



**UCGE Reports  
Number 20230**

Department of Geomatics Engineering

**Wetland Mapping through Semivariogram Guided  
Fuzzy Segmentation of Multispectral Satellite Imagery**

(URL: <http://www.geomatics.ucalgary.ca/links/GradTheses.html>)

by

**Wen-Ya Chiu**

**September 2005**



UNIVERSITY OF CALGARY

Wetland Mapping through Semivariogram Guided Fuzzy Segmentation  
of Multispectral Satellite Imagery

by

Wen-Ya Chiu

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES  
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE  
DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF GEOMATICS ENGINEERING

CALGARY, ALBERTA

SEPTEMBER 2005

© Wen-Ya Chiu 2005

## ABSTRACT

To protect wetlands from loss, managers need tools to understand the status and trends of wetland resources. Remote sensing techniques provide a cost-effective way for wetland mapping and inventory establishment. However, a robust classification algorithm is the key to generate a reliable map from remotely sensed imagery. To identify wetlands from multispectral imagery, classifiers should take the natural phenomenon, i.e. spatial and spectral vagueness, into account. The Fuzzy C-Means (FCM) clustering algorithm is better suited for dealing with the imprecise data than traditional “hard” classifiers, but it completely ignores the spatial variability inherent in an image. In this thesis, the Semivariogram Guided Fuzzy C-Means (SGFCM) classifier, a modification of the FCM algorithm with spatial variances involved, has been developed for wetland mapping.

Two major tasks are included: replacing the Euclidean distance by the Mahalanobis distance and incorporating the semivariogram texture as spatial guidance in the fuzzy clustering algorithm. Two Landsat 7 ETM+ subscenes are used to examine the effectiveness of the developed SGFCM algorithm. The SGFCM classifier shows an improvement by increasing the overall accuracy from 70 percent to 93 percent and decreasing the commission error by 20 to 40 percent compared to the standard FCM classifier. The SGFCM has an ability to highlight ambiguous pixels that normally lead to the classification uncertainty.

**Keywords:** wetland mapping, Fuzzy C-Means, semivariograms, spatial vagueness

## **ACKNOWLEDGEMENTS**

My sincere gratitude is given to all those who gave me the possibility to complete this thesis. First, I am deeply indebted to my supervisor, Dr. Isabelle Couloigner, for her constant encouragement and inspiration to me in all the time of research and writing of this thesis. Without the supervision and the support from her, this study cannot be done.

My colleagues, Dr. Qiaoping Zhang, Ms. Valarmathy Meenakshisundaram, and Mr. Santosh Phalke, support me in my research work. I want to thank them for all their help, discussions, and valuable hints.

I am very grateful for my dear family, especially my husband Jau-Hsiung, for their patience and love enabled me to complete this thesis.

I furthermore wish to acknowledge Dr. Henry Leung and Dr. Greg McDermid for their stimulating suggestions on the thesis.

## **DEDICATION**

To the Lord who leads me to the “Canaan”

To my little one who is coming into the world

# TABLE OF CONTENTS

<b>Abstract</b> .....	<b>ii</b>
<b>Acknowledgements</b> .....	<b>iii</b>
<b>Dedication</b> .....	<b>iv</b>
<b>Table of Contents</b> .....	<b>v</b>
<b>List of Tables</b> .....	<b>viii</b>
<b>List of Figures and Illustrations</b> .....	<b>ix</b>
<b>List of Symbols</b> .....	<b>xi</b>
<b>List of Abbreviations and Nomenclature</b> .....	<b>xii</b>
<b>CHAPTER 1 Introduction</b> .....	<b>1</b>
1.1 Background Information .....	1
1.1.1 The need for wetland inventory .....	1
1.1.2 Remote sensing for offsite wetland mapping.....	4
1.2 Problem Statement.....	6
1.3 Research Objectives.....	6
1.4 Thesis Outline .....	7
1.5 Summary.....	8
<b>CHAPTER 2 Literature Review</b> .....	<b>9</b>
2.1 Image Classification.....	9
2.1.1 Supervised parametric classification.....	10
2.1.2 Unsupervised parametric classification .....	11
2.1.3 Partially unsupervised classification.....	11
2.2 Classification Uncertainties .....	12
2.2.1 Spatial vagueness of natural objects .....	12
2.2.2 Spectrally mixed pixels.....	13
2.3 Fuzzy Logic for Imprecise Nature .....	14
2.4 Image Texture .....	16
2.4.1 Statistical texture analysis.....	16
2.4.2 Semivariogram.....	18
2.4.3 Semivariogram texture classification.....	20

2.5 Accuracy Assessment .....	21
2.5.1 Accuracy and error.....	22
2.6 Summary.....	23
<b>CHAPTER 3 Fuzzy Clustering Theory and Developed Modifications.....</b>	<b>25</b>
3.1 Concept of Clustering Analysis .....	25
3.1.1 Multidimensional feature space .....	26
3.1.2 Similarity measures.....	27
3.2 Theory of Fuzzy Clustering .....	29
3.3 Overview of Fuzzy C-Means.....	30
3.3.1 Fuzzy C-Means clustering algorithm (FCM).....	31
3.4 Modification of Fuzzy C-Means.....	33
3.4.1 Semivariogram Guided Fuzzy C-Means clustering algorithm (SGFCM) .....	34
3.5 Defuzzification of Fuzzy Membership Function .....	39
3.5.1 Maximum membership defuzzy principle .....	39
3.5.2 Alpha ( $\alpha$ )-cuts defuzzy rule.....	40
3.6 Measurement of Uncertainty.....	42
3.7 Summary .....	42
<b>CHAPTER 4 Methodology.....</b>	<b>44</b>
4.1 Study Area Description.....	44
4.2 Data.....	46
4.2.1 Satellite imagery .....	46
4.2.2 Reference data.....	46
4.3 Image Pre-processing.....	48
4.3.1 Radiance conversion .....	49
4.3.2 Reflectance conversion .....	49
4.4 Tasseled Cap Transformation .....	50
4.4.1 Overview.....	50
4.4.2 At-satellite reflectance-based tasseled cap transformation.....	52
4.5 Water Area Extraction .....	54
4.6 Training Site Selection.....	56

4.7 Implementation of Classification Algorithms.....	58
4.8 Summary .....	59
<b>CHAPTER 5 Experimental Results and Discussions .....</b>	<b>61</b>
5.1 Examination of Tasseled Cap Features.....	61
5.1.1 The plane of vegetation.....	64
5.1.2 The plane of transition zone.....	64
5.1.3 The plane of soil.....	65
5.2 Preliminary Analysis of Semivariogram Behaviors .....	66
5.3 Fuzzy Classification of Land Cover Types.....	69
5.3.1 Analysis of the fuzzy membership values .....	69
5.3.2 Visual evaluation of the fuzzy classification .....	73
5.3.3 Analysis of the fuzzy class dispersion .....	77
5.3.4 Analysis of the confusion index and mixed pixels .....	85
5.4 Accuracy assessment .....	91
5.5 Summary .....	94
<b>CHAPTER 6 Conclusions and Future Scope .....</b>	<b>96</b>
6.1 Conclusions.....	96
6.2 Future Scope .....	99
<b>References .....</b>	<b>100</b>
<b>Appendix A: Evaluation of the data-driven window size to incorporate texture features into wetland mapping.....</b>	<b>107</b>

## LIST OF TABLES

Table 2-1. Example of a confusion matrix.....	22
Table 4-1. Ancillary data of the LANDSAT 7 ETM+ scene.....	50
Table 4-2. Tasseled cap coefficients for Landsat 7 ETM+ at-satellite reflectance.....	53
Table 5-1. Confusion matrix for wetland mapping of the test area A .....	92
Table 5-2. Confusion matrix for wetland mapping of the test area B.....	93

## LIST OF FIGURES AND ILLUSTRATIONS

Figure 2-1. Illustration of an omnidirectional semivariogram.....	20
Figure 3-1. Illustration of feature space.....	27
Figure 3-2. Illustration of the membership function.....	30
Figure 3-3. Illustration of maximum membership defuzzification.....	40
Figure 3-4. Illustration of two different $\alpha$ -cut sets for classification .....	41
Figure 4-1. Area of interest.....	45
Figure 4-2. Color composite images (TM 432) showing the test areas.....	47
Figure 4-3. The framework of the image data processing.....	48
Figure 4-4. Dispersion of the six-band Thematic Mapper data.....	51
Figure 4-5. Approximate locations of important scene classes in TM tasseled cap feature space.....	53
Figure 4-6. Typical spectral reflectance of common earth features.....	55
Figure 4-7. Training sites for deriving semivariograms.....	58
Figure 5-1. Planar dispersion of the test area A in TM tasseled cap feature space.....	62
Figure 5-2. Planar dispersion of the test area B in TM tasseled cap feature space.....	63
Figure 5-3. Omnidirectional semivariograms derived from the TM tasseled cap features .....	67
Figure 5-4. Images of fuzzy membership values (FMVs) of the wetland class .....	70
Figure 5-5. Digital Elevation Model (DEM) of test areas.....	72
Figure 5-6. Reference data and two classification maps for test area A.....	74
Figure 5-7. Reference data and two classification maps for test area B.....	76
Figure 5-8. Class dispersion in the TM tasseled cap feature space for test area A using the SGFCM classifier.....	78
Figure 5-9. Class dispersion in the TM tasseled cap feature space for test area A using the FCM classifier.....	79
Figure 5-10. Class dispersion in the TM tasseled cap feature space for test area A using the FCM classifier with alpha-cuts rule defuzzification.....	80
Figure 5-11. Class dispersion in the TM tasseled cap feature space for test area B using the SGFCM classifier.....	82

Figure 5-12. Class dispersion in the TM tasseled cap feature space for test area B using the FCM classifier.....	83
Figure 5-13. Class dispersion in the TM tasseled cap feature space for test area B using the FCM classifier with alpha-cuts defuzzification. ....	84
Figure 5-14. Confusion index (CI) maps displayed in (a) gray level and (b) mixel level for test area A.....	86
Figure 5-15. Class level uncertainty maps of test area A .....	87
Figure 5-16. Confusion index (CI) maps displayed in (a) gray level and (b) mixel level for test area B.....	89
Figure 5-17. Class level uncertainty maps of test area B.....	90

## LIST OF SYMBOLS

<b>Symbol</b>	<b>Definition</b>
$A_k$	Norm matrix
$C$	Cluster numbers
$D, d_{ik}, d'_{ik}$	Distance measure
$J, J_m$	Objective function
$M_k$	Fuzzy covariance matrix
$m$	Fuzzy index
$N$	Total pixel numbers of an image
$n_1, n_2$	Image size in pixel and line
$p$	Dimensions of feature space
$U, U^*$	Partition matrix
$\mu_{ik}, \mu_{ik}^*$	Fuzzy membership
$V$	Vector of cluster center
$v_k, v_k^*$	Cluster center
$\gamma_i(h)$	Semivariance
$h$	Lag distance
$\alpha_{low}, \alpha_{high}, \varepsilon$	Thresholds
$\lambda$	Lagrange multiplier
$\rho_k$	Constant
$s$	Window size
$w_{ik}$	Semivariogram weighting factor
$i, j, \text{ and } k, g$	Variables of pixels, dimensions, and clusters
$T$	Transpose
$X$	Data matrix
$x_i$	Feature point

## LIST OF ABBREVIATIONS AND NOMENCLATURE

<b>Abbreviation</b>	<b>Full term</b>
LANDSAT	Name of the satellite
ETM+	Enhanced Thematic Mapper plus
FCM	Fuzzy C-Means
SGFCM	Semivariogram Guided Fuzzy C-Means
TP	True positive
FP	False positive
TN	True negative
FN	False negative
BOREAS	Boreal Ecosystem-Atmosphere Study

## CHAPTER 1

# INTRODUCTION

*“The dreams of scientists and engineers are not easily confined. Numerous people thirst for the opportunity to conduct a continental and/or global inventory of wetlands. Such an inventory would be of great interest to the scientific community and the public.”*

*John G. Lyon and Jack McCarthy*

*Editors of “Wetland and Environmental Applications of GIS”, 1995*

### **1.1 Background Information**

#### **1.1.1 The need for wetland inventory**

Wetlands are a major feature of the landscape in almost all parts of the world. However, the extent of wetlands is decreasing in many countries due to human-generated stresses such as urban development, resource extraction, toxic chemicals pollution, and non-native species introduction (Mitsch and Gosselink, 1993). The global climate changes that result in the rising of sea water level have also some impacts on wetland resources (Day, 2000). To conserve and protect the remained wetlands, thoughtful planning and regular monitoring are required.

Why are wetlands so valuable? Wetlands have been considered transitional seres that have extensive boundaries with both terrestrial and aquatic ecosystems. Because wetlands are neither aquatic nor terrestrial, such transitional ecotones reflect a complex and

---

dynamic interaction of physical and biotic forces acting at the landscape scale. They should be considered a distinctive class. In the 1980's, as wetlands were recognized for their societal values that motivate their protection, scientists were asked to give the definition of wetlands to determine the wetland boundaries for regulatory purposes. The functions of wetlands that directly benefit the society include storing floodwater, protecting shorelines, improving water quality, and recharging groundwater aquifers (Daily, 1997). Associations between functions and values may also be indirect. For example, wetlands provide critical habitats for fish and wildlife in sustaining a rich biodiversity. People can enjoy the recreational opportunities and aesthetics based on healthy wetland ecosystems.

To protect wetland resources, managers should conduct a 'status and trends' assessment of wetlands involving the analysis of the past, current and future landscape changes. The assessment needs ancillary information and data, such as maps or field surveys, for decision making. Inventory and monitoring are the key elements for effective implementation of wetland management programs, policies and sustainable development indicators (Milton and Hélie, 2003). The common objective of compiling wetland inventories is to determine wetland locations, to delineate the extent of various types of wetlands in a region, and to identify important habitats for wildlife, economic interests and other functions. An inventory can be made either for a small watershed or an entire nation. Whatever the size of the area to be surveyed, the inventory must provide information on wetlands to meet the needs of specific users.

In Canada, wetlands occupy about 18 percent of the total land area (Natural Resource Canada, 1986). Natural Resource Canada has published a map of the wetland regions in Canada in 1986. The exact amount varies somewhat according to the techniques and methods of survey. Seven major regions are recognized: Arctic, Sub-Arctic, Boreal, Prairie, Temperate, Mountain, and Oceanic. These regions are further divided into 20 sub-regions according to their geographic location. This general inventory does not identify specific wetlands; rather it delineates wetland regions based on similar ecological

characteristics. Another category instead of the general inventory is wetland type inventory that focuses on the specific wetland types such as peatlands or forested wetlands.

Regional wetland type inventories and national data sets do exist. However, no nationwide wetland type inventories and reports have been undertaken to track or report on the status and trends in wetland resources. The barriers of the different classifications and mapping standards cause the result of the difficulties in developing a nationwide wetland inventory, because each political jurisdiction and institution may have a different availability of technology and digital wetland inventory information (Milton *et al.*, 2003). Although the previous attempts at a national inventory have been frustrated due to these difficulties, the development of a national inventory is still called for in understanding and managing wetland resources. Davidson *et al.* (1999) concluded, “This lack of a national inventory with a standard classification system makes it virtually impossible to monitor wetlands at the ecosystem level, except in those areas that have developed their own inventories, but then only if the same methodologies are adhered to over time.”

Wetland identification and classification need a standard approach either a classification technique or an inventory system as a basis for wetland monitoring in the future. Developing a widely adopted wetland classification scheme such as the Cowardin system (Cowardin *et al.*, 1979) used for the National Wetland Inventory (NWI) in the United States or the Canadian Wetland Classification System (CWCS) goes beyond the scope of this thesis study. Whatever the classification system used for wetland types identification, extracting wetlands from landscape is the first step in building up a nationwide wetland resource database. The needs call for wetland mapping. Superior land cover classification techniques and data management systems can facilitate synoptic wetland inventories. This thesis study narrows the scope to the image processing techniques for wetland mapping by using remotely sensed imagery.

### **1.1.2 Remote sensing for offsite wetland mapping**

Wetland mapping is a prerequisite for wetland inventory. Aerial photography, satellite imaging, or maps can be utilized as an alternative to collect field data in order to map wetlands. Offsite methods are recommended for use in areas where an inspection is not possible due to time constraints or other reasons. Sometimes offsite wetland determinations may be the only source of information for environmental planning decisions. However, delineation of wetlands by offsite methods is subject to some errors; it thus should be used only when their inherent limitations are recognized. The overall accuracy of offsite wetland determinations is a function of the quality of the sources used and the ability of data interpretation (Tammi, 1994).

A mapping approach that relies upon remote-sensing technology holds the promise of greater efficiency, lower costs, and ease of repeatability. Because the remotely sensed images provide a synoptic view of wetlands and their surrounding terrain, they facilitate rapid boundary determination. The use of remote sensing platforms may include aircraft at various altitudes, and satellites. The choice of which platforms to use depends on the spatial resolution required, the area to be covered, and the cost of the data.

Aerial photography has been used to map wetlands for at least three decades (National Research Council, 1995). Before the second-generation satellite imagery from the “Thematic Mapper” that provided multispectral data at 15m or 30m resolution at a reasonable cost was available, wetland scientists preferred to use aerial photographs for wetland mapping. As aerial photography provided the basemap and framework for most other offsite sources, wetland scientists relied primarily on aerial photography. For example, the National Wetland Inventory (NWI) project, begun in 1975 by the United States Fish and Wildlife Service for conserving and managing wetlands, chose high altitude aerial photography over satellite imagery because of the desired level of details (Tiner, 1990).

While satellite imagery has advantages for the continued monitoring of wetland resources, aerial photographic coverage and availability significantly limit its use for offsite wetland identification. Because the coverage of wetlands is unpredictably varying from complete chronological coverage over several years to no coverage at all, the costs of purchasing aerial photographic imagery can be quite high with some firms charging access fees for database reviews (Tammi, 1994). Satellite imaging offers excellent opportunities to obtain cost-effective data at landscape-scale or watershed-scale to meet the demand for monitoring large wetland extents repeatedly.

Because reflectances of ground objects make associated responses to the status of physical characteristics, the changes of wetland extents across the landscape can be detected from multispectral imagery. Many satellite sensors, such as Landsat, SPOT, IKONOS, and radar systems, have been used to study wetlands. For example, Gluck *et al.* (1996) used Landsat TM imagery to determine the technology required for wetland classification and mapping. Parmuchi *et al.* (2002) evaluated the suitability of using multi-temporal RADARSAT-1 data with a decision-tree classifier to map wetlands. To take advantages of both the optical sensors, which contain information about the reflectivity of objects, and the SAR sensors, which are sensitive to moisture differences, Töyrä *et al.* (2001) investigated the use of multispectral SPOT and Radarsat imagery for mapping the extent of standing water in the Peace-Athabasca Delta wetland. High-resolution satellite imageries have also been used for wetland mapping. Dechka *et al.* (2002) investigated the use of IKONOS imagery that provides 4m-resolution multispectral and 1m-resolution panchromatic image data to assist in the assessment of wetland habitat classification in the Canadian Prairies region. Readers can refer to Ozesmi and Bauer (2002) for the summary of reviews on satellite remote sensing of wetlands and the classification techniques for wetland identification in the past two decades.

