for rocky desertification of karst Changes in aboveground biomass and diversity between different stages of secondary succession of a karst vegetation in Guangxi, China. *Journal of Soil Water Conservation*, 16(2), 29–32 (in Chinese with English abstract).

Theseira, M. A., Thomas, G., and Sannier, C. A. D. 2002. An evaluation of spectral mixture modeling applied to a semi-arid environment. *International Journal of Remote Sensing*, 23, 687–700.

Tompkins, S., Mustard, J. F., Pieters, C. M., and Forsyth, D. W. 1997. Optimization of endmembers for spectral mixture analysis. *Remote Sensing of Environment*, 59, 472–489.

Tong, X., Wang, K., Brandt, M., Yue, Y., Liao, C., and Fensholt, R. 2016. Assessing future vegetation trends and restoration prospects in the karst regions of southwest China. *Remote Sensing*, 8(5), 357.

Tong, X., Wang, K., Yue, Y., Liao, C., Xu, Y., and Zhu, H. 2014 Trends in vegetation and their responses to climate and topography in northwest Guangxi. *Acta Ecologica Sinica*, 34, 3425–3434 (in Chinese with English abstract).

Wan, J., and Cai, Y. 2003. Apply linear spectral mixture analysis to identify land cover changes in karst region—A case study in Guanling, Guizhou. *Geographical Research*, 22, 439–446 (in Chinese with English abstract).

Wang, J., Wang, K., Zhang, M., and Zhang, C. 2015. Impacts of climate change and human activities on vegetation cover in hilly southern China. *Ecological Engineering*, 81, 451–461.

Wang, M., Liang, X., Wei, X., and Li, X. 2014. Landscape pattern analysis on change of fractional vegetation cover between karst and no-karst areas: A case study in Hechi District, Guangxi Zhuang Autonomous Region. *Acta Ecologica Sinica*, 34, 3435–3443 (in Chinese with English abstract).

Wang, S. J., Li, R. L., Sun, C. X., Zhang, D. F., Li, F. Q., Zhou, D. Q., Xiong, K.N., and Zhou, Z. F. 2004. How types of carbonate rock assemblages constrain the distribution of karst rocky desertified land in Guizhou Province, PR China: Phenomena and mechanisms. *Land Degradation and Development*, 15(2), 123–131.

Wen, Y. G., Sun, D. J., Zhu, H. G., Liu, J. T., Liu, S. R., and Shi, Z. M. 2011. Changes in aboveground biomass and diversity between different stages of secondary succession of a karst vegetation in Guangxi, China. *Advances in Biomedical Engineering*, 1–2.

Xiao, J., and Moody, A. 2005. A comparison of methods for estimating fractional green vegetation cover within a desert-to-upland transition zone in central New Mexico, USA. *Remote Sensing of Environment*, 98(2), 237–250.

Xie, X., Xue, Z., Wang, D., Liu, W., and Du, P. 2015. Land cover classification in karst regions based on phenological features derived from a long-term remote sensing image series. *Journal of Remote Sensing*, 19, 627–638 (in Chinese with English abstract).

Xiong, Y., Yue, Y., and Wang, K. 2013. Comparative study of indicator extraction for assessment of karst rocky desertification based on Hyperion and ASTER images. *Bulletin of Soil and Water Conservation*, 33, 186–190 (in Chinese with English abstract).

Yang, J., and He, Y. 2017. Automated mapping of impervious surfaces in urban and suburban areas: Linear spectral unmixing of high spatial resolution imagery. *International Journal of Applied Earth Observation and Geoinformation*, 54, 53–64.

Yang, Q. Q., Wang, K. L., Zhang, C., Yue, Y. M., and Tian, R. C. 2011. Spatio-temporal evolution of rocky desertification and its driving forces in karst areas of Northwestern Guangxi, China. *Environmental Earth Sciences*, 64(2), 383–393.

Yue, Y., Wang, K., Zhang, B., Liu, B., Chen, H., and Zhang, M. 2011. Uncertainty of remotely sensed extraction of information of karst rocky desertification. *Advances in Earth Science*, 26, 266–274 (in Chinese with English abstract).

Zhang, C., Qi, X., Wang, K., Zhang, M., and Yue, Y. 2017. The application of geospatial techniques in monitoring karst vegetation recovery in southwest China: A review. *Progress in Physical Geography*, 41(4), 450–477.

Zhang, M., Wang, K., Zhang, C., Chen, H., Liu, H., Yue, Y., Luffman, I., and Qi, X. 2011. Using the radial basis function network model to assess rocky desertification in northwest Guangxi, China. *Environmental Earth Sciences*, 62(1), 69–76.

Zhang, Y., Zhou, Z., Ma, S., and Zhang, Y. 2014. Response of vegetation to climate change in karst region based on NDVI—A case study at Liupanshui City in Guizhou Province. *Bulletin of Soil and Water Conservation*, 34, 114–118 (in Chinese with English abstract).

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

17 Using High Spatial Resolution Imagery to Estimate Cherry Orchard Acreage in Michigan

Kin M. Ma

CONTENTS

17.1 INTRODUCTION

This chapter will showcase how high spatial resolution imagery and related image processing techniques and methods can be applied to an agricultural context. Within the United States, the Michigan agricultural industry was valued at $7 billion in 2011 (MDARD 2015), and by the year 2016 it had nearly doubled, resulting in a $13 billion impact (MDARD 2017). The fruit industry has a $758 million impact and is important in feeding Michigan's economy (MDARD 2017). In 2016, Michigan cherries generated $72 million, which was 9.5% of Michigan's fruit industry (NASS 2017). In 2005, Michigan tart cherry orchards produced 77% of all the United States' tart cherry goods (Pollack and Perez 2009). Located in the northwest region of Michigan, the Grand Traverse Bay region has the highest concentration of cherry orchards in the entire state, and this region spans from Northport south to Traverse City, to the Old Mission Point peninsula, and then to East Bay's Elk Rapids region (Figure 17.1). The combined factors of Lake Michigan's lake effect snow and the moist mesoclimate of this area contribute to the high fertility of this cherry growing region.

FIGURE 17.1 Study area of Grand Traverse Bay, Michigan.

Since the early part of the twenty-first century, high-resolution commercial satellites, such as IKONOS and QuickBird, have produced 1 m or finer resolution images (Lillesand and Kiefer 2008). The IKONOS satellite has a 1.0 m panchromatic and a 4.0 m multispectral spatial resolution. With these high-resolution images, forest and agricultural inventories and monitoring can be effectively performed. Coops et al. (2006) utilized high-resolution QuickBird imagery to detect the damage to forests by mountain pine beetle infestations in British Columbia, Canada. The use of high-resolution imagery was able to easily identify tree crowns that had been infected by the pine beetle, by using a red-green index (RGI), which is the ratio of the reflectance between the red and green channels. The RGI can then signify the color changes that happen at the leaf level, in which diseased pine needles change from healthy green to diseased red (Coops et al. 2006).

In addition, in central California's Marin County, high-resolution airborne data acquisition and registration (ADAR) imagery (1 m) were used to identify dead tree crowns that had succumbed to Sudden Oak Death. They were able to extract the dead tree crowns from surrounding bare soil patches by introducing a region-based

subtraction algorithm to distinguish the surrounding bare soil patches from true dead tree crown patches (Sun et al. 2005).

Furthermore, Pasher and King (2009) utilized hybrid classification techniques, object-based classification (OBC), and spectral unmixing to detect canopy-level deadwood objects and had achieved field-validated accuracy of 94%. Since 2009, for remote sensing analyses of urban and agricultural regions, there has been an increased use of OBC techniques for image analyses and information extraction (Myint et al. 2011; Sasaki et al. 2012; Lizarazo 2014).

In regards to cherry fruit production, Imanishi et al. (2010) have investigated the use of reflectance and absorption spectra to determine the chlorophyll content of blooming cherry plants. They determined that a specific leaf chlorophyll index that incorporated a "difference and ratio combination type may be a useful method of chlorophyll content estimation" (Imanishi et al. 2010). Therefore, these remote sensing techniques were used to estimate the vibrancy of the chlorophyll in healthy cherry trees. Also, crop load estimates of sweet cherry orchards have been investigated by Santesteban et al. (2007).

In another agricultural application, Panda et al. (2016) used high-resolution 1 m National Agricultural Imagery Program (NAIP) imagery to help identify blueberry orchards in southeastern Georgia. They utilized eCognition software and object-based image analyses (OBIAs) for analyzing and classifying the textural row patterns of the planted blueberry shrubs, and also differentiated the maturity growth levels of these shrubs.

Cherry orchard trees are also planted in a row-length pattern and the largest concentration of Michigan cherries are located in the northwestern Michigan region of Grand Traverse Bay. The use of high-resolution imagery and OBIA for cherry orchard acreage estimation has yet to be completed for this important agricultural region. One of the goals of this research was to estimate crop acreage because the U.S. Department of Agriculture's (USDA's) Marketing Order 930, which mandates that cherry fruit growers monitor and estimate their annual crop acreage (Agricultural Marketing Service 2012). This project aims to establish a baseline of cherry orchard crop acreage from previous historical reporting data and then develop specific algorithms and classification techniques to estimate the 2011 season cherry orchard crop acreage.