## 1.2 Problem Statement

Interpretation of satellite images is difficult, especially in areas where changes of vegetation, soil, or hydrology are indistinct or variable through time. Such areas are generally difficult to delineate by field methods as well (National Research Council, 1995). The boundary between wetland and upland can be identified scientifically as a transition zone that incorporates a hydrologic gradient as well as gradients in soil type and in community composition of plants. However, image classification techniques for wetland identification cannot ensure that any resultant definition will be precise in its ability to distinguish wetlands from all other kinds of ecosystems, or in its ability to specify the exact boundary of a wetland. Especially when using the traditional hard classification techniques, these areas may introduce classification errors and reduce the mapping accuracy.

The classification uncertainties result from two facts. First, the spatial vagueness is inherited in wetland ecosystems and the wetland definition itself (Carter *et al.*, 1994). The characteristics of hydrology, soil and vegetation in an ecotone from wetlands to uplands vary transitionally rather than abruptly. Second, a spectral mixture of sub-pixel elements leads to an uncertainty of the classification. Traditional image classification techniques assume pixels to be pure and thus carry out the classification results based on binary logic, in which a pattern is either a full member or not of the class. Such a hard classification algorithm is not the best way to deal with data that are mixed and imprecise in nature. The spatial extent of each object can be defined ambiguously and may contain unidentified areas not belonging to the object (Burrough, 1996; Cheng and Molenaar, 1999).

## 1.3 Research Objectives

It is hypothesized that because a fuzzy approach can measure the uncertainty between the class boundaries to identify the degree of mixed classes by membership values and because image texture contains complementary information for image classification,

---

wetland mapping will be more efficient when the applied fuzzy classifier takes image textures into account. The research goal of the study is to develop a more robust fuzzy clustering algorithm for land cover classification. A fuzzy classification can be treated more realistically when the data point is difficult to be assigned to one single class. In addition, the fuzzy classifier should consider information not only from the data set itself but also from the variances between the data points. The variances represent the spatial variability inherent in the data set and provide additional information for the classification.

The major objective of this thesis is to develop a fuzzy classifier incorporating image texture information for wetland mapping and to investigate its effectiveness in dealing with the data that are mixed and imprecise in nature. In order to fulfill the major objective, the following related objectives will be accomplished:

- (1) To modify the Fuzzy C-Means clustering algorithm as a partial supervised classifier.
- (2) To highlight the spectrally vague areas that are sensitive to disturbances by replacing the maximum defuzzifier with a threshold mode approach.
- (3) To examine the applicability and effectiveness of the modified FCM algorithm with two test areas.

## **1.4 Thesis Outline**

This thesis is structured in six chapters. Chapter 1 gave an introduction of the need for wetland mapping. The motivation for developing a texture involved fuzzy classifier for wetland mapping and the objectives of this thesis research are also introduced.

Chapter 2 reviews the literature related to this research along with spatial vagueness that leads to classification uncertainty, image textures, geostatistics in remote sensing, and accuracy assessment.

Chapter 3 portrays the methodology used and developed for this research. The chapter describes the fuzzy clustering algorithm - Fuzzy C-Means (FCM) - and its disharmony issues. The modifications of the FCM clustering algorithm - Semivariogram Guided Fuzzy C-Means (SGFCM) - are demonstrated with derivations.

Chapter 4 describes the study areas and the data processing for this thesis. The framework of the study is given and details of each procedure are addressed in this chapter.

Chapter 5 discusses the experimental results for test areas. The discussions begin with the preliminary examination of the data set dispersion and the semivariogram behaviors of the land cover classes of interest. The classifications from the two fuzzy classifiers are compared qualitatively (visual assessment) and quantitatively (accuracy assessment).

Finally, Chapter 6 summarizes the findings of this thesis study, and draws the important conclusions and the future scope of the research.

## **1.5 Summary**

This chapter has discussed the background of wetland mapping. Because managers need the tools for wetland resources management, wetland scientists are developing a remote-sensed wetland inventory using satellite imagery. However, identifying wetlands from multispectral imagery is not a trivial task, because such a complex ecosystem is characterized by a mixing of the signatures of water, soil, and vegetation, which vary gradually and continuously between aquatic and terrestrial systems. The main objective of this research is to develop a fuzzy approach incorporating image texture information for improving the accuracy of wetland mapping.

## CHAPTER 2

# LITERATURE REVIEW

*“You must not know too much or be too precise or scientific about birds and trees and flowers and watercraft; a certain free-margin, and even vagueness - ignorance, credulity - helps your enjoyment of these things.”*

**Henry David Thoreau, 1817-1862**  
*American author, poet and philosopher*

*“How many problems can we say that the information content is known absolutely ... There is uncertainty that arises because of complexity; there is uncertainty that arises from ignorance, from chance, from various classes of randomness, from imprecision, from inability to perform adequate measurements, from lack of knowledge, or from vagueness, like the fuzziness inherent in our natural language.”*

**Timothy J. Ross**  
*Professor, Civil Engineering, University of New Mexico, 1995*

### 2.1 Image Classification

The goal of classification in general is to select the most appropriate category for an unknown object. In principle, classification of remotely sensed image should be

---

straightforward. To achieve an acceptable mapping accuracy, it should be concerned to choose the appropriate analytical tools first and then apply them. The classification techniques may be categorized either on the basis of training processes (supervised and unsupervised) or on the basis of theoretical models (parametric and non-parametric). The difference between supervised and unsupervised classification depends on whether the classifier has been trained by user-based training data. Making a certain assumption of the probability distribution of data can differentiate the parametric and non-parametric classification. The parametric classifier provides a way to estimate the error of the classification process itself since the decision boundaries between classes are established in the distribution assumption. With the non-parametric classifiers, classes are to be found with no prior knowledge of the probability density function of each class. To attain optimal results, more often a hybrid strategy of combining the two will be necessary (Richards, 1993). In a hybrid classification supervised and unsupervised classifier may be used in different stages of the classification procedure.

### **2.1.1 Supervised parametric classification**

Supervised classification procedures are the essential analytical tools used for the extraction of quantitative information from remotely sensed images. An important assumption in supervised parametric classifications adopted in remote sensing is that each spectral class can be described by a probability distribution in the multispectral space (Richards, 1993). The most commonly used distribution is the Normal or Gaussian distribution. For example, Maximum Likelihood Classification is the most commonly used classifier among the other parametric classifiers despite its longer computational time and inherent assumptions. A supervised classification needs existing reference data to establish user-selected training sets for land cover classes. Reference data can be obtained from site visits, maps, air photographs or even photo-interpretation of a color composite product formed from the images. When pixels are selected as training site, these training pixels are used to derive various statistics (e.g. mean and standard deviation) for each land cover class to represent the prototype of the associated class. According to the statistics generated from the training data, the parameters of the

particular classification algorithm are estimated and the classifier is taught to recognize the spectral characteristics of each class.

### **2.1.2 Unsupervised parametric classification**

Unsupervised classification is a means by which pixels in an image are assigned to spectral classes without the user having foreknowledge of the existence or names of those classes. Clustering analysis is usually the common method to perform an unsupervised image classification. The user determines the number and the location of the spectral classes into which the data falls, as well as the spectral class of each pixel according to statistical similarity. After the completion of the clustering, it is the task of the analyst to identify those classes by associating a sample of pixels in each class with available reference data, which could be maps or information from ground visits (Richards, 1993). Often unsupervised classification is used on its own, particularly when the training data for supervised classification cannot be obtained or is too expensive to acquire.

### **2.1.3 Partially unsupervised classification**

Contrasting the supervised and unsupervised learning processes, intermediate modes of data structure searching are called for the real-world applications. Especially when the domain knowledge about the classification problem is available, a prudent use of the foreknowledge could substantially enhance the classification efficiency (Pedrycz and Waletzky, 1997). Semi-supervised/unsupervised classifiers, therefore, have been proposed by some researchers to deal with the fact that gathering training samples or labelling training samples is very expensive in terms of time and manpower. For examples, Jeon and Landgrebe (1999) in their study had prior knowledge of only one particular class of interest. They thus developed a partially supervised clustering method as a special case of unsupervised clustering with one particular cluster initially known for classification. Bruzzone *et al.* (2002) proposed a multiple-classifier architecture that first trained the classifier in a supervised way, and then retrained the classifier in an unsupervised way with the obtained parameters to updated land cover maps. Their hybrid classification system was capable of producing accurate land cover maps even for images

without available ground-truth information. The combination of supervised and unsupervised classifiers can be a useful strategy in classification, especially when training samples of only a particular class of interest is available.

## **2.2 Classification Uncertainties**

Remote sensing offers the potential to study the ecotone dynamics over a large space and long time periods. Here, the ecotone in the general sense represents the merged ecological classes that neighbour with each other. The potential of remote sensing for land cover mapping, however, may be restricted if an inappropriate classification method is chosen for the specific mapping purpose. Especially the inherent uncertainty from various sources propagates during the image processing. Recognizing the limitations of classification techniques and understanding the sources of mapping errors are important to both producers and users. In land cover mapping, either natural phenomenon or classification techniques would lead to classification uncertainty. First, no exact boundary exists between spatially vague objects. Second, spectrally mixed pixels enclosing more than one object are contradictory to the pure pixel assumption of image generation. In order to have reliable mapping results, these uncertainty sources need to be considered in the choice of an appropriate classifier for wetland mapping.

### **2.2.1 Spatial vagueness of natural objects**

Many geographical applications describe the spatial extensions of natural geographic objects by sharp regions that have a unique boundary. However, the determinate and crisp concepts are not adequate logic for mapping the spatial phenomena of the real world. The reason is that the intrinsically spatial indeterminacy and spatial vagueness are inherent in many geometric and geographic data (Schneider, 2002). Almost every natural object has a gradual boundary; two vague regions can share a common gradual boundary like the transition of wetland and upland. For example, the presence of hydrophyte vegetation, hydric soil, and flooding is commonly used as conditions to classify wetlands

and to determine their boundaries. Typically, these characteristics of wetlands are distributed gradually and continuously rather than abruptly over space in ecotones.

Carter *et al.* (1994) analyzed the changes of hydrogeology, soil and vegetation along the moisture/elevation gradient to select tentative boundaries based on field survey. Their investigation showed that these three parameters demonstrated a gradual change from wetlands to uplands and that the locations of boundaries based upon the individual parameters could differ in broad gradual transition zones. They thus concluded that the basic problems appearing to the placement of exact wetland boundaries were the vagueness and circularity of the wetland definition itself. This conclusion is coincident with the statement of Kulik (2003), "At least two perspectives of spatial vagueness to be considered: the objects themselves are vague, or the concepts or representation of objects are vague." The ambiguity of nature land cover composition in the transitional area leads to confusion and thus to classification errors. Indeed many errors in image classification may be concentrated spatially around boundaries of both continuous and discrete classes (Zhang and Foody 1998).

### **2.2.2 Spectrally mixed pixels**

As the vagueness of natural objects introduces confusion and classification errors, this natural phenomenon is also reflected spectrally on a remotely sensed image. When collecting the spectral information from the ground objects, the remote sensor records the comprehensive reflectances of an area based on a pixel unit. A problem in remote sensing is that the land cover may vary more spatially than the sampling interval between pixels in the imagery, particularly with imagery of moderate and coarse spatial resolution. This implies that many of the pixels may represent a mixture of land cover classes and may be referred to as mixed pixels.

Mixed pixels can represent a significant problem in image analysis and they are not sensible to assign a single pixel to a single class. Conventional image classification techniques, however, assume that the study area is composed of a number of unique,

---

internally homogeneous classes (Fisher and Pathirana, 1990). They also presume that the classification analysis based on reflectance data can identify these unique classes. Unfortunately, these untenable assumptions cannot give a realistic description of data because they violate the real facts. Furthermore, the inherent errors emerge from the classification uncertainty (Cheng et al., 2001). The mixed pixels degrade the image classification accuracy because a mixed pixel contains more than one class, and because a mixed pixel displays a composite spectral response that is similar to each of its component classes. The pixel sometimes may not be allocated to one of the component classes, but in the conventional classification a pixel can be allocated to only one class.

A number of approaches have been developed and tested for solving pixel unmixing to reduce the classification errors. Foody and Cox (1994) in their study estimated the sub-pixel land cover composition from Landsat MSS imagery. Their results showed that pixel unmixing approaches were more accurate than traditional image classifiers in the assessment of class coverage over a region. This means that the traditional classification method such as maximum likelihood classification is often incapable of performing satisfactorily in the presence of mixed pixels.

### **2.3 Fuzzy Logic for Imprecise Nature**

Recognizing that the classical theoretical notion of classification techniques as a discrete view of space, researchers are trying to model and reason about areas with “indeterminate boundaries” (Burrough and Frank, 1996). With this perception both the spatial continuity and the fuzziness of spatial data can be involved in the classification techniques. When the two or more classes occupy a single pixel especially in a coarse spatial resolution remotely sensed imagery, the mixed pixel would be appropriate to conceive the different landscapes as a set of fuzzy classes. Burrough (1996) considered fuzzy sets as an appropriate method to deal with ambiguity, vagueness, and ambivalence in mathematical or conceptual models of empirical phenomena.

---

Originally, fuzzy logic was developed by Zadeh (1965) to provide a general representation of vagueness based on different degrees of membership values. In contrast to a traditional “hard” classification constructed based on binary logic, a classification method using fuzzy logic is called a “soft” classifier. Fuzzy approaches have been successfully utilized in pattern recognition (Cuevas *et al.*, 2004; Leung *et al.*, 2004) or classification in different study fields, such as medical image analysis (Pham, 2001; Ahmed *et al.*, 2002) and climate study (McBratney and Moore, 1992). Fuzzy classification is now widely accepted as a ubiquitous solution to the mixed and imprecise problems in nature because of its ability to describe the classification uncertainty. De Bruin and Stein (1998) applied the fuzzy set approach to soil data derived from a digital elevation model (DEM) to represent transition zones in the soil-landscape. They concluded that the fuzzy approach could reveal the spatial pattern of soil-landscape and improve the conventional soil-landscape modelling. Zhang *et al.* (2004) applied the fuzzy approaches and the statistical method to ecological habitats classification with the field data collected from the mixed-species forest stands. Their study showed the ability of the fuzzy classifiers in improving the classification accuracy and the flexibility in classifying ecological habitats that have a mixture of overstory and understory species.

In the applications of capturing the gradual transition areas from remote sensing imagery, Foody and Boyd (1999) compared the fuzzy classification to the conventional “hard” classification while using Advanced Very High Resolution Radiometer (AVHRR) data to identify the changes in the apparent position of the forest-to-savanna transition. Their study revealed that fuzzy classifications could characterize the migration of the transitional area because richer information emerged in the class memberships. To delineate ‘meaningful’ objects from uncertain classification results, Cheng *et al.* (2001) elucidated three fuzzy object models (*i.e.* fuzzy-fuzzy, fuzzy-crispy and crispy-fuzzy object model) to represent objects with indeterminate boundaries. Recently, fuzzy approaches also illustrate their potential in sub-pixel mapping by interpreting the spectral data in terms of membership grades to a class. For example, Lucas *et al.* (2002) applied the fuzzy approach to airborne image spectrometer (CASI) imagery for coastal dune

habitat mapping. Schmidt and Schoettker (2004) utilized the fuzzy membership function for spectral unmixing to optimize a land cover classification.

## **2.4 Image Texture**

The linkage between landscape structures and ecological functions is based on the spatial arrangement of landscape elements rather than on an identical square area. In landscape ecology, the basic unit of a landscape is a landscape element - a “homogeneous surface area that differs from its surroundings in nature or appearance” (Turner *et al.*, 2001). Using remote sensing images, homogeneity is defined in relation to pixel size. A patch refers to a compact element consisting of a few pixels and is also a homogeneous surface area. Pixels comprising the patch are supposed to have similar digital numbers. Intrinsically, the local spatial variability of image pixels provides a practical means of analyzing the textural properties of objects since the grey level is not randomly distributed within an image.

In digital images, the characteristics of a texture can be sensed via variations in the captured intensities. It is a valuable feature in discriminating among different land cover types. Image texture, which provides a complementary tool to satellite images, has received great attention in land cover mapping (Treitz *et al.*, 2000; Franklin *et al.*, 2000; Arzandeh and Wang, 2002). Classical image classification algorithms based on the “pixel-paradigm”, however, focus on the statistical analysis of individual pixels rather than on the spatial patterns they build up (Blaschke and Strobl, 2001). The potential of the spatial information existing between a pixel and its neighbours is totally ignored. Therefore, image attributes within a landcover type over its neighbourhoods should be characterized to achieve a more reliable and a more accurate mapping result.

### **2.4.1 Statistical texture analysis**

Haralick *et al.* (1973) noted that texture has been extremely refractory to precise definition. So far no one has been able to define digital texture in mathematical terms

although a number of people have been trying to find ways of categorizing textures (Mäenpää, 2003). Traditionally, textural processing algorithms can be divided into two categories: structural and statistical methods. Structural methods consider texture as a repetition of basic primitive patterns. Statistical methods treat texture as statistical properties of intensities and positions of pixels.

Grey-level co-occurrence matrix (GLCM), one of the most widely used statistical texture measures, has been well implemented in commercial software, such as *PCI Geomatica<sup>TM</sup>*. The idea of the algorithm is to consider the relative frequencies for which two neighbouring pixels are separated by a distance on the image. Since the GLCM collects information about pixel pairs instead of single pixels, it is called a second-order statistic. Texture measures, such as homogeneity, contrast, and entropy derived from the co-occurrence matrix can be incorporated with spectral features for classification purposes (Haralick, 1979). The GLCM approach has been utilized in different applications such as crop classification in agriculture, forest species classification, and wetland mapping in nature resources management. By using the GLCM approach for crop classification, Treitz *et al.* (2000) noted that combination of texture features in classification was superior to the original image. Franklin *et al.* (2000) also demonstrated the improvement of the accuracy for forest stands classification while incorporating derived texture measures from airborne multispectral video images into classifier.

Two parameters, such as the combinations of textural features and window size, associated with the GLCM method should be considered directly. Examining the effects of textural feature combination, Arzandeh and Wang (2002) observed that the combination of three or four texture features provided a better performance than only one or two texture features. However, the most appropriate combinations of texture features strongly depend on the surface properties for a particular application; no rules are recommended for the texture measures selection. With an appropriate window size for texture feature derivation, unique texture patterns of a patch can be extracted for land cover type discrimination. Marceau *et al.* (1990) and Peng *et al.* (2003) have examined

the effects of different moving window sizes, from 3×3 to 25×25 pixels, on classification accuracy. They concluded that large window size could capture the spatial patterns of each land cover type better, but might contain more than one land category which could introduce systematic errors. The window should be small enough, however, to keep the variance low and to maximize the potential for class separability.

In general, the window size is a critical factor in remote sensing. Instead of a time intensive trial-and-error method, a promising strategy, the semivariogram, based on the actual image data can be used into the procedure to derive an optimal geographic window to use in remote sensing. The semivariogram is a property used in regionalized variable theory to express the degree of relationship between pixels on image digital numbers. The semivariance at some distance, called range of the regionalized variable, is approximately equal to the variance of the image digital numbers. This range shows the greatest distance over which the value of a pixel on the image is related to the value of another pixel. The range can thus be used as a measure of homogeneity and be used to derive an optimal geographic window.

#### **2.4.2 Semivariogram**

Geostatistics is a set of techniques for the analysis of spatial data. The techniques consider the spectral-response patterns existing between a pixel and its neighbours. They are characterized by their dependence on a model of the spatial covariance function: the semivariogram function. The semivariogram, a data-driven graphical representation of the spatial variability, provides a method to measure the spatial dependency of continuously varying phenomena (Curran, 1988; Franklin *et al.*, 1996). In remote sensing, digital numbers of remotely sensed imagery can be geostatistically interpreted as a regionalized variable, and characterized by both random and spatial correlation aspects. Under the intrinsic hypothesis, both aspects can be jointly studied through the semivariogram function concept.