17.2 STUDY AREA AND METHODS

17.2.1 Study Area

The northwest region of Michigan has the highest concentration of cherries in the state. With the expertise of our research team, the study area of 152,000 acres was selected across the eastern side of the Leelanau Peninsula, from the Old Mission Point Peninsula of Grand Traverse County, to the Elk Rapids region on the western section of Antrim County (Figure 17.1). In order to best capture the acreage and vitality of this cherry tree growth, IKONOS image acquisition and field research occurred on the same day, July 17, 2011. During July 17–18, 2011, our research team traveled to randomly distributed representative cherry orchard farms throughout this region. At each orchard location, the

research team, using a Trimble Juno global positioning system (GPS) receiver, acquired 10–12 GPS point samples of cherry orchard trees in a rectangular pattern, beginning on an outside row, and moving four trees down that row, then continuing inward two rows, and coming back and then returning to the initial sample tree. Photographs were also taken of the cherry trees at each field site. Healthy cherry trees averaged 2.5–3.0 m (8.2–10.0 ft) in height, and mature cherry trees averaged 5 m in diameter. The field sites averaged about 1,000 m^2 in size. A total of 25 field sites and 250 GPS points were collected. Since these collected field GPS points represented a small section of the study area, they were not used for reference and accuracy assessment, but as supplemental information to determine land cover classification categories.

17.2.2 Methods

High-resolution IKONOS 1.0 m panchromatic and 4.0 m multispectral 16-bit radiometric images were acquired for Michigan's Grand Traverse Bay region during the height of the harvest period, July 17, 2011 (Figure 17.1). The IKONOS 4.0 m multispectral imagery has four electromagnetic bands including blue, green, red, and infrared sections of the spectrum. The red and infrared bands are especially useful in distinguishing vegetated areas such as cherry orchards. The finer 1.0 m panchromatic resolution can delineate the texture of the orchard tree crowns. In order to maximize the advantage of the 1.0 m panchromatic image, the 4.0 m multispectral images were fused with the mosaicked panchromatic image through an ERDAS IMAGINE pan-sharpening process.

The ERDAS IMAGINE 2013 and OBIA eCognition 9.2 software packages were used to process the pan-sharpened IKONOS image with a large 13,000 MB size, consisting of processing hundreds of millions of pixels. Even with a very fast 64-bit server with 128 GB of available RAM, the eCognition processing of these large, complex images has caused the program to freeze and crash because of the very large total size of the mosaicked and pan-sharpened images. Therefore, these large images were subdivided into six small subsections for processing. For the case study within this chapter, image processing and analyses focused on the eastern section of the Grand Traverse Bay within Antrim County (Figure 17.1).

From previous image processing experience, we know that dense forest stands would often be misclassified as very dense orchard trees because they were spectrally and texturally similar. Therefore, prior to eCognition's image processing, the East Bay image from Figure 17.1 was subset by ERDAS IMAGINE 2013 to exclude natural forest stands on the western section of the original Figure 17.1 IKONOS image. The exclusion of forest regions, which would affect cherry orchard acreage estimation, reduced the size of the East Bay image by about 15%. Figure 17.2 gives the outline of the IKONOS image that was processed.

Through trial and error, and with a goal to optimize the classification of the IKONOS pan-sharpened images at the orchard level, eCognition 9.2's segmentation process was conducted on the East Bay image with the following parameters: scale = 150, shape = 0.7, and compactness = 0.5. In OBIA, the scale parameter captures the relative size of the landscape feature, and the shape parameter focuses on the relative shape of the landscape feature, i.e., long rectangular roads versus round cherry tree crowns.