The semivariogram, developed from the theory of regionalized variables, displays the average change of a property with changing lags. The relation between a pair of pixels that are lag distance  $h$  apart can be given by the average variance of the difference between all such pairs and expressed as follows:

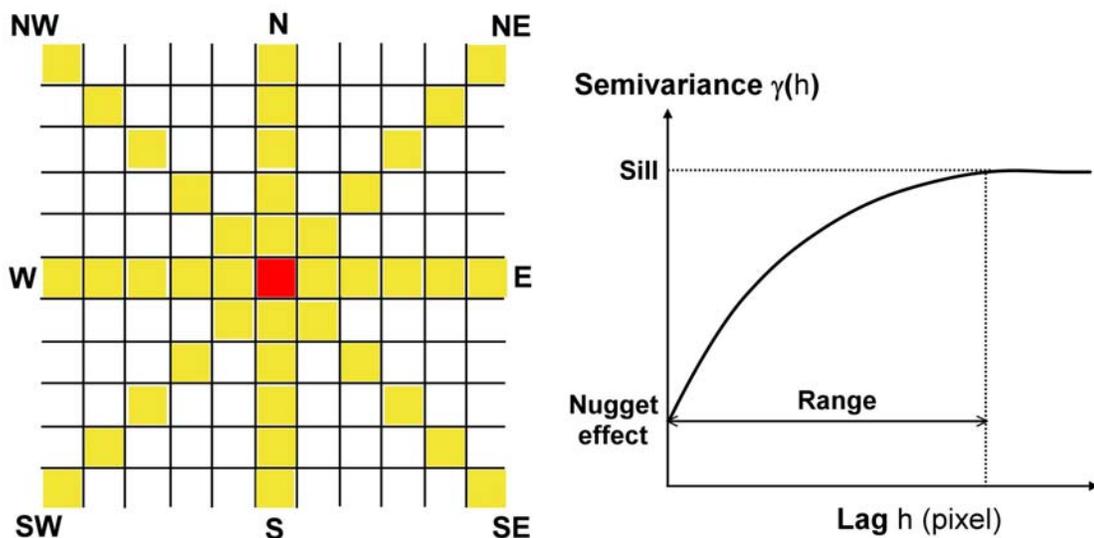
$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x_i) - Z(x_{i+h})]^2 \quad (2-1)$$

In remote sensing,  $\gamma(h)$  estimates the variability of radiance as a function of the spatial separation in lag distance,  $Z(x_i)$  represents the digital number at a pixel location  $x_i$ , and  $N(h)$  means the total number of pixel pairs. Because real spatial phenomena may show directional effects depending on the data properties, a spatial process of a directional semivariogram may reflect anisotropy phenomenon. Directional semivariograms define the variation of two point pairs depending not only on the separation distance, but also on the orientation of the two points. However, omnidirectional semivariograms considers that the separate lag distance  $h$  is a scalar rather than a vector in the space.

The interpretation of a semivariogram usually focuses on the relating parameters: range, sill and nugget effect (Figure 2-1). Fig 2-1 shows an omnidirectional semivariogram computed according to the average variance of eight directions. At a lag distance called the range, the variogram would typically stabilize. Groups of pixels within the range are highly correlated. Thus the range can be used to measure the homogeneity of a patch and further to predict the optimal geographic window sizes (Curran, 1988; Franklin *et al.*, 1996). The value of the variogram at this range is called the sill. The sill can be viewed as a mean value of the digital number variations. At distances greater than the range, any two pairs of values are independent of each other. In general, the value of the variogram would also be smaller at smaller lag distances because values separated at small distances tend to be more similar. However, a plot of the experimental semivariogram may indicate a discontinuity at the origin,  $h = 0$ . Mathematically the variance of the pixel itself is equal to zero. However, due to microscale variation a nugget effect occurs (i.e. the variance at

the origin is not equal to zero). This nugget effect is considered as random noise due to microscale processes and measurement error.

Chiu and Couloigner (2004a) used the omnidirectional semivariograms derived from Landsat image data to find out the optimal window size(s) for deriving texture features. The result showed that the incorporation of texture measurements into multispectral data could improve the classification accuracy. Also, using the predicted range as the window size for deriving texture features for a specific class could provide superior discrimination and correlation results compared to those obtained using randomly selected identical windows.



**Figure 2-1. Illustration of an omnidirectional semivariogram showing the parameters of nugget effect, sill, and range in image application.**

### 2.4.3 Semivariogram texture classification

Recently, classification techniques have involved the use of semivariogram computed within a local window as a measure of texture (Atkinson and Lewis, 2000). Carr (1996) presented the concept of using the semivariogram to aid digital image classification.

Miranda *et al.* (1998) employed the semivariograms of training classes as textural signatures to perform image classification. Chica-Olmo and Abrarca-Hernandez (2000) computed a set of textural measures based on several semivariogram estimators. However, in their study these semivariogram measures were calculated only for a specific lag of distance in a neighborhood. While geostatistical measures were compared to traditional co-occurrence based textural measures, Carr and Miranda (1998) found that the texture measure achieving the greatest accuracy depended upon the nature of the data and texture. The semivariogram textures were employed as additional input layers to the classifiers, such as the maximum likelihood algorithm or the minimum-distance-to-mean classifier. Image classifications in these studies were performed with hard classifiers in a fully supervised manner.

## **2.5 Accuracy Assessment**

Accuracy assessment allows a degree of confidence to be attached to the mapping results. It also serves to indicate whether the analysis objectives have been achieved. Often reference data is referred to as ground truth. From the percentage of pixels labelled correctly by a classifier, the results can be expressed in a confusion matrix to illustrate the effectiveness of the classifier. A confusion matrix is used for checking the accuracy of a classification. It contains information about actual and predicted classifications done by a classification system as shown in Table 2-1. There are four different possible outcomes of a single prediction for a two-class case: true positive (TP), true negative (TN), false positive (FP), and false negative (FN). Among the four outcomes, only true positives and true negatives are obviously correct classifications. False negative means an incorrect prediction that an instance is positive, and false positive defines an incorrect outcome when an instance negative is misclassified as positive.

The accuracy will be affected by many factors, such as the degree of spectral separability of the classes under study, the characteristics of the test site, and the image processing techniques used. These factors introduce somewhat errors due to either systematic

problem in image generation or personal misunderstanding during image processing. On all accounts, the accuracy of each classification is an obvious limit to its value for the provision of land cover information.

**Table 2-1. Example of a confusion matrix.**

Confusion Matrix		Predicted Class	
		Positive	Negative
Actual Class	Positive	TP	FN
	Negative	FP	TN

### 2.5.1 Accuracy and error

Accuracy refers to the agreement between a measurement and the true or correct value; error refers to the disagreement between a measurement and the true or accepted value (Richards, 1993). Both the accuracy and the error are often calculated in percentage. In image classification, overall classification accuracy is the widely used estimator to illustrate the average percentage of correct classifications. As in Table 2-1, both TP and TN are correctly classified outcomes. Therefore, the proportion of the total number of correct predictions (i.e. TP plus TN) defines the overall accuracy. However, the overall accuracy method considers the classes as having equivalent or similar accuracies whereas classes in the real application often exhibit drastically differing accuracies. More specific measures are thus needed because the overall accuracy does not indicate the accuracy distributed across the individual class.

Producer's accuracy and user's accuracy are the two methods to determine the accuracy of individual classes. The producers are more interested in how correctly the reference samples are classified by the classifiers, while the users are concerned about what

percentage of the classes has been correctly classified and correspond to the reality. As shown in Table 2-1, TP and TN that define the proportion of correct predictions of an individual class respectively represent the producer's accuracy of each class. User's accuracy means the precision that the predicted positive instances are correct.

When only a small number of cover types is of interest, a more appropriate way of presenting the individual classification accuracies is through the use of commission error and omission error. Omission errors correspond to those pixels belonging to the class of interest that the classifier has failed to recognize. As shown in Table 2-1, the proportion of FN to actual positive instances has the equivalent meaning of omission errors. Omission error is directly related to producer's accuracy as:

$$\text{Producer's accuracy} = 1 - \text{Omission error} \quad (2-2)$$

Commission errors are those pixels labelled by the classifier from other classes to the belonging class of interest. As shown in Table 2-1, the proportion of FP to total positive predictions is the commission error. Commission errors are associated with user's accuracy as:

$$\text{User's accuracy} = 1 - \text{Commission error} \quad (2-3)$$

## 2.6 Summary

This chapter has shown some of the relevant theories that are employed in this thesis study. Image classification is the major objective in mapping application. Classification models to be selected for the image classifiers have been introduced. The sources of the classification uncertainty are inherent in either the target objects or the acquired imagery. A "soft" classifier has an ability to describe these uncertainties in the imprecise nature phenomena by fuzzy membership functions. Image textures derived from geospatial techniques can be incorporated with spectral information in the classification procedures

to improve the accuracy. Finally, the accuracy assessment is a quantification method to evaluate the effectiveness of classifiers. Each section in this chapter has reviewed some associated studies corresponding to the topics to highlight the potential strengths and weaknesses of each technique being used. The framework of developing this research is, therefore, formed: the semivariogram texture can be employed as a weighting factor in a partially unsupervised fuzzy classifier for wetland mapping, so that the vagueness of ecological transition zones can be interpreted through the fuzzy membership degrees.

## CHAPTER 3

# FUZZY CLUSTERING THEORY AND DEVELOPED MODIFICATIONS

*“There is one thing even more vital to science than intelligent methods; and that is, the sincere desire to find out the truth, whatever it may be.”*

*Charles Sanders Pierce, 1839-1914*

*American philosopher and scientist*

### 3.1 Concept of Clustering Analysis

Clustering analysis is an important exploratory technique for searching the structures of natural groupings in data. It can provide information for assessing dimensionality, identifying outliers, and suggesting interesting hypotheses that concern relationships. The common application of clustering methods is based on the similarity or distance measures computed from the input data to partition the data set into classes. Similar data are assigned to a same cluster whereas dissimilar data are allocated to different clusters. Clustering analysis thus is to discover natural groupings of a certain data set and to obtain a meaningful partition according to the given definition of similarities.

The two major categories of clustering techniques are the hierarchical clustering methods and the non-hierarchical clustering methods. The hierarchical clustering methods proceed by either a series of successive mergers or a series of successive divisions. The non-

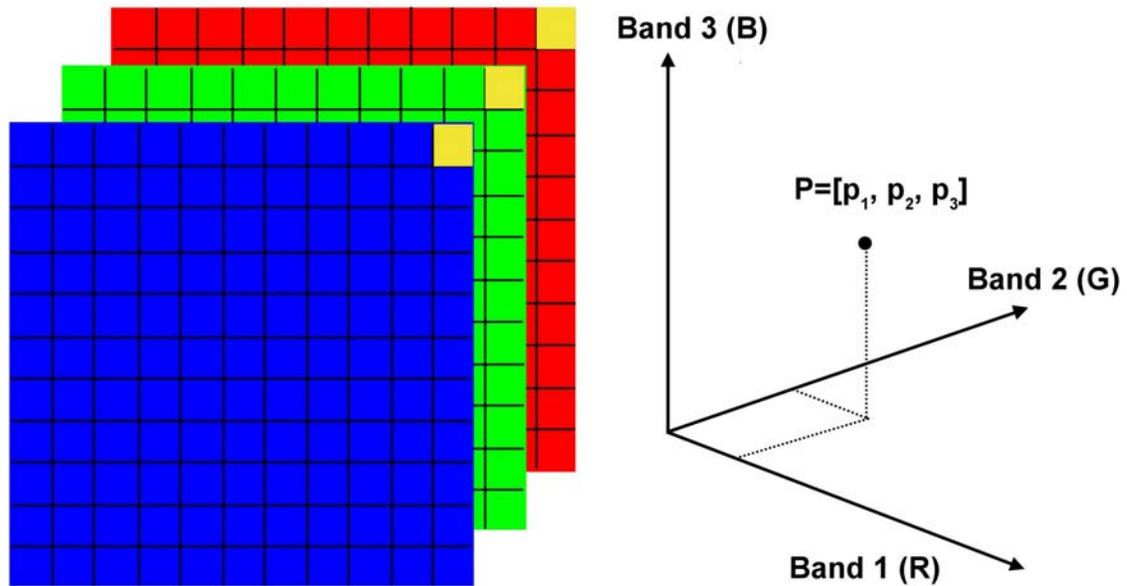
hierarchical clustering process is based on a monotonically increasing ranking of strengths and the clusters themselves progressively become members of larger clusters. The non-hierarchical clustering methods are also called partitioning methods. The similarity matrix, in non-hierarchical clustering algorithm, does not need to be determined initially and data do not have to be stored during the computer run. The non-hierarchical clustering methods have more flexibility and can be applied to larger data sets than the hierarchical techniques. The non-hierarchical clustering proceeds iteratively. In general, partitioning methods first start from either an initial partition of the data set or an initial set of seed points, which will form the nuclei clusters (Johnson and Wichern, 1988). Each pixel are then classified and assigned to the closest cluster in the second step. In the third step the new cluster mean vectors are calculated based on all the pixels belonging to the cluster. The second and third steps are repeated until the “change” of the objective function is smaller than a predetermined threshold.

### 3.1.1 Multidimensional feature space

Remote sensing is a technology for sensing and recording reflected or emitted energy, which is called Electro-Magnetic (EM) radiation, at various wavelengths or polarizations. Because many ground materials reflect and emit EM radiations in unique ways that form their spectral signatures, they can be characterized and identified by analyzing the spectral information of a multispectral imagery. In an image, a pixel is the basic unit to represent ground objects. However, the pixel size is associated with the spatial resolution that defines the smallest level of spatial detail perceived in an image. The smallest ground objects may be distinguished as separate entities in the image when having sizes larger than a single pixel.

Pixels can then be viewed as a picture element in a  $p$ -dimensional vector space, which is called the feature space. If a multispectral imagery has  $p$  bands, each pixel is a  $p \times 1$  vector. Figure 3-1 shows a pixel of a multispectral imagery viewed as a picture element in a 3D multispectral feature space. This space is the basis of clustering analysis. When the location of all image pixels are plotted in the feature space, pixels that belong to a

same class will form a cluster. After a clustering process, all pixels in the image will be assigned to one of the predefined information class.



**Figure 3-1. Illustration of feature space. A pixel of a multispectral imagery can be viewed as a picture element in a feature space.  $p_1$ ,  $p_2$  and  $p_3$  are the brightness value of the pixel located at row  $i$ , column  $j$  and band 1, 2 and 3 respectively.**

### 3.1.2 Similarity measures

A measure of the similarity between the pixels is necessarily required to produce a simple group structure from a complex data set. Similarity measures can be defined as functions that automatically map data points represented by feature vectors to indicate the degree of similarity between the objects. The distance between two data points is the central idea for the mathematical representation of similarities. Two techniques are provided for distance measurement in the following.

### 3.1.2.1 Euclidean distance

The Euclidean distance is a straight-line distance measure. If we consider two arbitrary data points  $\mathbf{x}$  and  $\mathbf{v}$  in a  $p$ -dimensional feature space, the straight-line distance between the ordinates  $\mathbf{x}=(x_1, x_2, \dots, x_p)$  and  $\mathbf{v}=(v_1, v_2, \dots, v_p)$  is given by:

$$D^2(\mathbf{x}, \mathbf{v}) = (x_1 - v_1)^2 + (x_2 - v_2)^2 + \dots + (x_p - v_p)^2 \quad (3-1)$$

However, the Euclidean distance may not be satisfactory for many purposes because of the assumption of an equal contribution of each coordinate to the distance results. When using the Euclidean distance in a clustering analysis, it gives good results only when all clusters are spheroids of the same size or when all clusters are well separated in the feature space (Krishnapuram and Kim, 1999). Once the measurement of each coordinate is subject to random fluctuations of scale magnitudes, the variability of these coordinates should be taken into account when computing the distance; that is, without considering the scale variability among the coordinates, the coordinate with larger number scale dominates the distance computation and results in a bias outcome. This suggests a different distance measure: the Mahalanobis distance.

### 3.1.2.2 Mahalanobis distance

The Mahalanobis distance is a statistical distance measurement invented by P.C. Mahalanobis in 1936. Because the correlations of a data set are taken into account, the variability of the coordinates are standardized on an equal footing with one another. It is a useful way of determining the similarity of an unknown sample set to a known one. Thus the Mahalanobis distance of two data points  $\mathbf{x}$  and  $\mathbf{v}$  is of the form:

$$D^2(\mathbf{x}, \mathbf{v}) = (\mathbf{x} - \mathbf{v})^T \mathbf{M}^{-1} (\mathbf{x} - \mathbf{v}) \quad (3-2)$$

where  $\mathbf{M}$  denotes the covariance matrix and  $T$  indicates the transpose of a vector. Since the Mahalanobis distance can give the description for an ellipsoidal distributed data set, it was introduced to design the Gustafson and Kessel fuzzy clustering algorithm (1979). It

is to be noted that the Mahalanobis distance is the same as the Euclidean distance if the covariance matrix is the identity matrix.

### 3.2 Theory of Fuzzy Clustering

“Fuzzy” is a word that means uncertain, incomplete, and not fully determined. Fuzzy logic is a means to handle the concept of partial truth, that is, true values between "completely true" and "completely false". Zadeh (1965) initiated the concept of a fuzzy set theory in the early 1960s. He stated that the development of a fuzzy set theory was motivated in large measure by problems in pattern classification and cluster analysis. He thus extended the notion of binary membership to accommodate various degrees of membership on the real continuous interval from 0 to 1, where the endpoint of 0 and 1 conform to no membership and full membership, respectively. Later in the beginning of the 1970s, Zadeh presented the foundations of a linguistic synthesis and showed how vague logical statements could be used to construct computational algorithms that might be used to derive inferences from vague data (Zadeh, 1973).

Clustering techniques without introducing the fuzzy theory enables a clear distinction between objects. However, geographic objects in the natural world are not often illustrated by a sharp boundary in contrast to man-made objects. Rather, almost every natural geographic object has a vague boundary (Couclelis, 1996). Fuzzy clustering is often suited better than crisp assignments of the data to the clusters. A key difference between crisp and fuzzy sets is their membership function: a crisp set has a unique membership function, whereas a fuzzy set can have an infinite number of membership functions to represent it (Ross, 1995).

Figure 3-2 gives a simple hypothetical example to illustrate the membership function for a crisp and a fuzzy set. The variations in land covers run along an environmental gradient and control the distribution of three classes: upland, wetland, and open water. The membership value,  $\mu(x) \in [0,1]$ , measures the similarity degree of the element  $x$

belonging to the class. For a crisp set function, the binary logic is applied to the land cover distribution; that is, the strength of membership is either 0 or 1. A crisp boundary is given to divide a gradual geography into two classes. By contrast, a fuzzy set function allows the classes to intergrade and co-exist spatially between the “core” areas of the two classes. The “core” areas of each class are characterized by a complete coverage of the relevant class. The membership to one class should be expected to decline from the class’s core area while the membership of the other class should rise.