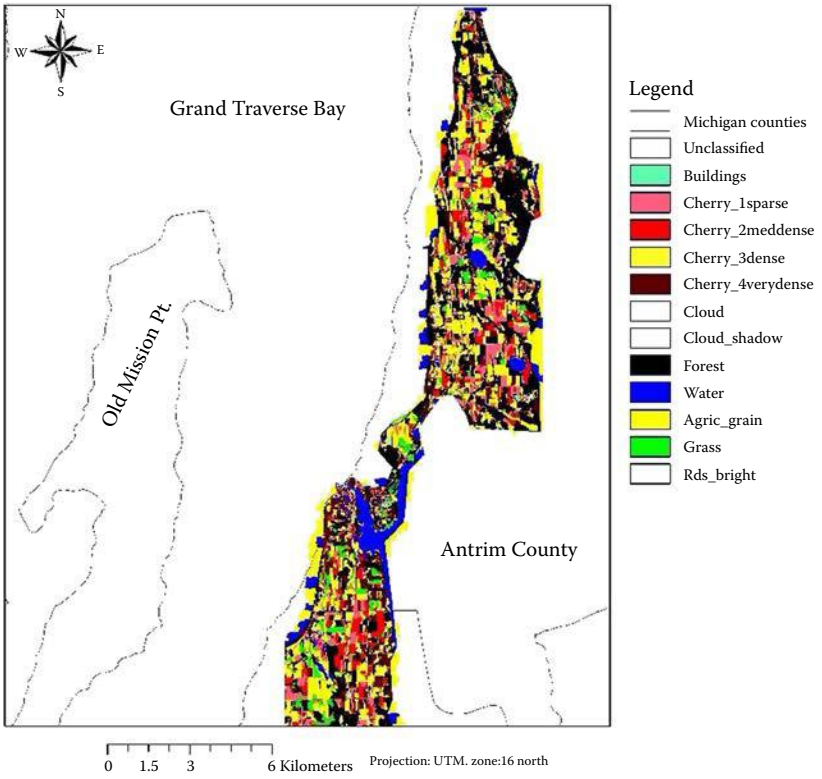


FIGURE 17.2 eCognition segmented classification with land cover categories.

Supervised classification, a nearest neighbor classifier, was then performed to divide the segmented image into 12 class categories: (1) grain agriculture; (2) buildings and driveways; (3) roads and impervious surfaces; (4) grass; (5) cloud; (6) cloud shadow; (7) water; (8) forest, which combines both deciduous and coniferous forests; and four categories that include the four density levels of orchard patterns (Table 17.1), (9) Cherry_1 sparse, (10) Cherry_2 medium density, (11) Cherry_3 dense, and (12) Cherry_4 very dense. The Cherry_1 sparse orchard density was characterized by orchard stands with dead and very young trees and open soil and grassy areas; Cherry_2 medium density had some rows of missing trees, though it was approximately 50% full; the Cherry_3 dense orchard category had various polygons filled with 75% of the orchard tree rows; and the Cherry_4 very dense orchard classification category had mature orchard stands with large crown canopies that touched the adjacent orchard tree, and more than 75% of the orchard tree rows were covered with orchard trees. Since there were no clouds or cloud shadows in this image, these assigned categories (Categories 5 and 6) were eliminated from the reference sample categories.

For each classification category, representative sample polygons (i.e., segments) were visually identified throughout the image. The eCognition software would ingest the data statistics of the sample polygons such as the median multispectral reflectance

TABLE 17.1
Area Estimates for Elk Rapids Region, Antrim County, Michigan

Class Name	Pixels	Hectares	Acres	Percentage
Agric	19,186,571	1,918.66	4,741.00	21.62
Buildings	1,008,216	100.82	249.13	1.14
Cherry_1	8,110,962	811.10	2,004.22	9.14
Cherry_2	8,520,242	852.02	2,105.35	9.60
Cherry_3	6,463,879	646.39	1,597.22	7.28
Cherry_4	5,883,904	588.39	1,453.91	6.63
Mixed Forest	28,062,168	2,806.22	6,934.16	31.62
Grass	3,791,744	379.17	936.94	4.27
Roads	973,794	97.38	240.62	1.10
Water	6,760,247	676.02	1,670.46	7.62
Total Estimates	88,761,727	8,876.17	21,933.02	100.00

Buildings = buildings and impervious surfaces; Cherry_1 = sparse low-density cherry orchards; Cherry_2 = medium-density cherry orchards; Cherry_3 = dense cherry orchards; Cherry_4 = very dense cherry orchards; Agric = grain agriculture and bare soil; Forest = deciduous, coniferous, and mixed forest; Grass = grass and golf course greens; and Roads = roads and driveways.

values and the standard deviation of the reflectance values. These statistic values were then used by the nearest neighbor classifier to assign classification categories to all of the remaining polygon objects. After all of the image's polygon objects were classified, the completed object-level classification category names and related data statistics were then exported into a raster ERDAS IMG format. Within ERDAS, the accuracy assessment module was launched to evaluate the accuracy of the classified image. Within the module, 127 stratified random points were generated. Because some of the class land cover classification categories, such as roads and buildings, only occupied a small area within the classified image, a stratified random reference sampling was chosen to better represent the diversity and dominance of the natural land cover types (McCoy 2005). All of the random reference points were then visually identified and assigned to the land cover categories referred to previously. Then the accuracy assessment module generated a report on the producer's, user's, and overall accuracy between the classified and reference land cover category names (Congalton 1991).

With eCognition, the classified layer was exported out to a shapefile to be processed with ArcGIS 10.4, and within ArcGIS the shapefile was dissolved according to the land cover name and spatial statistics were generated for display of the estimated acreage for each of the classification categories, as well as the various cherry orchard density levels. The cherry orchard acreage estimate amounts were then compared with the U.S. Department of Agriculture's 2007 and 2012 Fruits and Nuts reports, detailing the acreage amounts of agricultural orchards (Census of Agriculture 2007, 2012a,b,c). The 2011 cherry orchard acreage estimates were compared with data from both censuses because the estimates were not within one of the agricultural census years.