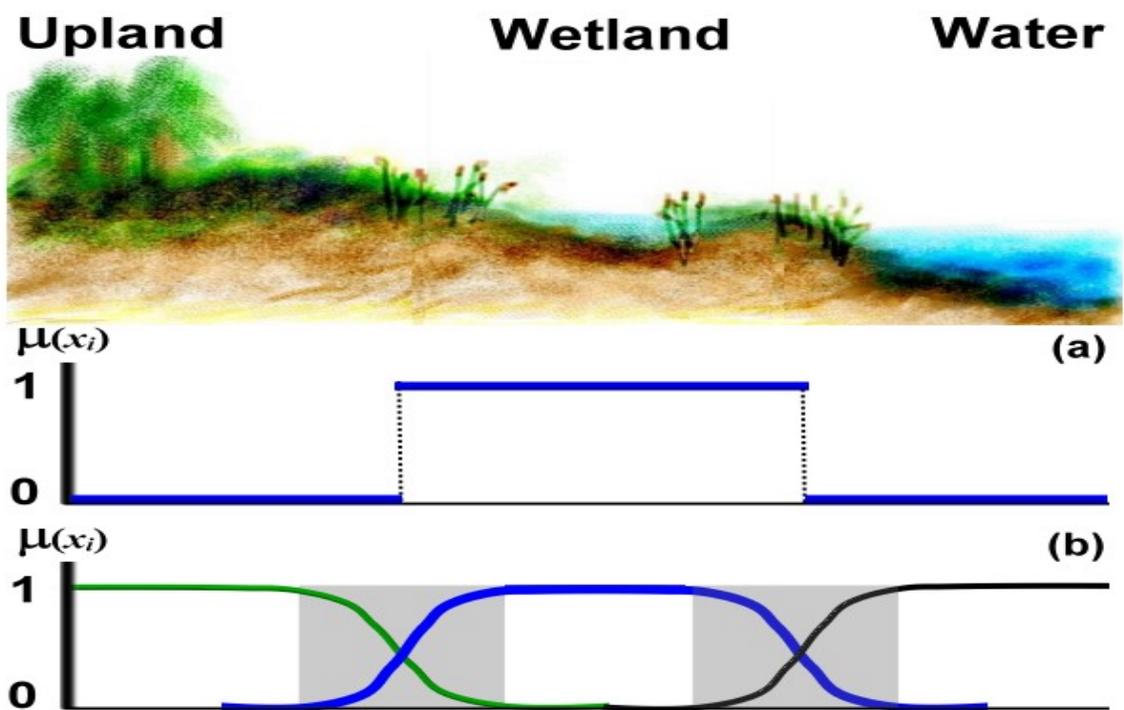


Figure 3-2. Illustration of the membership function for (a) a crisp and (b) a fuzzy set.

### 3.3 Overview of Fuzzy C-Means

The Fuzzy C-Means (FCM) clustering algorithm developed by Bezdek (1984) is commonly implemented and employed as a constrained soft partition to provide a cluster description. Bezdek and Pal (1992) have stated that feature vectors should be allowed to have degrees of membership in more than one class. This statement strongly supports the

fuzzy classification of a multispectral image. This clustering algorithm uses image attributes as information to divide the image domain into a predetermined number of clusters. It then assigns every pixel in the image with a certain degree of membership to the clusters. In the last decades, fuzzy partitions have been the approaches applied to image classification and intended to overcome the intrinsic limitations of crisp partitions. The terminology “fuzzy segmentation” thus can be found in several papers related to image analysis and pattern recognition (Tolias and Panas, 1998; Noordam *et al.*, 2000; Chumsamrong *et al.*, 2000; Thitimajshima 2000; Gordan *et al.*, 2002; Ahmed *et al.*, 2002; Chamorro-Martinez *et al.*, 2003; Cuevas *et al.*, 2004)

### 3.3.1 Fuzzy C-Means clustering algorithm (FCM)

The idea of the FCM clustering technique is to minimize the objective function that can be expressed in the form of a generalized least-squared error function. The objective function specifies a sum of distances, characterized by the Euclidean norm, between the feature points and the corresponding cluster centres. This algorithm is noted  $J(X;U,V)$ , where  $U$  is the partition matrix and  $V$  is a vector of the cluster centres. Given a  $n_1 \times n_2$  sized image, the data set can be expressed as a  $p$ -dimension feature vector  $X = \{x_1, x_2, \dots, x_N\} \in R^p$ , where  $N = n_1 \times n_2$  is the number of total pixels in the image. The objective function is then given by:

$$J(X;U,V) = \sum_{i=1}^N \sum_{k=1}^C (\mu_{ik})^m (d_{ik})^2 \quad (3-3)$$

where  $C$  is the number of clusters and  $d_{ik}$  is the Euclidean distance measure (in a  $p$ -dimension feature space,  $d_{ik} \in R^p$ ) between the  $i^{\text{th}}$  data sample  $x_i$  and the  $k^{\text{th}}$  cluster centre  $v_k$ , written as:

$$d_{ik} = d(x_i - v_k) = \|x_i - v_k\| = \left[ \sum_{j=1}^p (x_{ij} - v_{kj})^2 \right]^{\frac{1}{2}} \quad (3-4)$$

Since the location of each data sample  $x_{ij}$  can be described in the  $R^p$  feature space, each cluster centre should be represented in the same feature space. Therefore, the  $k^{\text{th}}$  cluster centre is a vector of length  $p$ ,

$$v_k = \{v_{k1}, v_{k2}, \dots, v_{kp}\} \quad (3-5)$$

The optimum partition matrix,  $U^*$ , that produces the minimum value of the objective function  $J$  is that of interest in the fuzzy c-means clustering algorithm. That is:

$$J(X; U^*, v^*) = \min_{(U, V)} J(X; U, V) \quad (3-6)$$

The iterative optimization, devised by Bazdek (1984), is an effective algorithm for fuzzy classification. The solution to Eq. (3-6) cannot be promised to be a global optimum but rather a local minimum within a given level of accuracy. Because fuzzy sets allow for degrees of membership, a single point can have partial membership in more than one class. It will be useful to describe the membership value that the  $i^{\text{th}}$  feature point has in the  $k^{\text{th}}$  class with the following notation:

$$\mu_{ik} \in [0, 1] \quad \text{for all } k = 1, 2, \dots, C \text{ and } i = 1, 2, \dots, N \quad (3-7)$$

In the optimization process, the partition matrix is subject to two constraints. First, no cluster should be empty and no cluster should contain all the feature points. This qualification is manifested in the following expression:

$$0 < \sum_{i=1}^N \mu_{ik} < N \quad \text{for all } k = 1, 2, \dots, C \quad (3-8)$$

Second, the overall weight of all membership values for a single feature point in all of the clusters has to be unity, i.e.:

$$\sum_{k=1}^C \mu_{ik} = 1 \quad \text{for all } i = 1, 2, \dots, N \quad (3-9)$$

By taking these constraints into consideration, the partition matrix,  $U$ , can be solved with an iterative scheme leading either to a local minimum or a saddle point of  $J$ . Meanwhile high membership values are assigned to pixels whose intensities are close to the centroid of its particular cluster and low membership values are assigned to pixel data located far from the centroid. The membership degrees associated to each class are given in the partition matrix  $U^* = \{\mu_{ik}^*\}$ , which follows from the overlapping character of the classes and the infinite number of membership values.

In Eq. (3-3) the weighting parameter,  $m$ , is called the fuzzy index whose range for the membership exponent is  $m \in [1, \infty)$ . It controls the degree of fuzziness in the classification process. When  $m \rightarrow 1$ , the clusters tend to be crisp, *i.e.* the membership degree  $\mu_{ik}$  equals either 1 or 0. Conversely, as  $m \rightarrow \infty$ , larger powers of the membership values that are less than or equal to 1 make the objective function approaching to zero,  $J \rightarrow 0$ . As  $m$  increases, the membership function becomes increasingly fuzzier and the convergence of the algorithm tends to be slower. Pham (2001) compared different values of fuzzy indices in the FCM clustering algorithm for magnetic resonance image (MRI) segmentation. His study denoted that when the selection of  $m$  was greater than 1, the membership function did not become binary. However, there is no theoretical optimum choice of  $m$  in the literature. The literature seems to report that  $m = 2$  is usually chosen (Pham, 2001; Ahmed *et al.*, 2002). Therefore, the fuzzy index is chosen to be 2 in this study.

### 3.4 Modification of Fuzzy C-Means

When the FCM classifier is applied for wetland mapping, two major problems need to be clarified. First, the Euclidean distance used in the FCM clustering algorithm for similarity measurement assumes that every cluster comes from a spherical normal distribution with different means but an identical variance (Krishnapuram and Kim, 1999). If clusters

having spherical distributions fit in with the assumption, the Euclidean distance works better for the clustering. Furthermore the implementation of the clustering algorithm with the Euclidean distance is easier. Unfortunately, spherically distributed clusters do often not happen in a remotely sensed image data set, especially for natural land cover classes. For example, a cluster of “water” pixels is compact and well detected as one distinct cluster while a “wetland” cluster is usually more or less elongated with a much larger variability compared to the “water” cluster. This is because the inherently spectral characteristics of the land cover itself that are spectrally mixed and that vary gradually in the landscape.

Second, the standard FCM clustering algorithm uses only the spectral information in the objective function. Although using only the spectral information in clustering is intuitive and allows the algorithm to be easily applied to many applications, spatial variability inherent in an image implies that some structure properties of landscape elements should be taken into account. As image textures highlight the relationship between a pixel and its neighbourhood, they characterize the spatial variability of data and provide additional information for image classification. However, the objective function of the standard FCM clustering algorithm does not take the spatial information into account.

These two concerns motivate the modification of the FCM clustering algorithm for wetland mapping from two aspects: replacing the Euclidean distance by the Mahalanobis distance and incorporating the semivariogram texture as spatial guidance in the conventional FCM clustering algorithm.

### **3.4.1 Semivariogram Guided Fuzzy C-Means clustering algorithm (SGFCM)**

Since only the pixel intensity is considered in the distance function of the FCM classifier, we propose a modification to Eq. (3-4) by introducing the texture component as a weighting factor. This introduction allows the labelling of a pixel to be influenced by the semivariograms of the training sites and thus by the spatial information of the classes under consideration. The semivariogram texture involves the regional homogeneity of

observations. A  $s \times s$  pixel region surrounding a pixel  $x_i$  to be classified is extracted from an image for the semivariogram computation. In this instance,  $s$  is of the same size as the one used for the training sites. For each changing lag  $h$ , the average distance is computed between the semivariogram for the extracted region  $\gamma_i(h)$  and the semivariogram for each of the training classes  $\gamma_k(h)$ . The semivariogram texture parameter  $w_{ik}$  can be formulated as:

$$w_{ik} = \frac{I}{s-I} \sum_{h=1}^{s-I} \|\gamma_i(h) - \gamma_k(h)\| = \frac{I}{s-I} \sum_{h=1}^{s-I} \left[ \sum_{j=1}^p (\gamma_{ij}(h) - \gamma_{kj}(h))^2 \right]^{1/2} \quad (3-10)$$

where  $\gamma(h) \in R^p$  is referred to the semivariance expressed in Eq. (2-1). The modified distance function of Eq. (3-4) corresponding to pixel  $i$  and cluster  $k$  can thus be written as:

$$d'_{ik} = d'(x_i - v_k) = \|x_i - v_k\| \times (w_{ik})^{1/2} \quad (3-11)$$

As mention above, the Euclidean distance is only suitable for spherically distributed data set. To overcome the drawback due to the Euclidean distance, the Mahalanobis distance is used as the distance measure, i.e.:

$$(d'_{ik})^2 = w_{ik} \|x_i - v_k\|^T A_k^{-1} \|x_i - v_k\| \quad (3-12)$$

where the norm matrix  $A_k$  is a positive definite symmetric matrix. Since the norm matrix determines the size and the shape of the points enclosed within a given distance of the centre, the above distance measure is meaningful only when all clusters are expected to be ellipsoids with same orientation and size. Although the Mahalanobis distance is well known for its invariance to linear transformations, it cannot be used directly in the clustering algorithm by making the covariance matrix  $A_k$  diagonal as

$$A_k = \begin{bmatrix} \sigma_{k1}^2 & 0 & \cdots & 0 \\ 0 & \sigma_{k2}^2 & \cdots & 0 \\ \vdots & \vdots & \sigma_{kj}^2 & \vdots \\ 0 & 0 & \cdots & \sigma_{kp}^2 \end{bmatrix}$$

where:

$$\sigma_{kj}^2 = \frac{\sum_{i=1}^N (\mu_{ik})^m (x_{ij} - v_{kj})^2}{\sum_{i=1}^N (\mu_{ik})^m} \quad \text{for all } j = 1, 2, \dots, p \quad (3-13)$$

In this case, every point has been proven to be shared equally by all clusters, that is, meaning that  $\mu_{ik} = 1/C$ . This result is not expected for fuzzy membership values (Krishnapuram and Kim, 1999). Instead, the data set is linearly transformed to have the covariance matrix of cluster  $k$  becoming diagonal in the Gustafson and Kessel (1979) fuzzy clustering algorithm. In the general case, the distance measure becomes:

$$(d'_{ik})^2 = w_{ik} (\rho_k \det M_k)^{1/p} \|x_i - v_k\|^T M_k^{-1} \|x_i - v_k\| \quad (3-14)$$

where  $\rho_k$  is a constant and  $M_k$  denotes a fuzzy covariance matrix associated with a given class  $k$  and is defined by:

$$M_k = \frac{\sum_{i=1}^N (\mu_{ik})^m \|x_i - v_k\|^T \|x_i - v_k\|}{\sum_{i=1}^N (\mu_{ik})^m} \quad (3-15)$$

Gustafson and Kessel (1979) recommend that  $\rho_k = 1$  to preserve the volume of the ellipsoidal clusters after transformation. Substituting Eq. (3-14) into (3-3), the modified objective function is obtained as

$$\begin{aligned}
J_m(X;U,V) &= \sum_{i=1}^N \sum_{k=1}^C (\mu_{ik})^m (d'_{ik})^2 \\
&= \sum_{i=1}^N \sum_{k=1}^C (\mu_{ik})^m w_{ik} (\rho_k \det M_k)^{1/p} \|x_i - v_k\|^T M_k^{-1} \|x_i - v_k\|
\end{aligned} \tag{3-16}$$

As shown by Gustafson and Kessel (1979), minimizing the objective function by using the distance in Eq. (3-14) is equivalent to minimizing

$$J_m(X;U,V) = \sum_{i=1}^N \sum_{k=1}^C (\mu_{ik})^m w_{ik} \|x_i - v_k\|^T A_k^{-1} \|x_i - v_k\| \tag{3-17}$$

Here, the norm matrix  $A_k = \frac{M_k}{\sqrt[p]{\rho_k \det M_k}}$  is subject to  $\det A_k = \rho_k$  with  $\rho_k = 1$ . The function  $J_m$  can be solved while enforcing the two constraints given in Eq. (3-8) and (3-9) by means of Lagrange multipliers:

$$F_m(X;U,V) = \sum_{i=1}^N \sum_{k=1}^C (\mu_{ik})^m (d'_{ik})^2 + \lambda \left( 1 - \sum_{k=1}^C \mu_{ik} \right) \tag{3-17}$$

Taking the partial derivatives of  $F_m$  with respect to  $\mu_{ik}$  and setting the result to zero for  $p > 1$  can derive the formula for updating the membership degrees.

$$\frac{\partial F_m}{\partial \mu_{ik}} = m(\mu_{ik})^{m-1} (d'_{ik})^2 - \lambda = 0 \tag{3-18}$$

Solving for  $\mu_{ik}$ , it can be written as:

$$\mu_{ik} = \left[ \frac{\lambda}{m(d'_{ik})^2} \right]^{\frac{1}{m-1}} \tag{3-19}$$

Since the sum of the membership values for a single feature point in all classes equal to one, it leads to:

$$\sum_{g=1}^C \left[ \frac{\lambda}{m(d'_{ig})^2} \right]^{m-1} = 1 \quad (3-20)$$

and  $\lambda$  is solved as:

$$\lambda = \frac{m}{\left[ \sum_{g=1}^C (d'_{ig})^{-2/(m-1)} \right]^{m-1}} \quad (3-21)$$

Substituting Eq. (3-21) into Eq. (3-19) results in the following formula for updating the membership functions:

$$\mu_{ik}^* = \frac{(d'_{ik})^{-2/(m-1)}}{\sum_{g=1}^C (d'_{ig})^{-2/(m-1)}} \quad (3-22)$$

Taking the derivative of  $F_m$  with respect to  $v_k$  and setting the result to zero, we will have the differential equation leading to:

$$\frac{\partial F_m}{\partial v_k} = \sum_{i=1}^N (\mu_{ik})^m w_{ik} \|x_i - v_k\| = 0 \quad (3-23)$$

Thus, the zero-gradient condition for the cluster centre is expressed as

$$v_k^* = \frac{\sum_{i=1}^N (\mu_{ik})^m w_{ik} x_i}{\sum_{i=1}^N (\mu_{ik})^m w_{ik}} \quad (3-24)$$

and the covariance matrices are updated according to

$$A_k = \frac{M_k^*}{\sqrt[p]{\det M_k^*}} \quad (3-25)$$

where

$$M_k^* = \frac{\sum_{i=1}^N (\mu_{ik}^*)^m \|x_i - v_k^*\|^T \|x_i - v_k^*\|}{\sum_{i=1}^N (\mu_{ik}^*)^m} \quad (3-26)$$

The iterative procedures repeatedly replace the membership function and the cluster centre until the change in memberships drops below a given threshold:

$$|U^{(t+1)} - U^{(t)}| \leq \varepsilon \quad (3-27)$$

### 3.5 Defuzzification of Fuzzy Membership Function

Once the fuzzy clustering procedure has stopped, the partition matrix  $U$  must be defuzzified to obtain a final classification of the pixels. The data sets have been computed, reasoned, and modeled with fuzzy information. However, most of the actions and decisions implemented by human or machines are crisp or binary. “Defuzzifying” the fuzzy results generated through a fuzzy set analysis is, therefore, necessary for various applications to reduce a fuzzy quantity into a single scalar quantity. The output of a fuzzy process can be the logical union of two or more fuzzy membership functions defined on the universe of discourse of the output variable.

#### 3.5.1 Maximum membership defuzzy principle

In the application of image classification, the maximum membership defuzzy principle is the most commonly used defuzzification criterion for either fuzzy classifications or possibilistic approaches. Even the widely used Maximum Likelihood Classifier employs

the concept of maximum operator for the classification. The operation is known as the highest method, which is limited to peaked output functions. It is given by the algebraic expression:

$$\mu_{x_i}(k^*) \geq \mu_{x_i}(k) \quad \text{for all } k \in C \quad (3-28)$$

and is shown graphically in Figure 3-3. The maximum membership defuzzification does not consider the relative strength of the memberships for other classes; it has to assign a pixel into a class without taking the coexisting classes into consideration at all. But how can the pixel be assigned to a class over the others if the highest membership value is extremely low? The classification error is, therefore, committed because the maximum membership method ignores the similarity between the coexisting classes. To emphasize the pixel ambiguity leading to uncertain classification, a method using the alpha ( $\alpha$ )-cut set for partition matrix defuzzification is proposed in the following.

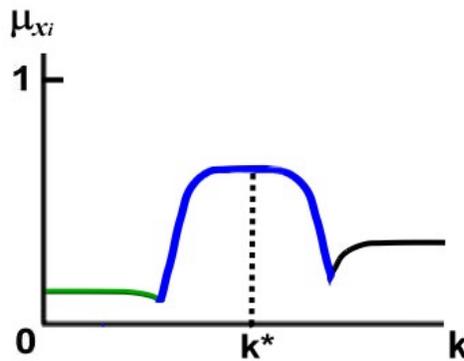


Figure 3-3. Illustration of maximum membership defuzzification.