17.3 RESULTS AND DISCUSSION

17.3.1 Results

After the image segmentation of the East Bay region using the parameters scale = 150, shape = 0.7, and compactness = 0.5, eCognition generated a segment layer with 14,978 object polygons within the East Bay region. When this segmented image was classified and exported to a raster-based shapefile, the estimated percentage of 10 land cover types were calculated and displayed in Table 17.1.

When the estimated acreages of the four density levels of cherry orchards were added together, their sum totaled 32.65% of this image within the Elk Rapids region of Antrim County. This clearly shows the high concentration of cherry orchards in this northwestern region of Michigan because nearly 1/3 of this region has cherry production, and 13.91% of this area has Cherry_3 dense and Cherry_4 very dense cherry orchards.

When the ERDAS IMAGINE 2013 accuracy assessment of the classified image was run on the East Bay region image, it produced an overall classification accuracy of 74.80%, with a kappa statistic of 0.678 (Table 17.2). Within the Cherry_1 sparse orchard category, the user's accuracy was only 33%. The polygon objects that were classified within the Cherry_1 sparse category have a very low percentage of orchard trees, and more than 75% of the areas within the polygons had fallow land or dead trees. This larger proportion of bare soil or grassy vegetation would be more similar to the reflectance characteristics of land devoted to grain agriculture, like corn, oats, or rye within Antrim County (Census of Agriculture 2012a,b,c). Since the segmentation scale parameter was set to 150, some of the polygon objects were large, and multiple land cover and land use types may occur within a large polygon object.

For the Cherry_2 medium-density orchard category, the producer's accuracy was 87.5%, though the user's accuracy was 53.85% (Table 17.2). A close investigation disclosed that some polygons were classified as Cherry_2 medium-density orchard, but were labeled as the Cherry_3 dense category. Since the polygons can span two adjacent orchards, it may be difficult to distinguish a Cherry_2 medium-density orchard, having about 50% orchard trees, and the next adjacent category having 75% orchard trees in the Cherry_3 dense orchard category. At two additional reference points, they were designated forest instead of the Cherry_2 medium-density category.

The Cherry_3 dense category was classified most accurately, with both 77.78% for the producer's and user's accuracy percentage. This orchard density was most easily identifiable because 75% of the polygon objects had orchard trees, and the ratio of bare soil or nonorchard vegetation was relatively low. On the other hand, when the density of the orchard becomes very dense, within the Cherry_4 very dense category, both the user's and producer's accuracy were lower at 44.44% (Table 17.2). Very dense orchards have often been misclassified as forest since the older, mature cherry orchard tree crowns have begun to touch adjacent tree crowns filling up the spaces, and this was demonstrated by the continuous leafy canopy from one end of the orchard row to the far end. These mature fully crowned orchard trees look very similar to a dense forest stand of deciduous trees (see the green forest colored polygon in the middle of Figure 17.3a).

In another misclassification case, the segmentation scale and pattern adversely affected the classification. At the eCognition scale factor of 150, larger polygons

TABLE 17.2
Classification, Producer's and User's Accuracy of Elk Rapids Region, Antrim County, Michigan

	Reference												
Classification	**Buildings**	**Cherry_1**	**Cherry_2**	**Cherry_3**	**Cherry_4**	**Forest**	**Water**	**Agric**	**Grass**	**Roads**	**TOTAL**	**Producer's Accuracy (%)**	**User's Accuracy (%)**
Buildings	2	0	0	0	0	0	0	0	0	0	2	66.67	100.00
Cherry_1	0	4	0	0	0	2	0	5	0	1	12	100.00	33.33
Cherry_2	0	0	7	2	0	2	0	1	0	1	13	87.50	53.85
Cherry_3	0	0	0	7	1	1	0	0	0	0	9	77.78	77.78
Cherry_4	0	0	0	0	4	4	0	1	0	0	9	44.44	44.44
Forest	0	0	1	0	3	41	0	0	2	0	47	78.85	87.23
Water	0	0	0	0	0	1	2	0	0	0	3	100.00	66.67
Agric	0	0	0	0	0	0	0	22	0	1	23	75.86	95.65
Grass	0	0	0	0	1	1	0	0	5	0	7	71.43	71.43
Roads	1	0	0	0	0	0	0	0	0	1	2	25.00	50.00
TOTAL	3	4	8	9	9	52	2	29	7	4	127		

Classification accuracy = 74.80% (stratified random sampling); overall kappa statistics = 0.678.