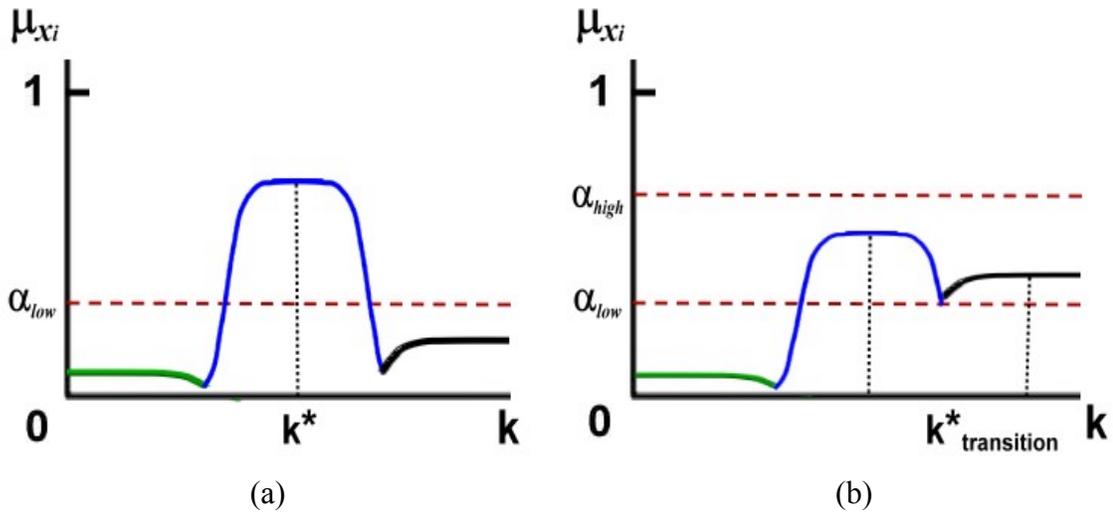
### 3.5.2 Alpha ( $\alpha$ )-cuts defuzzy rule

Addressing the vagueness in the transition zones, this study utilizes the  $\alpha$ -cuts as a threshold mode defuzzification method. A  $\alpha$ -cuts method rescale the membership values to one for all elements of the fuzzy set having membership values greater than or equal to  $\alpha$ , and zero for all elements of the fuzzy set having membership values less than  $\alpha$ . Therefore, in this study, two threshold values according to the maximum ambiguity are

given as the lower threshold  $\alpha_{low}$  and the upper threshold  $\alpha_{high}$ . Since the maximum ambiguity of each pixel's membership depends on the number of clusters, we can have the lower threshold  $\alpha_{low}$  defined as  $1/C$ . Any pixel having a membership value lower than  $\alpha_{low}$  would not be assigned to the associated class, whereas a pixel is assigned to a certain class only when the associated membership value is higher than the upper threshold  $\alpha_{high}$ , where  $\alpha_{high} = 1 - 1/C$ , or when the associated membership value is the only one larger than the lowest threshold. As shown in Figure 3-4, the ambiguous pixels can be expressed according to their membership values as follow:

$$\alpha_{low} \leq \mu_{x_i}(k_{transition}^*) < \alpha_{high} \quad \text{for all } k \in (2^C - 1) \quad (3-29)$$

To classify the ambiguous pixels, which are not assigned to any of the major classes, new classes are created. By creating the “transition” classes, the classifier allows the transition areas between any two major classes to be captured. For example, if the given number of the major classes is three, the maximum number of classes after defuzzification will be seven or lower (i.e.  $2^3 - 1 = 7$ ).



**Figure 3-4. Illustration of two different  $\alpha$ -cut sets for classification: (a) maximum membership criteria with a lower threshold restriction and (b) “transition” classes are created to allocate the ambiguous pixels.**

### 3.6 Measurement of Uncertainty

These vague areas are the regions with highly spatial uncertainty in the classification map. Burrough (1996) suggested the term “confusion index ( $CI$ )” to inspect interference between multiple membership maps. The measure calculates the difference between the highest and the second highest membership class per pixel. It is given as

$$CI_i = 1.0 - (\mu_{ik, \max} - \mu_{ik, 2^{nd} \max}) \quad (3-30)$$

The confusion index values are scaled between zero and one, that is,  $CI \in [0, 1]$ . Any pixel with  $CI$  values close to one has a higher uncertainty, meaning that the two classes are similar. For a pure pixel, it will have a maximum membership value for the associated class and for the other classes it will have very low membership values. Then  $CI$  will be close to or equal to zero.

### 3.7 Summary

This chapter demonstrates the principal fuzzy classifier—the Semivariogram Guided Fuzzy C-Means clustering algorithm—used in the study. From the point of view of fuzzy clustering analysis, the definition of similarity measurement is an important factor in a clustering approach. Similarity is defined as a distance function to decide to which cluster the data point belongs in the clustering algorithm. Because the standard FCM clustering algorithm does not take the spatial information into account and assumes that the clusters inherent in the data set are well separated from each other, a more robust fuzzy clustering algorithm has been developed for wetland mapping. The modification of the standard FCM algorithm includes two tasks: replacing the Euclidean distance by the Mahalanobis distance and incorporating the semivariogram texture as spatial guidance in the fuzzy clustering algorithm. The derivatives of the algorithm are shown in this chapter. Furthermore, a threshold defuzzification method has been proposed to emphasize the

ambiguous pixels in the transition zones by allocating them to newly created “transition” classes since the classification uncertainty is always accompanied with these pixels.

## CHAPTER 4

# METHODOLOGY

*“The practice of conservation must spring from a conviction of what is ethically and aesthetically right, as well as what is economically expedient. A thing is right only when it tends to preserve the integrity, stability, and beauty of the community, and the community includes the soil, waters, fauna, and flora, as well as people.”*

*Aldo Leopold, 1887-1948*

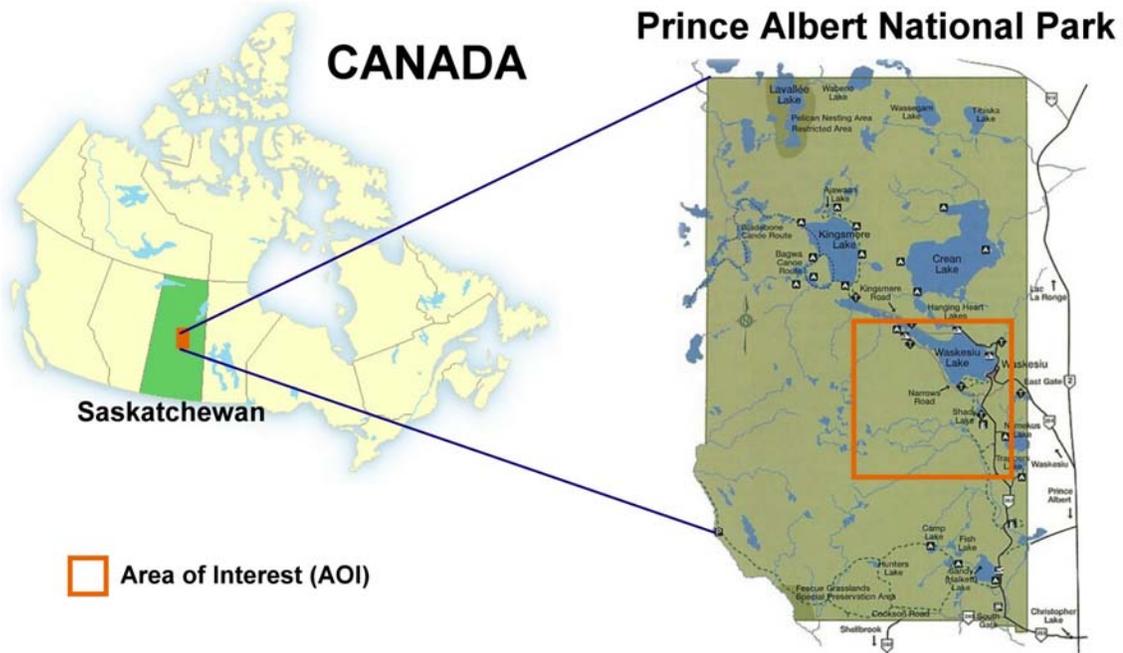
*American ecologist, wildlife biologist, and forester*

### 4.1 Study Area Description

The area of interest is located within the boundary of Prince Albert National Park, Saskatchewan, Canada. The geographic location of the area is predominantly situated in the south-Boreal Plains ecoregion. The Churchill River basin runs through the study site that has a geographical extent of 53°45'00''N to 54°00'00''N and 106°00'00''W to 106°25'00''N as shown in Figure 4-1. Because of its northern mid-continental location, the mean monthly temperatures range from approximately -17.2°C in January to 17.5°C in July according to the 7-year Meteorological Service of Canada (MSC) Normals for 1996-2002. The mean monthly precipitations can vary significantly from 80.2mm in July to only 14.7mm in November.

The relatively simple topography is defined by low hills and ridges and by lake basins. The ground relief of the area varies gradually from the western hills to the eastern valleys

and the altitude ranges from 512m to 689m above sea level. The area is usually well rounded so that the mean slope is about 6 percent except for some rugged areas that have steep slopes varying from 20 to 70 percent. In the south, the land is gently undulating or nearly flat. Open to semi-open expanses of true prairie are found there too (Soper, 1952).



**Figure 4-1. Area of interest is located within the boundary of Prince Albert National Park, Saskatchewan, Canada.**

The lake system of the park is remarkable. Hundreds of water bodies vary from ponds to fair-sized lakes. Since the region was once heavily glaciated, there are numerous bodies of water, bogs, sand and gravel ridges, and deep deposits of boulder clay. Many small ponds dominate Waskesiu Hill located in the west of the study area and Lake Waskesiu in the east side of the park is one of the largest and most important lakes.

The landscapes of the area consist of a range of vegetation types. Two of the major vegetation zones are the mixed wood section of the boreal forest region and the aspen grove section. Forest canopies are often controlled by small changes in relief and soil

drainage. For example, aspen occurs on the uplands while jack pine are on minor ridges. Because some small (10 to 30m) ponds occur in the canopy, local wet areas characterize the site. In the poorly drained areas throughout the study area, black spruce with some tamarack is found in bogs, while sedge vegetation with discontinuous cover of tamarack or swamp birch is found in the fen areas (ORNL DAAC, 2001). The commonest emergent aquatic plant is the roundstem bulrush (*Scirpus*), which often forms a narrow belt along the shores of lakes and ponds. Cattail (*Typha*) is represented in widely scattered stands and pondweeds (*Potamogeton*), water-milfoil (*Myriophyllum*), coontail (*Ceratophyllum*) and arrowhead (*Sagittaria*) are well grown in few small lakes, ponds, and streams (Kiil *et al.*, 1973).

## 4.2 Data

### 4.2.1 Satellite imagery

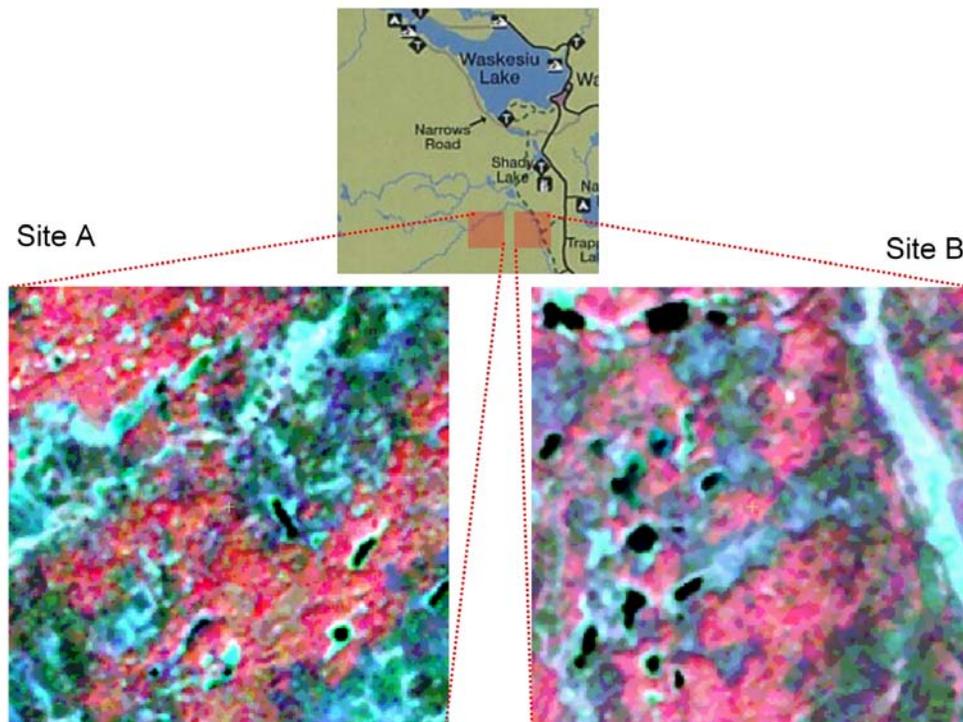
The available imagery data for this study was acquired in August 1999 by Landsat 7 Enhanced Thematic Mapper Plus (ETM+) satellite. Except for the thermal infrared, this multispectral imagery consists of six bands in different spectral bandwidth: blue (0.45-0.52  $\mu\text{m}$ ), green (0.53-0.61  $\mu\text{m}$ ), red (0.63-0.69  $\mu\text{m}$ ), near infrared (0.78-0.90  $\mu\text{m}$ ), and two middle infrareds (1.55-1.75  $\mu\text{m}$  and 2.09-2.35  $\mu\text{m}$ ). All the bands have a spatial resolution of 25m. The image was Universal Transverse Mercator (UTM) projected under Zone 13. Two 200  $\times$  200 pixel subscenes of the area of interest were used for the experiment in order to test the effectiveness of the developed classifier. Figure 4-2 shows the colour composite images of TM 432 of the two testing sites, *i.e.* test area A and test area B.

### 4.2.2 Reference data

The reference data used in this thesis research were acquired from two sources. First, the topographic data at a 1:50,000 scale was obtained from the National Topographic Database (NTDB) developed by Geomatics Canada. The topographic data includes wetland thematic maps, which can be used as the reference data for the validation of the

classification results, and digital elevation model (DEM) data. The accuracy of the NTDB data is about 25 metres (NRCan, 2003). In other words, about one pixel pixel error exists in the topographic maps.

The second source of the reference data was referred to the data of the project “Boreal Ecosystem-Atmosphere Study (BOREAS)” conducted in central Canada from 1993 to 1996 (ORNL DAAC, 2001). The project is a large-scale experiment to investigate interactions between the boreal forest biome and the atmosphere. Our test areas are covered in the portion of the BOREAS Southern Study Area (SSA). According to the description of the project, the classification map derived from the Landsat TM imagery acquired in 1990 was used as a reference map in our study for selecting training sites of information classes.



**Figure 4-2. Color composite images (TM 432) showing the test areas subset from the area of interest.**

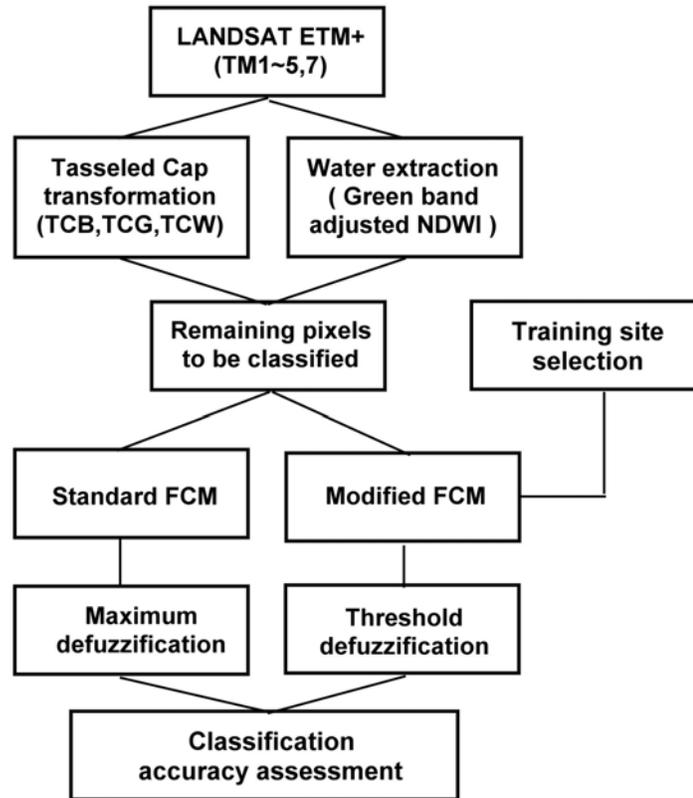


Figure 4-3. The framework of the image data processing.

### 4.3 Image Pre-processing

Figure 4-3 presents the framework of the image preparation. First, the digital numbers of the multispectral imagery are converted to at-satellite reflectance to achieve radiometric consistency between the images. This standard measurement unit (reflectance) allows us to compare between dates and sensors, i.e. our reference map was established in 1990 after an at-satellite reflectance transformation of the Landsat images. Then the data are transformed into the at-satellite reflectance-based tasseled cap features. Then water pixels are excluded from the data set before the fuzzy clustering algorithm is applied. Very low reflectance responses make water bodies easy to distinguish from the other ground features. Although water class often shows higher producer's accuracy, excluding the water pixels can promise that misclassified water pixels would not reduce the overall

classification accuracy. The remaining pixels are used as the input data set for both the standard Fuzzy C-Means (FCM) classifier and the developed Semivariogram Guided Fuzzy C-Means (SGFCM) classifier. Furthermore, defuzzification with the maximum function and the alpha-cuts method is applied to harden the fuzzy output into a single scalar quantity (*i.e.* land cover class) and the classification accuracy is computed to compare the classification results of the two classifiers.

#### 4.3.1 Radiance conversion

After a  $3 \times 3$  median filter was applied to the image to remove any noise such as “salt and pepper” shown on the image, raw digital numbers were converted to at-satellite reflectances for further usage according to Landsat 7 Users Handbook (Irish, 2000). The digital numbers were converted back to the radiance unit according to:

$$L_{\lambda} = Gain \times DN + Offset \quad (4-1)$$

where

- $L_{\lambda}$  is the spectral radiance at the sensor aperture in  $watts / m^2 - ster - \mu m$  ;
- $Gain$  is the rescaled gain provided in the ancillary data record from Table 4-1 in  $watts / m^2 - ster - \mu m$  ;
- $Offset$  is the rescaled bias provided in the ancillary data record from Table 4-1 in  $watts / m^2 - ster - \mu m$  ;
- $DN$  is the raw digital number of each pixel.

#### 4.3.2 Reflectance conversion

Accordingly, the spectral radiance was converted to the planetary reflectance as normalization for the solar irradiance to reduce the variability between scenes. The combined surface and atmospheric reflectance could be computed according to:

$$\rho = \frac{\pi \cdot L_{\lambda} \cdot d^2}{ESUN_{\lambda} \cdot \cos \theta_s} \quad (4-2)$$

where

- $\rho$  is the planetary reflectance (unitless);
- $L_\lambda$  is the spectral radiance at the sensor aperture in  $watts/m^2 - ster - \mu m$ ;
- $d$  is the earth-sun distance in astronomical units ( $d = 1.0112$  for the available image data in this study);
- $ESUN_\lambda$  is the mean solar spectral irradiances from Table 4-1 in  $watts/m^2 - \mu m$ ;
- $\theta_s$  is the solar zenith angle in degrees ( $\theta_s = 43.48^\circ$  for the available data in this study).

**Table 4-1. Ancillary data of the LANDSAT 7 ETM+ scene acquired in August 1999 for the radiance conversion showing the gain and offset values.**

Band number	1	2	3	4	5	7
<i>Gain</i>	0.786274	0.817255	0.639608	0.635294	0.128471	0.044439
<i>Offset</i>	-6.2	-6.0	-4.5	-4.5	-1.0	-0.35
$ESUN_\lambda$	1969.9	1840	1551	1044	225.7	82.07

*Source: data of  $ESUN_\lambda$  is obtained from Irish (2000).*

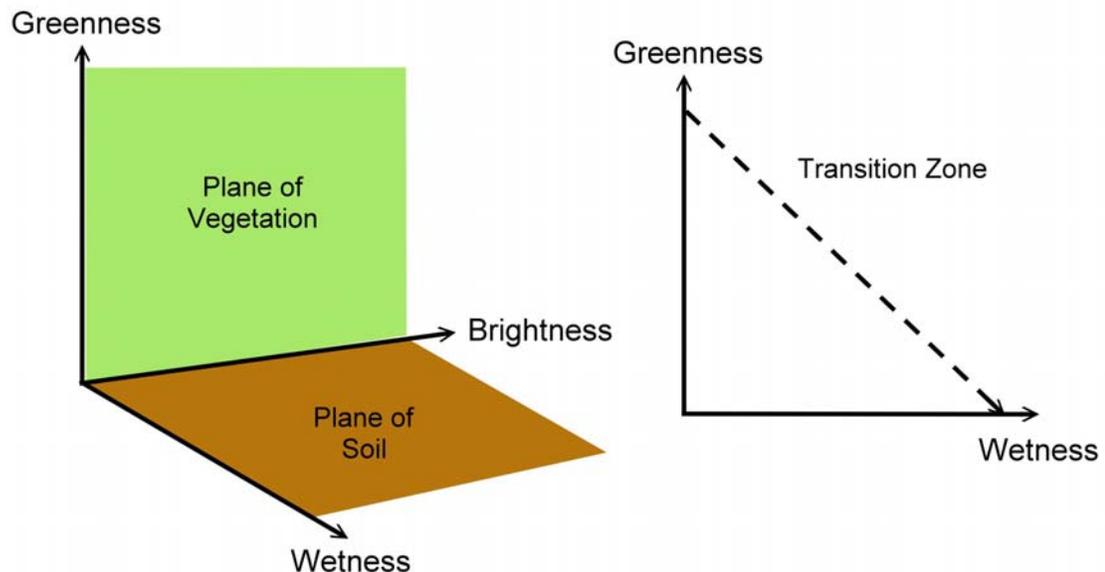
## 4.4 Tasseled Cap Transformation

### 4.4.1 Overview

The tasseled cap transformation in remote sensing is the conversion of the readings in a set of bands into composite values. The transformation linearly combines the readings in the multiple bands to a weighted sum according to the given coefficients. The composite weighted sums represent the tasseled cap features, that is, brightness, greenness, and wetness. The tasseled cap transformation is inspired by the method of the principal component analysis. The principal component analysis is often used to evaluate data dimensionality. It decomposes the data set into a new coordinate system with a new set of orthogonal axes and uses the first two principal components to define the plane into which the data are dispersed. However, the principal component analysis may fail to

define the actual planes into which the data are dispersed because the variation of data density in the other planes will change the results to a degree (Crist and Cicone, 1984).

By contrast, the tasseled cap transformation has a more analytical basis because it combines a generalization from empirical observations. Although used mainly for vegetation studies, tasseled-cap transformation can separate urban, water, and wetland classes (Jensen, 1996). Usually there are just three composite variables. Brightness, greenness, and wetness are the most important composite indices of a tasseled cap transformation. In a three-dimensional space, two perpendicular planes and a “transition zone” between the two define the feature space. The axes of brightness and greenness form a “vegetation plane” while the axes of brightness and wetness form a “soil plane”. Between the two planes are the data from partially vegetated plots where both vegetation and soil are visible. Usually, the transition zone is roughly filling out a right triangle. Figure 4-4 illustrates this relationship between the three tasseled cap indices in a feature space.



**Figure 4-4. Dispersion of the six-band Thematic Mapper data.**

The tasseled cap transformation was originally developed for understanding important phenomena of crop development in the spectral space (Kauth and Thomas 1976). However, with the information from the third dimension, i.e. the wetness feature, the distinction between forest vegetation and cultivated vegetation is enhanced. Figure 4-5 adapted from Crist *et al.* (1986) illustrates the approximate locations of the scene classes in the TM tasseled cap feature space. It gives the primary ideas about the types of land cover enclosed in the test areas when the data distributions are presented in the feature space. In the figure, the forest class is always distributed on the front of the “cap” while the water class is located in the corner of the “cap”. The tasseled cap transformation thus has potential in revealing key forest attributes including species, age and structure (Cohen *et al.* 1995) and in extracting water and wetland pixels (Civco and Hurd, 1999).

#### **4.4.2 At-satellite reflectance-based tasseled cap transformation**

Since the brightness feature highlights the areas of high reflectance, the greenness feature the areas that are vegetated, and the wetness feature the areas that have high canopy and soil moisture content, the wetland pixels can be extracted by using a tasseled cap transformed imagery. An at-satellite reflectance-based tasseled cap transformation compresses the Landsat 7 ETM+ multispectral data into a few bands associated with the physical scene characteristics. Huang *et al.* (2001) developed a new tasseled cap transformation based on at-satellite reflectance. They noted that their transformation was more appropriate for regional applications where atmospheric correction was not feasible. It also improves the ability to differentiate bright soil pixels from some dark green vegetation pixels. The tasseled cap features can be derived through linear combinations of the at-satellite reflectance coefficients as given in Table 4-2. Brightness is a partial sum of all bands; greenness describes the contrast between the near infrared bands and the visible bands; wetness depicts the contrast between the middle infrared bands that is sensitive to water and other bands.

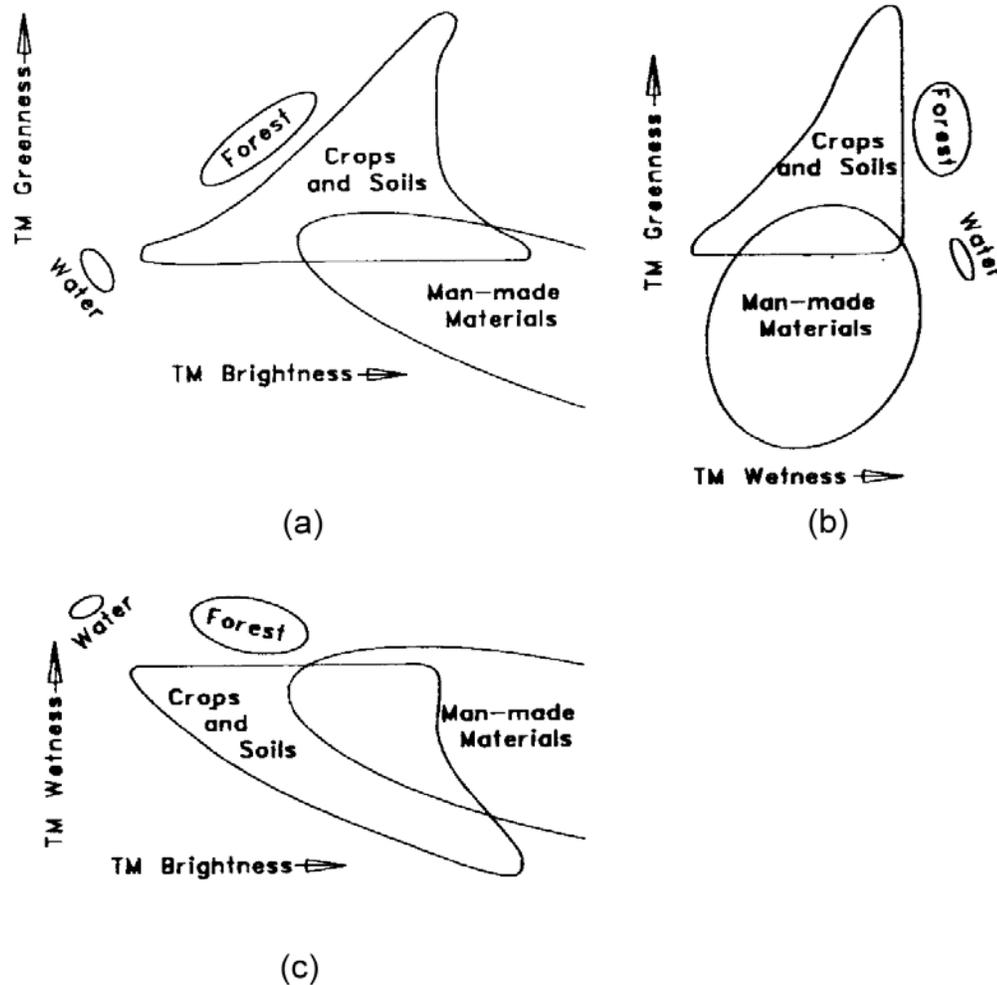


Figure 4-5. Approximate locations of important scene classes in TM tasseled cap feature space: (a) the plane of vegetation, (b) transition zone, and (c) the plane of soil. (From Crist *et al.*, 1986).

Table 4-2. Tasseled cap coefficients for Landsat 7 ETM+ at-satellite reflectance

Band number	1	2	3	4	5	7
Brightness	0.3561	0.3972	0.3904	0.6966	0.2286	0.1596
Greenness	-0.3344	-0.3544	-0.4556	0.6966	-0.0242	-0.2630
Wetness	0.2626	0.2141	0.0926	0.0656	-0.7629	-0.5388

Source: data is obtained from Huang *et al.* (2001).

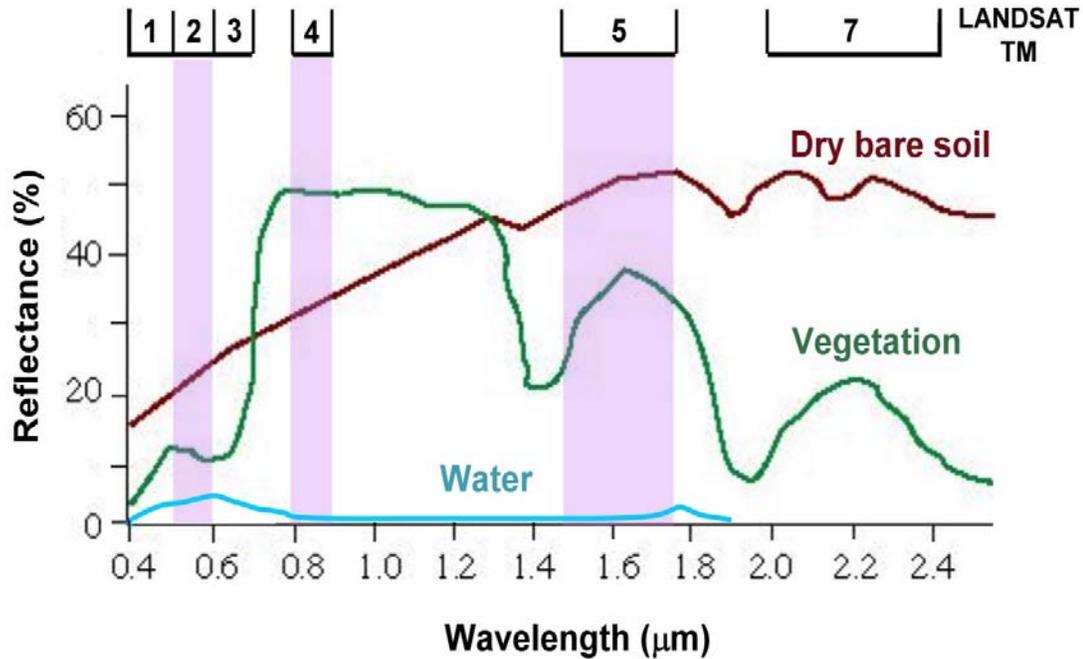
#### 4.5 Water Area Extraction

A band ratio-typed index was used to extract out the water pixels. Because water is a strong absorber in the near infrared and shows a higher reflectance in the middle infrared region of the Electro-Magnetic spectrum, the Normalized Difference Water Index (NDWI), derived from band 4 and band 5 reflectances, was proposed by Gao (1996). The index could be useful for discriminating water bodies from land. Anderson *et al.* (2004) found the index was least susceptible to saturation at high levels of leaf area index. The NDWI is a broad-channel ratio and can be defined as:

$$NDWI = \frac{\rho_{TM4} - \rho_{TM5}}{\rho_{TM4} + \rho_{TM5}} \quad (4-3)$$

where  $\rho_{TM4}$  and  $\rho_{TM5}$  are the reflectances of TM band 4 and band 5 respectively. The NDWI is far less sensitive to atmospheric scattering effects and more sensitive to vegetation water content than the NDVI. In this study the NDWI was used as an index to mask out water bodies rather than the NDVI. However, both water bodies and closure canopies illustrated high water content on the image.

To improve this problem of detecting pure water bodies, the NDWI could be adjusted to a specific range so that pure water pixels could be easily distinguished from vegetation features by setting a threshold to the adjusted water index. Figure 4-6 shows the typical spectral reflectances of water and vegetation. In the near infrared and the middle infrared regions water appears with a darker tone and vegetation with a lighter tone. Both the water and the vegetation features have positive NDWI values. It becomes difficult to manually decide a threshold to separate these two features. However, the trends of the spectral signatures of water and vegetation are significantly opposite in band 2 and band 5. For example, water shows a higher reflectance in band 2, whereas vegetation shows a higher reflectance in band 5. Because of this difference, water and vegetation give an opposite sign (i.e. one is in positive and the other is in negative) when the normalized difference ratio of band 2 to band 5 is applied.



**Figure 4-6. Typical spectral reflectance of common earth features: water, vegetation, and soil.**

The ratio is known as the Normalized Difference Snow Index (NDSI), derived from the green band (TM2) and the middle infrared band (TM5) reflectances. It can be expressed as:

$$NDSI = \frac{\rho_{TM2} - \rho_{TM5}}{\rho_{TM2} + \rho_{TM5}} \quad (4-4)$$

Vegetation has negative NDSI values due to higher reflectance in TM band 5 than in TM band 2, whereas water has positive NDSI values. Because of this physical characteristic, water bodies can be easily distinguished from vegetation when the NDWI is multiplied by the NDSI. As given in Eq. 4-5, zero can be set as a threshold to extract water pixels out from the original data set. When the pixel has a multiplied value of two indices higher than zero, it should be defined as a water pixel and excluded from the original data set. In

contrast, if the pixel has a multiplied value lower than zero, the pixel should be included in the fuzzy input data set. After the water pixels are extracted out from the original data, the remaining data set will be used as the inputs of the two fuzzy clustering algorithms.

$$x_i \in \begin{cases} X_{fuzzy\ input}^* & \text{if } (NDWI \times NDSI) \leq 0, \forall x_i \\ X_{water} & \text{if } (NDWI \times NDSI) > 0, \forall x_i \end{cases} \quad (4-5)$$

#### 4.6 Training Site Selection

In this thesis study semivariograms were assumed to represent the landscape structure pattern of land cover types. In the partial supervised clustering algorithm, training site selection is crucial to the guidance of the clustering procedure. Before finally selecting the training sites, several training sites should be evaluated to find out the representative site. Except water class, three different classes of land cover were examined in this study: deciduous forest, wetland, and mixed stand. The information of these three land cover classes is available in the BOREAS project done over the area (ORNL DAAC, 2001). The forest inventory provides land cover information for selecting the training sites of these three land cover types. The land cover classes used in this study are defined as follows: (1) deciduous forest is comprised of areas covered by aspen groves; (2) wetland is comprised of areas covered by wet conifers such as black spruce in the poor-drained area; (3) mixed stand is comprised of areas covered by deciduous forest stands and dry conifers.

Our preliminary examination of this study has evaluated the applicability of utilizing texture features in hard classification for wetland mapping (Chiu and Couloigner, 2004a, see Appendix A) and the suitable window sizes for deriving geospatial texture features to extract wetland patches (Chiu and Couloigner, 2004b). After examining the semivariogram behaviors and comparing the classification accuracy obtained, the study concluded that an arbitrary window of  $7 \times 7$  pixels was suitable for deriving texture features for wetland mapping and that a combination of texture features derived from

---

multiscale window sizes could also help to improve the classification accuracy. The explanation was that a smaller window size highlights edges between different land cover classes and a larger window size characterizes the texture properties of land cover classes.

In this study since semivariograms are employed as landscape structure pattern in the classification, the idea inspired from the preliminary examination of using multiscale window sizes for texture features is to extend an arbitrary window size to a series of sizes.

In order to select the training sites for deriving the semivariogram structure patterns for each land cover class, several sites have been examined until the representative one has been found. To derive the semivariogram patterns, training sites of the different land cover categories were randomly selected from the area of interest instead of from the test areas used in the experiments. One of the reasons is that the test areas are too small (i.e. 200×200 pixels) to find out representative training sites for every class. Another reason is that using the same training sites as guidance in the clustering iterations for two different test areas gives an unbiased result. Figure 4-7 illustrates the locations and subscenes of the selected training sites. The subscenes show the differences in the morphological features and the texture appearance. The water class is also included as one of the training sites because of the intention to compare differences in the semivariogram patterns of different types of land cover.

By extending a moving window size from a lag distance of 1 to 23 pixels, the algorithm derived a series of omnidirectional semivariogram texture images. The maximum window size was restricted to 23×23 pixels because of the concern around computation time. The geospatial variations of the data could be counted on using multiple window sizes simultaneously rather than just an arbitrary window size. The omnidirectional semivariograms that consider the means of eight directions were computed in this study because of unnoticeable anisotropy and greater computation time for deriving directional semivariograms.

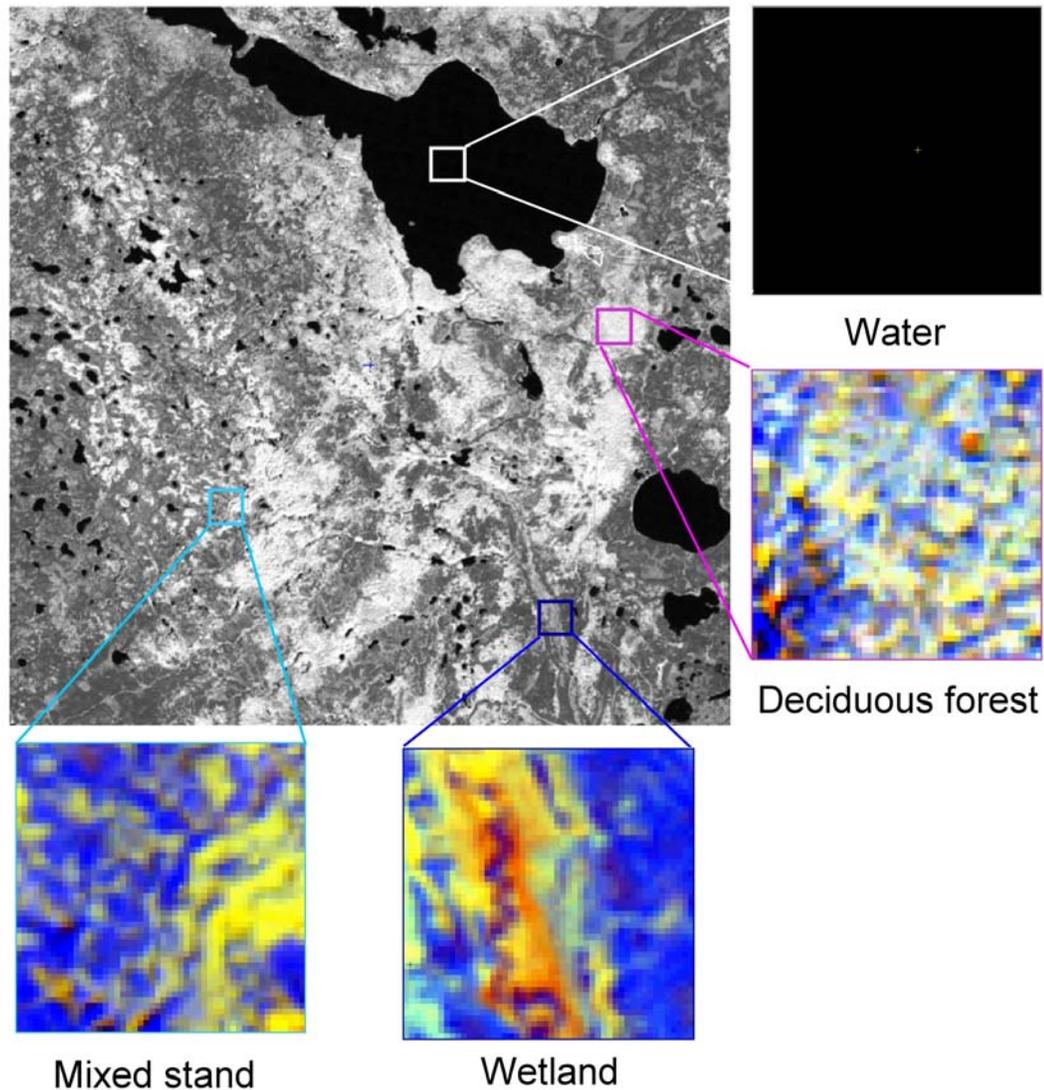


Figure 4-7. Training sites for deriving semivariograms.