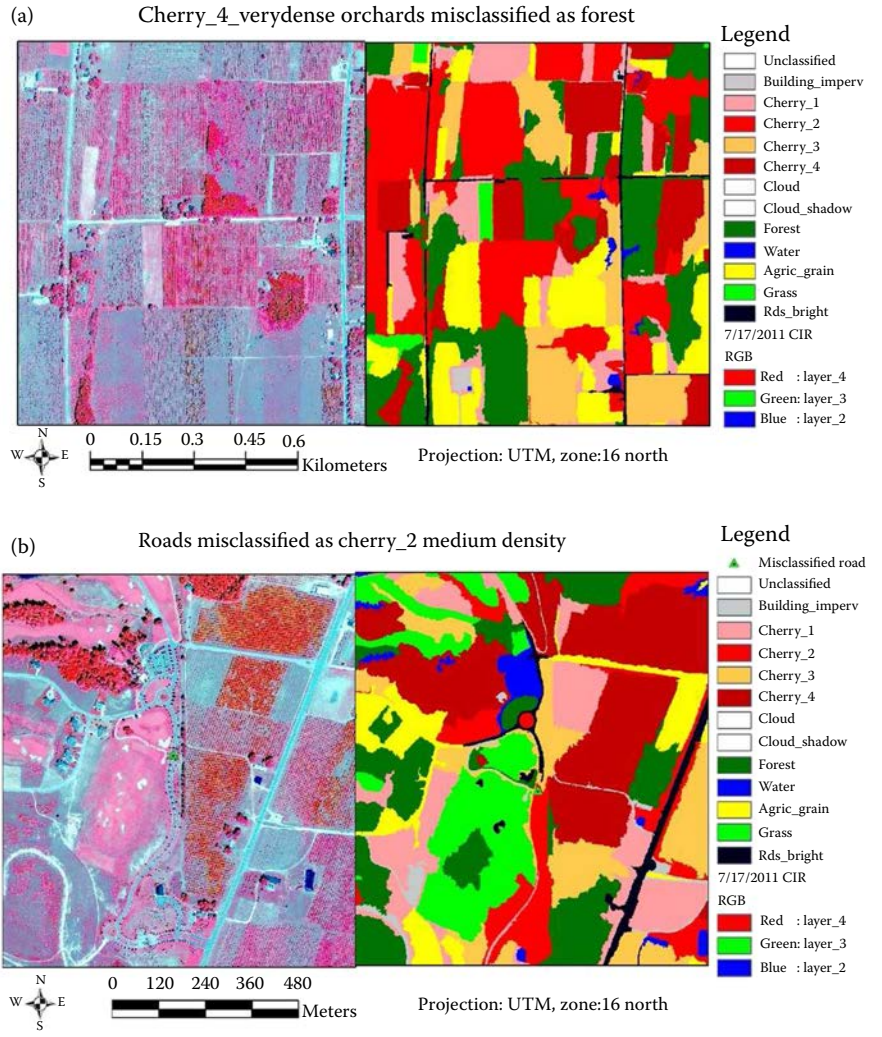


FIGURE 17.3 (a) left is color infrared, and right is an eCognition segmented classification. Toward the center of the image is an area of very dense cherry orchards that were misclassified as forest. (b) left is color infrared, and right is an eCognition segmented classification. Toward the center of the image is an area of roads, lined with trees, that were misclassified as Cherry_2 medium density.

may span multiple land use zones. In a forested suburban road, with many trees that were adjacent to a Cherry_2 density category orchard, the eCognition program is tricked in classifying it as a Cherry_2 orchard polygon (Figure 17.3b). Visually, an experienced researcher would not consider five to six trees lining a street as a cherry orchard, but the eCognition program could not make this distinction. Further segmenting undersegmented objects at a finer scale could potentially minimize this type of misclassification.

TABLE 17.3
Cherry Orchard Acreage Estimate and USDA Comparison for Antrim County, Michigan

Class Name	Acres	Acres Adjusted in 2007	2007 USDA Acreage Data[a]	Acres Adjusted in 2012	2012 USDA Acreage Data[b]
Cherry_1	2,004.22	1,615.0		1,697.4	
Cherry_2	2,105.35	1,696.5		1,783.0	
Cherry_3	1,597.22	1,287.0		1,352.7	
Cherry_4	1,453.91	1,171.6		1,231.3	
Total Orchard Acre Estimates	7,160.71	5,770.1		6,064.4	
USDA Antrim County Cherry Acreage			5,954.0		5,889.0
Percentage Difference from USDA		−3.09%		2.98%	

[a] Census of Agriculture (2007).
[b] Census of Agriculture (2012c).
Cherry_1 = sparse low-density cherry orchards; Cherry_2 = medium-density cherry orchards; Cherry_3 = dense cherry orchards; Cherry_4 = very dense cherry orchards.