#### 4.7 Implementation of Classification Algorithms

Both the standard FCM classification algorithm and the new developed SGFCM classification algorithm have been implemented into MATLAB. After the image pre-processing, a set of input vectors having  $N$  feature points was given. According to the theory described in Chapter 3, minimization of the objective function is based on the

suitable selection of a partition matrix  $U$  and a cluster centre  $V$  using an iterative process described by the following steps.

- (1) Determining values for the number of clusters  $C$ , fuzziness index  $m$ , and the convergence error  $\varepsilon$ .
- (2) Creating an initial  $C \times N$  partition matrix  $U$  of the input data set randomly.
- (3) Calculating cluster center of the clusters using Eq. 3-24 and the fuzzy covariance matrix using Eq. 3-25.
- (4) Obtaining the distance measure according to Eq. 3-14.
- (5) Updating the partition matrix  $U$  using Eq. 3-22 until Eq. 3-27 is true. Otherwise, repeating step (3) to (5) until convergence.
- (6) Performing a final defuzzification by assigning the image data to the cluster with
  - (a) the highest membership value for the FCM classification
  - (b) the alpha-cuts rule for the SGFCM classification.

In this study, the three tasseled cap features - brightness, greenness, and wetness - were used as the fuzzy inputs. The default values, including the fuzzy index, the number of classes, and the convergence error, used in the standard FCM clustering algorithm and the semivariogram guided FCM (SGFCM) clustering algorithm were identical, i.e.  $m = 2$ ,  $C = 3$  and  $\varepsilon = 10^{-5}$ . The standard FCM clustering algorithm belongs to a purely unsupervised classifier and used the maximum function in the defuzzification. The developed SGFCM clustering algorithm is defined as a partial supervised classifier. The alpha-cuts approach was applied in the defuzzification. In SGFCM, three semivariogram training features were used to guide the fuzzy classifier: deciduous forest, wetland, and mixed stand. Water was excluded because of its specific reflectance responses shown in the preliminary analysis of the semivariogram behaviors.

## 4.8 Summary

This chapter has given a description of the study site, the area within the boundary of Prince Albert National Park, including its topographic status and commonly distributed

flora. The available image data was a Landsat 7 ETM+ scene acquired in August 1999. Image pre-processing has been done through radiance and reflectance conversion and the tasseled cap transformation has compressed the six Landsat bands of interest into three spectral features: brightness, greenness, and wetness. These three tasseled cap features can represent the physical characteristics of wetland, such as soil, vegetation, and hydrology status, respectively. Because pure water bodies were easily distinguished, this land cover class has been excluded from the original data set and the remained data set was prepared for the fuzzy clustering procedure. The idea of how to extract water pixels by using an adjusted Normalized Difference Water Index (NDWI) has also been introduced. The selection of the training sites and the preparation of their semivariogram texture features have been introduced. Finally, the steps of the classification algorithms implemented in MATLAB have been described in this chapter.

## CHAPTER 5

# EXPERIMENTAL RESULTS AND DISCUSSIONS

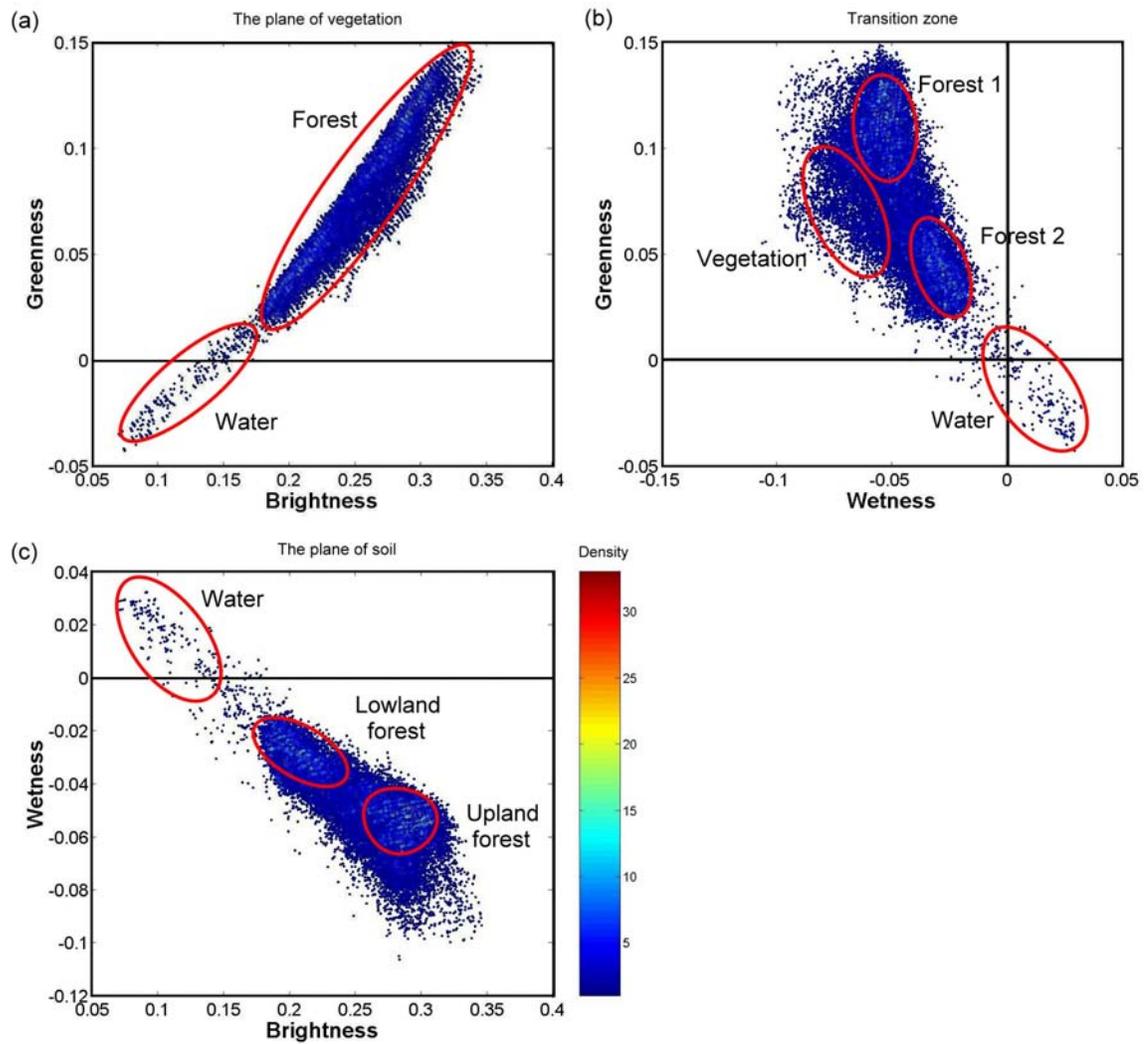
*“We must not forget that when radium was discovered no one knew that it would prove useful in hospitals. The work was one of pure science. And this is a proof that scientific work must not be considered from the point of view of the direct usefulness of it. It must be done for itself, for the beauty of science, and then there is always the chance that a scientific discovery may become like the radium a benefit for humanity.”*

*Marie Curie, 1867-1934*

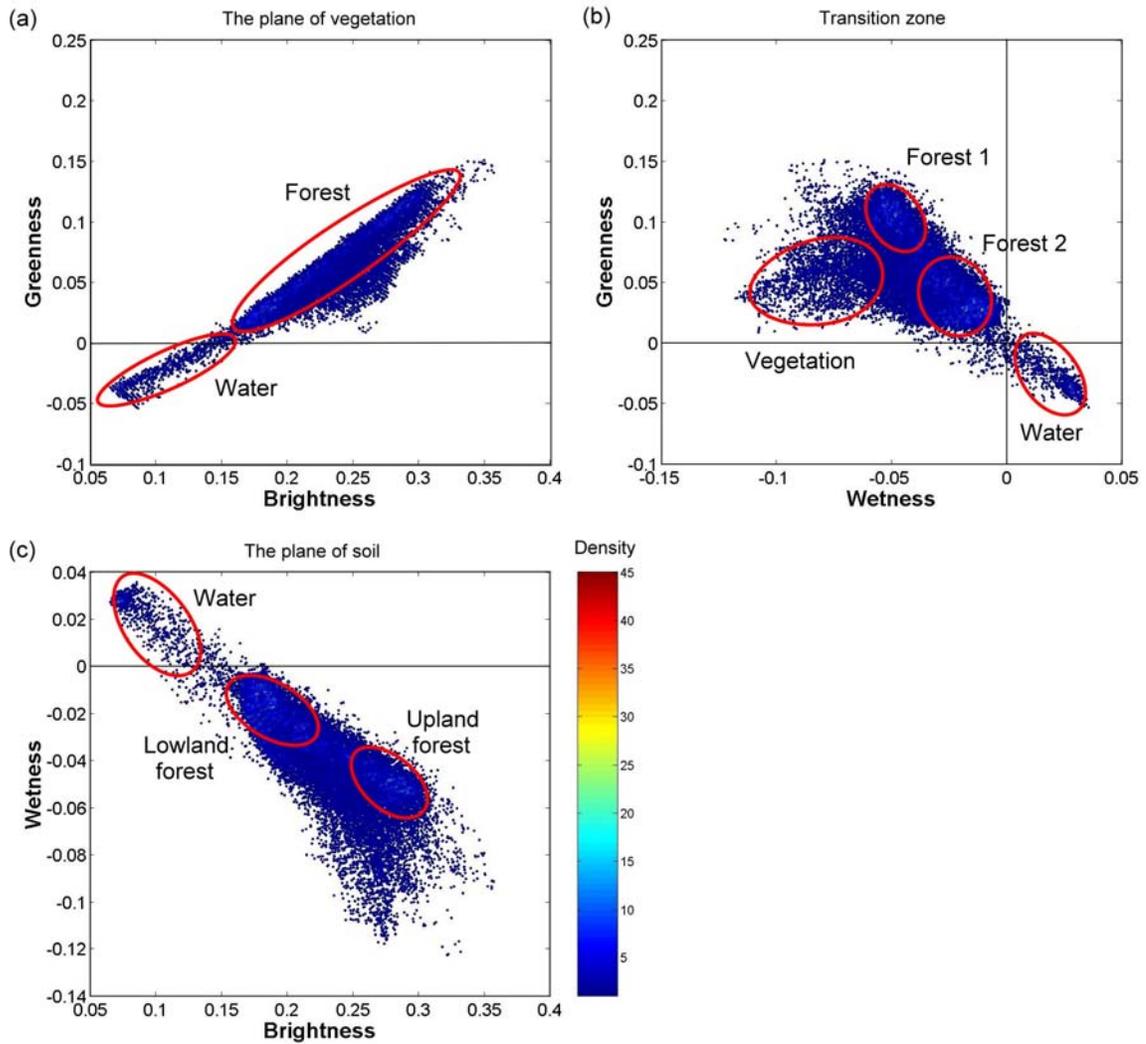
*Polish Scientist and physicist*

### 5.1 Examination of Tasseled Cap Features

The original Landsat TM image data sets of both subscenes have been converted to tasseled cap features based on at-satellite reflectance. These image data sets were plotted into the tasseled cap feature space to demonstrate the data distributions. Figure 5-1 and Figure 5-2 show the scatter plots of the data sets for test area A and test area B on the three TM tasseled cap feature planes. The color legend represents the density of the data points. If the occurrence of the data composition in the feature space is high, the data point with that composition has higher data density and is displayed in dark red. By contrast, the lower occurrence means lower data density and is displayed in dark blue. Compared to the study result of Crist *et al.* (1986) shown in Figure 4-5, “our” scatter plots give some initial cluster information according to the approximate locations of scene classes in the TM tasseled cap feature space.



**Figure 5-1. Planar dispersion of the test area A data set in TM tasseled cap feature space: (a) The plane of vegetation, (b) the transition zone, and (c) the plane of soil. The color legend represents the data density of the data points.**



**Figure 5-2. Planar dispersion of the test area B data set in TM tasseled cap feature space: (a) The plane of vegetation, (b) the transition zone, and (c) the plane of soil. The color legend represents the data density of the data points.**

### **5.1.1 The plane of vegetation**

Figure 5-1(a) and Figure 5-2(a) illustrate the planar data dispersions of the two data sets in the plane of vegetation. In both figures, the data points seem to be separated into two major groups. Both figures indicate that these two major clusters should be water and forest according to their locations in the TM tasseled cap feature space. The data points of the water cluster have the greenness values lower than zero and the brightness values close to zero, while the data points of the forest cluster have high greenness and brightness values. The data density of the forest cluster is much higher than the water cluster, which means that most of the data points in the data set belong to vegetation canopies. Even though forest is the major type of land cover in the subscene, the data dispersion illustrates that two or three subclasses emerged in the forest cluster according to the different density levels displayed. However, the subclasses are not easily distinguished because of the compactly elongated data distribution in the plane of vegetation. The other TM tasseled cap features should provide helpful information to separate these emerged subclasses.

### **5.1.2 The plane of transition zone**

The transition zone provides additional information and results by providing a clearer separation of compact clusters. This distinction can be seen particularly in color composite imagery. Partially vegetated data points mainly occupy the transition zone. Figure 5-1(b) and Figure 5-2(b) show the planar dispersion of the data points for both test areas in the TM tasseled cap transition zone. When the tasseled cap wetness index is added, the dispersion of the data points in the transition zone is not as compact as it is in the plane of vegetation. It spreads out in the transition zone and makes the vegetation structure to be more easily detected from that plane.

In both Figure 5-1(b) and Figure 5-2(b), the forest cluster is separated into at least two subclasses: forest and other vegetation cover. The forest class can be furthermore partitioned into two more subclasses, forest 1 and forest 2, according to the data density. The forest classes have higher wetness values than the other type of vegetation canopy.

However, greenness does not present such trends; it varies between different vegetation structures. Crist *et al.* (1986) has given the most probable explanation for the tasseled cap wetness difference, “as compared to cultivated crop or grass canopies, the relatively dense forest stands contains a higher percentage of opaque stems and thus increasing the incidence of deep shadows both on the lower layers of the canopy and on the leaves in the tree crowns.” Their hypothesis that increased shadowing should cause increased signal in the wetness feature has found support on theoretical grounds as well as in simulated and actual data. Thus, the forest structure has been at least one of the important factors determining the wetness value.

### 5.1.3 The plane of soil

The plane of soil is composed of the brightness and wetness projection. The brightness direction is correlated to the texture and the moisture content of the soil while the wetness direction is heavily weighted by the middle infrared reflectance that is sensitive to soil moisture. In the plane of soil, both the brightness and the wetness features show their spectral response to the soil moisture content. Cohen *et al.* (1995) noted that the tasseled cap wetness was essentially unaffected by the topographic variations in closed conifer stands. Wetness in the plane of soil is thus more powerful than brightness and greenness for distinguishing upland from lowland forest type.

Figure 5-1(c) and Figure 5-2(c) present the scatter plots of the two test areas in the plane of soil. Two subclasses of the forest cluster are highlighted in the plane of soil: upland forest and lowland forest. The cluster of upland forest shows lower wetness values while the cluster of lowland forest has wetness values close to zero. The cluster of upland forest has higher brightness values than the cluster of lowland forest. Usually deciduous stands have higher brightness and greenness values than coniferous stands due to the structure of their leaves. The cluster of upland forest has then a high probability to be defined as deciduous forest stand. The tasseled cap transformation can make a direct association between the feature response and the physical characteristics of the scene classes (Crist *et al.*, 1986). By capturing the majority of data variation in the TM tasseled cap features, at

least four land cover clusters can be approximately defined from the scatter plots: water, dense upland forest (i.e. deciduous forest), dense lowland forest, and less dense canopy.

## 5.2 Preliminary Analysis of Semivariogram Behaviors

From the previous examination of the tasseled cap enhanced images, at least four clusters - water, upland forest, lowland forest, and other vegetation canopy - have been found in both scenes. Compared to the reference data for the definition of land cover types in the study area, the cluster of lowland forest is equivalent to the type of mixed stands that include deciduous forest and dry coniferous stands, and the cluster of other vegetation canopy exactly represents the type of wetland that is covered by the wet coniferous stands in the poor-drained area. Therefore, the four categories of land cover classes are now defined as water, deciduous forest, wetland, and mixed stand.

From a textural point of view, land cover types can demonstrate their landscape structure through semivariogram patterns. In the preliminary analysis of the semivariogram behaviors, the semivariogram texture features are derived from the TM tasseled cap features. Figure 5-4 shows the examples of the semivariogram patterns computed from the TM tasseled cap brightness, greenness, and wetness, respectively. For the selected training sites (Figure 4-7) corresponding to the four land cover classes, Figure 5-3(a), Figure 5-3(b), and Figure 5-3(c) represent the omnidirectional semivariograms computed from the three tasseled cap features respectively to illustrate the relationship between the image texture and the semivariogram. The semivariograms computed for each land cover class are unique and have the following characteristics:

- (1) Water: water body has very low pixel values in the three tasseled cap features and shows almost no variation between the pixels and their neighborhoods when a visual inspection is done. The semivariograms of water body are essentially flat in Figure 5-3(a) to (c). The curve exhibits little or no spatial correlation for lag distance greater than 1 pixel.