17.3.1.1 Cherry Orchard Estimated Acreage Comparisons

For this study, the image analyses focused on the orchard level, but not necessarily to distinguish between the various fruit or tree orchard varieties. Therefore, after adding the acreage amounts of the four cherry orchard density levels, there was a total estimate of 7,160.71 orchard acres. Within this East Bay region of Antrim County, even though cherry crop production is the dominant commodity type, other row pattern orchards, including apple, plum, and cut Christmas trees, were present. The USDA Census of Agriculture also had acreage data for these non-cherry orchard types, and the sum of the 2007 census statistics tallied to 1,435 non-cherry acres, or 19.42% of Antrim County's orchard acres (Census of Agriculture 2007). Therefore, if the acreage estimates for the 2011 cherry orchards utilized the statistics from the 2007 census (Cherry Industry 2011), this study's cherry orchard estimate of 7,160.71 acres would be lowered by 19.42%, down to 5,770.1 acres. When the 5,770.1 estimated acres is compared with the 2007 USDA cherry orchard acreage of 5,954.0 acres, there is a small, −3.09%, underestimate of the USDA's calculation (Table 17.3).

However, if the acreage estimates for the 2011 cherry orchards were based on the non-cherry orchard statistics from the 2012 census, this study's 7,160.71 estimated acres would be lowered by 15.31%, to an adjusted estimate of 6,064.4 acres. When analyzing the adjusted cherry orchard acreage and comparing it with the 2012 USDA cherry orchard acreage amount of 5,889.0 acres, there is also a small 2.98% overestimation of the USDA acreage amount (see "Classified Image" in Table 17.3).

In addition, there are annual fluctuations in cherry production acreage from the 2007 and 2012 censuses. Various fruit growers may switch to a different crop, or choose to leave some acres fallow, or not harvest from low-producing cherry trees. The 2011 classification and orchard acreage estimate is one snapshot in time, and can

be used in a historical context to guide future cherry growing trends. The full view of the 2011 classification for the East Bay area is shown in Figure 17.2.

17.4 CONCLUSION

High spatial resolution imagery combined with OBIAs and techniques have been effective in estimating the cherry orchard acreage in the Grand Traverse Bay region of Northwest Michigan, and specifically for this case study in Antrim County.

There were some instances that very dense orchard plots were misclassified as deciduous forest stands, though the final adjusted cherry orchard acreage estimates were remarkably similar and within a ±3.1% error margin of the USDA's 2007 and 2012 Census of Agriculture statistics on fruit tree crop acreage.

Since the USDA's Census of Agriculture is undertaken every 5 years within the years ending in numbers 2 and 7, high-resolution imagery acquisition and OBIAs can supplement cherry orchard acreage estimates in this fruitful region in years not ending in the numbers 2 or 7. Future research should explore spectral signature and textural analysis techniques that can distinguish between different orchard tree types to improve the classification and acreage estimation of fruit orchards.

On the Census of Agriculture off years, cherry and other fruit growers who want to obtain crop acreage and additional monitoring of the orchard fields could employ a commercial surveying company to fly UAV drones using multispectral near-infrared sensors to obtain relevant high spatial resolution aerial imagery that would better inform them regarding fruit crop yield and/or potential diseases that may adversely affect their orchard trees. Within this book, you will find innovative uses of UAVs and high spatial resolution imagery in many chapters.

FUNDING SOURCES

Generous funding for the IKONOS imagery came from the Cherry Industry Administrative Board (CIAB). Field research funds came from Pesticide Alternatives Lab, Michigan State University, and the Center for Scholarly and Creative Excellence, Grand Valley State University.

ACKNOWLEDGMENTS

Research support and assistance was provided by the Geography and Sustainable Planning Department, Grand Valley State University, Dr. Wanxiao Sun, Dr. Elena Lioubimtseva, and Ms. Katherine Smigelski. Also, many thanks to Michigan State University's Dr. Mark Whalon, former director of the Pesticide Alternatives Lab (Pest Alt Lab), for his wisdom, agricultural expertise, and encouragement, and finally to the Pest Alt Lab research assistants, Brandon Betz and Tyler Ash, for their assistance with field research.

REFERENCES

Agricultural Marketing Service. 2012. Marketing Order 930: Cherries. U.S. Department of Agriculture. https://www.ams.usda.gov/rules-regulations/moa/930-tart-cherries (accessed April 10, 2017).

Census of Agriculture. 2007. Michigan Counties, Table 32: Fruits and Nuts. U.S. Department of Agriculture. https://www.agcensus.usda.gov/Publications/2007/Full_Report/Volume_1,_Chapter_2_County_Level/Michigan/st26_2_032_032.pdf (accessed July 17, 2017).

Census of Agriculture. 2012a. Michigan Counties, Table 25: Field Crops. U.S. Department of Agriculture. https://www.agcensus.usda.gov/Publications/2012/Full_Report/Volume_1,_Chapter_2_County_Level/Michigan/st26_2_025_025.pdf (accessed July 17, 2017).

Census of Agriculture. 2012b. Michigan Counties, Table 26: Field Seeds, Hay, Forage, and Silage. U.S. Department of Agriculture. https://www.agcensus.usda.gov/Publications/2012/Full_Report/Volume_1,_Chapter_2_County_Level/Michigan/st26_2_026_026.pdf (accessed July 17, 2017).

Census of Agriculture. 2012c. Michigan Counties, Table 31: Fruits and Nuts. U.S. Department of Agriculture. https://www.agcensus.usda.gov/Publications/2012/Full_Report/Volume_1,_Chapter_2_County_Level/Michigan/st26_2_031_031.pdf (accessed July 17, 2017).

Cherry Industry. 2011. *Weekly Raw Product Harvest Report Week 9*. http://www.cherryboard.org/Week92011.pdf (accessed July 17, 2017).

Congalton, R. G. 1991. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment* 37 (1): 35–46.

Coops, N. C., M. Johnson, M. A. Wulder, and J. C. White. 2006. Assessment of QuickBird high spatial resolution imagery to detect red attack damage due to mountain pine beetle infestation. *Remote Sensing of Environment* 103: 67–80.

Imanishi, J., A. Nakayama, Y. Suzuki et al. 2010. Nondestructive determination of leaf chlorophyll content in two flowering cherries using reflectance and absorptance spectra. *Landscape and Ecological Engineering* 6 (2): 1860–1871.

Lillesand, T. M., and R. W. Kiefer. 2008. *Remote sensing and image interpretation*. 6th edition. New York: John Wiley & Sons, pp. 392–481.

Lizarazo, I. 2014. Accuracy assessment of object-based image classification: another STEP. *International Journal of Remote Sensing* 35 (16): 6135–6156. http://dx.doi.org/10.1080/01431161.2014.943328

McCoy, R. M. 2005. *Field methods in remote sensing*. New York: The Guilford Press.

Michigan Department of Agriculture and Rural Development (MDARD). 2015. Facts about Michigan Agriculture. http://michigan.gov/mdard/0,4610,7-125-1572-7775--,00.html (accessed June 23, 2015).

Michigan Department of Agriculture and Rural Development (MDARD). 2017. Facts about Michigan Agriculture. http://michigan.gov/mdard/0,4610,7-125-1572-7775--,00.html (accessed July 17, 2017).

Myint, S. W., P. Gober, A. Brazel, S. Grossman-Clarke, and Q. Weng. 2011. Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. *Remote Sensing of Environment* 115: 1145–1161.

National Agricultural Statistics Survey (NASS). 2017. *State of Michigan Agricultural Overview, 2016*. https://www.nass.usda.gov/Quick_Stats/Ag_Overview/stateOverview.php?state=MICHIGAN (accessed July 17, 2017).

Panda, S. S., G. Hoogenboom, and J. Paz. 2016. Blueberry orchard delineation with high-resolution imagery and self-organizing map neural image classification. *Agricultural Research & Technology: Open Access Journal* 3 (1): 555602. doi:10.19080/ARTOAJ.2016.03.555602.

Pasher, J., and D. J. King. 2009. Mapping dead wood distribution in a temperate hardwood forest using high resolution airborne imagery. *Forest Ecology and Management* 258: 1536–1548.

Pollack, S., and A. Perez. 2009. *Fruit and Tree Nuts Situation and Outlook Yearbook 2008*, FTS-2008. Market and Trade Economics Division, Economic Research Service, U.S. Department of Agriculture, Washington, DC, pp. 48–49, 155.

Santesteban, L. G., C. Miranda, and J. B. Royo. 2007. Establishment of a model to estimate crop load on sweet cherry (*Prunus avium* L.) before bloom, *ISHS Acta Horticulturae 803: VIII International Symposium on Modelling in Fruit Research and Orchard*, pp. 33–39.

Sasaki, T., J. Imanishi, K. Ioki, Y. Morimoto, and K. Kitada. 2012. Object-based classification of land cover and tree species by integrating airborne LiDAR and high spatial resolution imagery data. *Landscape and Ecological Engineering* 8: 157–171. doi:10.1007/s11355-011-0158-z.

Sun, W., M. Kelly, and P. Gong. 2005. Separation of dead tree crowns from the oak woodland forest mosaic by integrating spatial information. *Geocarto International* 20 (3): 15–20.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Index

H

I

K

T

U

V

W