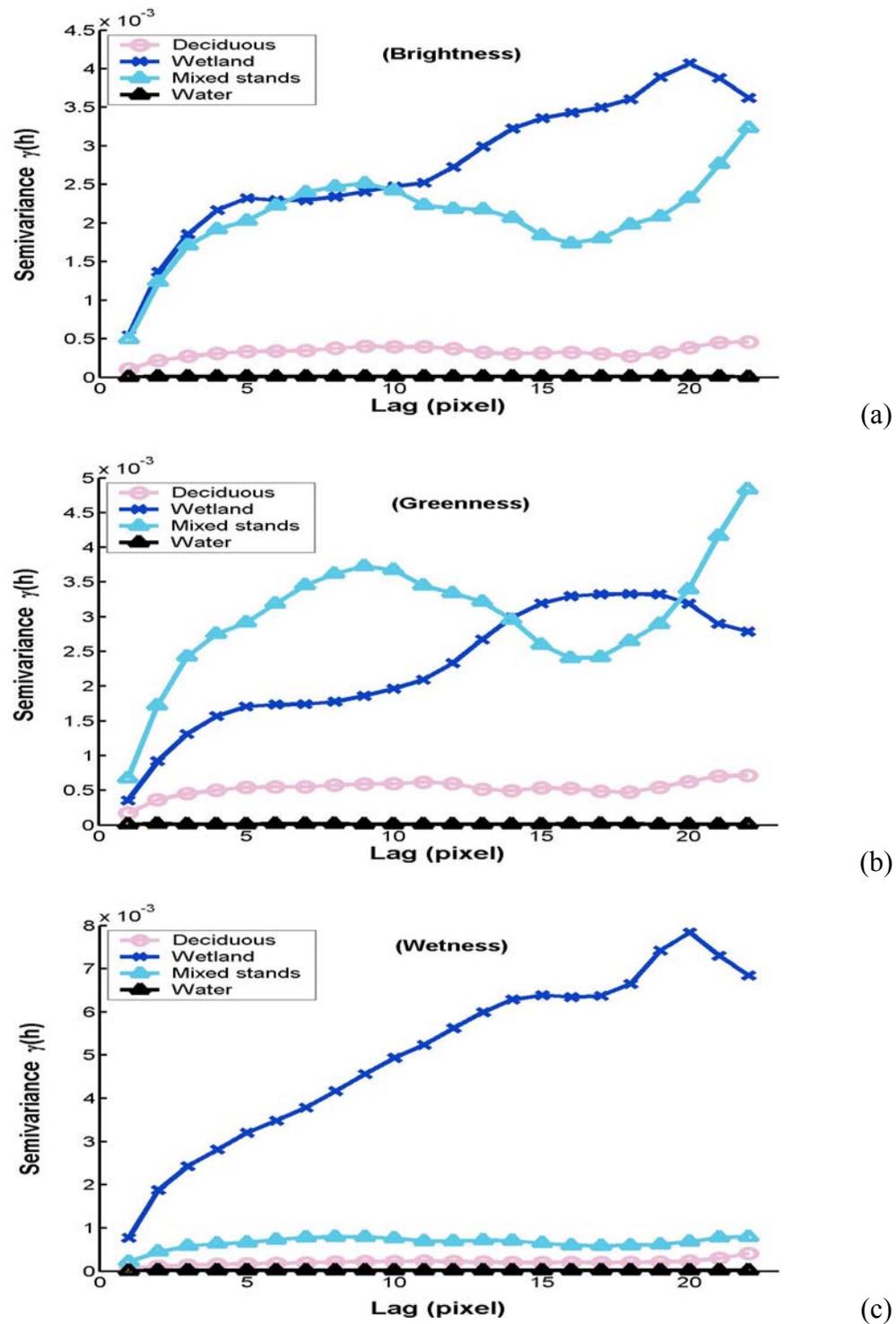


Figure 5-3. Omnidirectional semivariograms derived from the TM tasseled cap features: (a) brightness, (b) greenness, and (c) wetness for the four classes presented in the training sites (Figure 4-7).

- 
- (2) Deciduous forest: deciduous forest shows many bright forest stands in the color composite image. These brighter forest stands represent higher values in both the tasseled cap brightness and greenness. However, the forest structure such as stand density demonstrates somewhat degree variations. The semivariograms of deciduous forest class exhibit smooth curves that reach the limiting value (i.e. the sill) at a lag of 7 pixels. The semivariogram of this class has greater variances in the brightness and the greenness features as compared to the variances in the wetness feature. This finding indicates that the moisture content within the deciduous forest is much even. However, the difference between deciduous forest and mixed stand is not significant enough in the wetness feature when only the wetness information is used for class separation.
  - (3) Wetland: wetland does not show the homogeneity as the deciduous forest class and the water class do. Especially, it shows the greatest variances among the four training classes in the study area. Wetland is associated with the largest semivariance in the tasseled cap wetness feature. The semivariograms of wetland in both the brightness and the greenness features show a wave shape, while reaching first limiting values at a lag of 5 pixels and second limiting values at a lag of 15 pixels. This wave shape indicates that the landscape structure of wetland, such as vegetation compositions, has some changes occurring between the lag distances of 5 to 15 pixels. In contrast to the semivariogram shown in Figure 5-3(a) and 5-3(b), the semivariogram rises steadily upward, approaching the sill only at lag distances greater than 15 pixels. This trend indicates that the moisture status of the wetland class varies in a large areal extent: the transitional characteristics of wetland are spectrally reflected on the semivariogram.
  - (4) Mixed stand: mixed stand shows much greater semivariances in brightness and greenness than the deciduous forest class. Although these two classes both belong to dense canopies in the tasseled cap feature space, mixed stand has the largest variance of the semivariogram in greenness. The likely explanation of the difference in greenness is that mixed stand is composed of both deciduous and coniferous stands. The semivariograms rise smoothly upward up to the sill at a lag distance of 9 pixels,

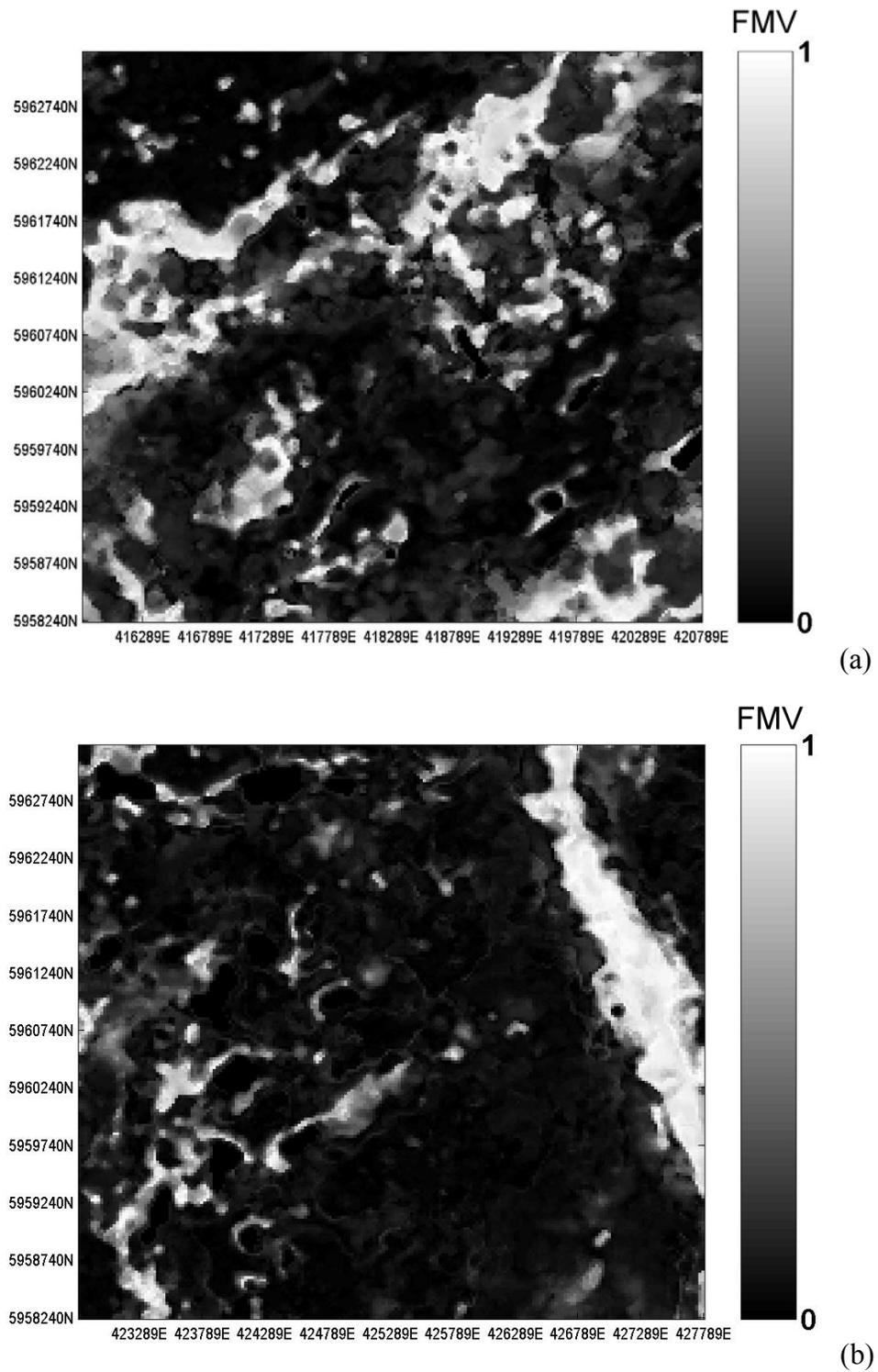
and rise upward again after curving downward to a lag distance of 16 pixels. This semivariogram behavior is similar to the behavior of planted coniferous woodland in the study of Curran (1988).

### **5.3 Fuzzy Classification of Land Cover Types**

#### **5.3.1 Analysis of the fuzzy membership values**

The partition matrix of a data set is obtained after the objective function converged in the fuzzy clustering iterations. The partition matrix gives each pixel a fuzzy membership value to the associated class. In this study, both the FCM and the SGFCM classifiers produce a set of three fuzzy membership values (FMVs) for three land cover class: deciduous forest, wetland, and mixed stand. These comparable fuzzy membership values can indicate the “likelihood” degrees to which land cover type the pixel belongs. One example of the fuzzy membership image is shown in Figure 5-4.

Figure 5-4(a) and 5-4(b) show the fuzzy membership images of the wetland class for test area A and test area B respectively. In these figures, the brighter the pixel is, the higher the value of its membership to wetland. Here, the fuzzy membership values of the wetland class are generated from the SGFCM clustering algorithm. In the same way, the fuzzy classifier gives to the other two classes, deciduous forest and mixed stand, a membership value at the same time. After examining the fuzzy membership values of the wetland class (Figure 5-4), the following patterns are found: (1) areas along the lakeshores or the riparian areas have higher membership values in the wetland class; (2) pixels nearby the water bodies have higher membership values in the wetland class. These patterns emphasize that the occurrence of wetlands is accompanied with ponding waters. This is satisfactory with the definition of wetlands, which states that wetland is referred to a transition zone between uplands and aquatic areas.



**Figure 5-4. Images of fuzzy membership values (FMVs) of the wetland class: (a) test area A and (b) test area B.**

---

Ancillary data such as digital elevation model (DEM) data can also help to clarify the findings. Figure 5-5 shows the 25m DEM from 1:50,000 scale maps of test area A and test area B. The brighter areas represent uplands while the darker areas depict lowlands. When a fuzzy membership map is compared to the DEM data, the pixels with the higher membership values in the wetland class are located in the lowlands while the pixels with the lower membership are in the uplands. As mentioned above, wetland represents poorly drained areas covered by wet coniferous stands in the area of interest. While deciduous and coniferous stands are mixed in some area, it may result in the increase of wetland membership value for a particular pixel. However, these pixels may be misclassified in the defuzzification procedure.

Parameters such as slope, curvature, and flow direction, derived from DEM data are helpful in wetland mapping, because the hydrology condition, which is one of the most important characteristics in wetland definition, is directly related to these topographic information. However, in this study DEM data was not included in the classification procedure. The main reason is that the 25m resolution of the DEM data is too coarse to present the variations of the topographic parameters. The parameters did not show distinguished variations in the area based on the 25m resolution DEM data. However, it is possible to involve DEM data in the post-classification process to adjust misclassified pixels. This would be one of the scopes for further study.

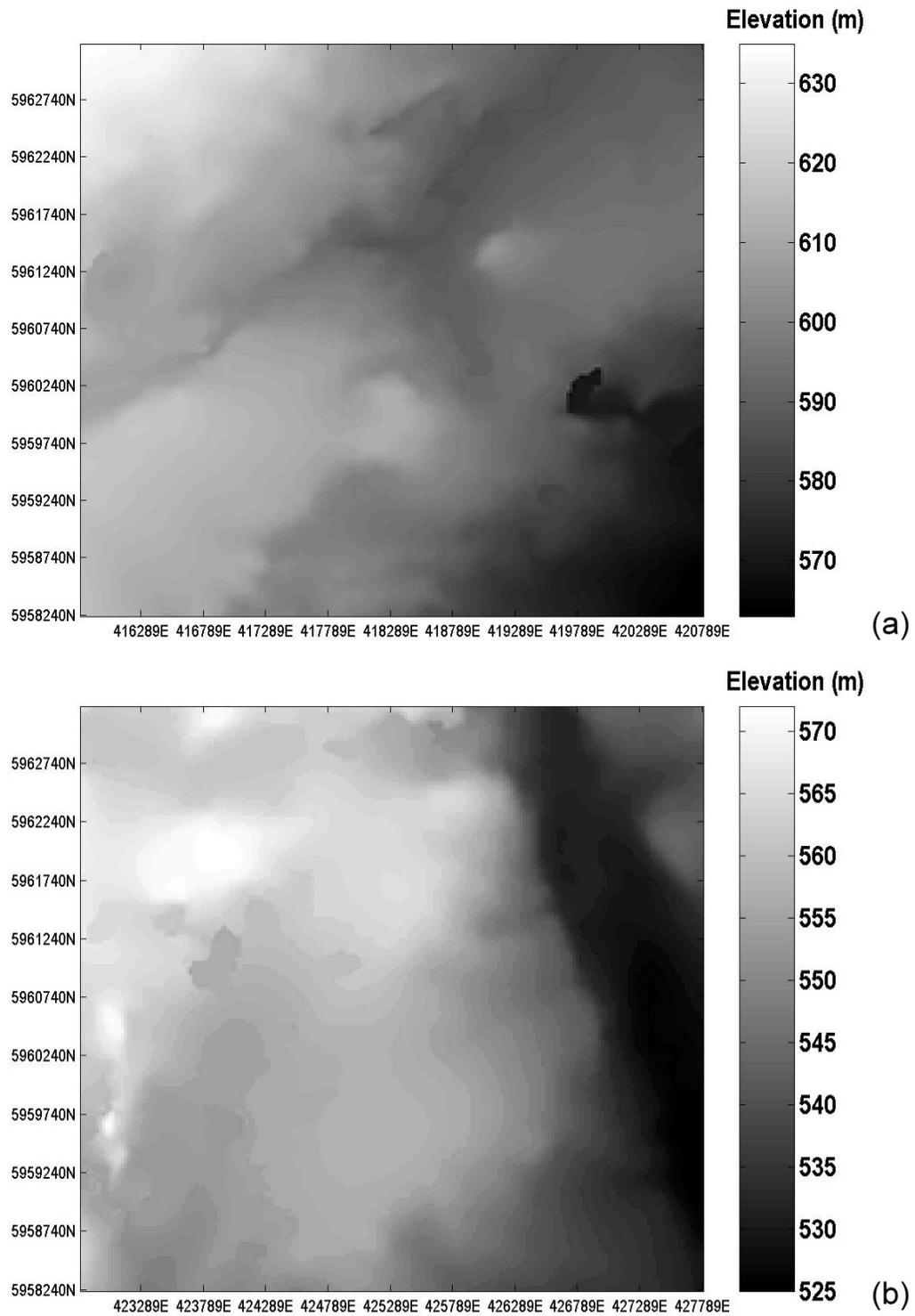


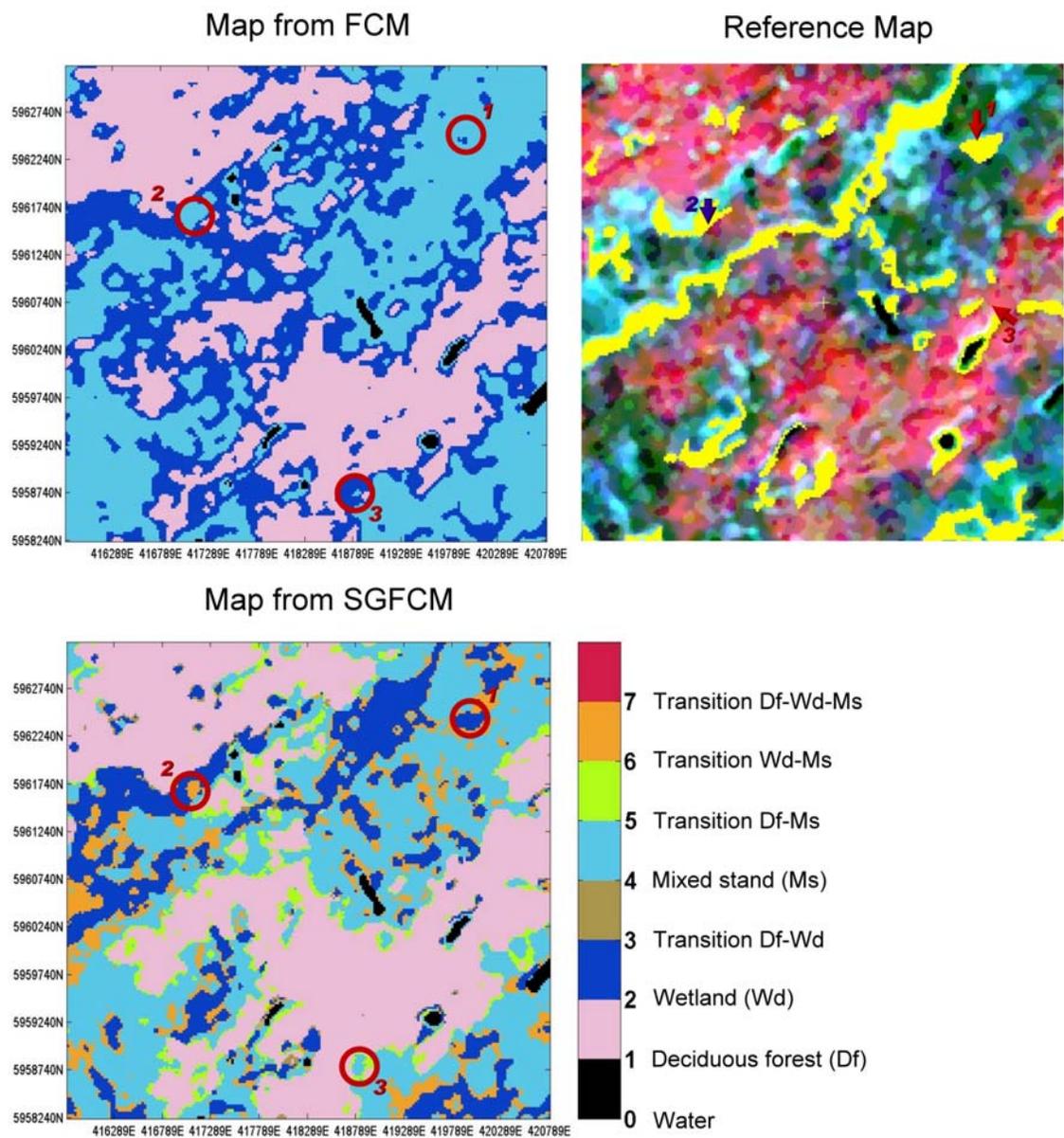
Figure 5-5. Digital Elevation Model (DEM) of test areas (a) A and (b) B.

### 5.3.2 Visual evaluation of the fuzzy classification

After the fuzzy classifier has given a set of membership values to each pixel for each class, all these fuzzy membership vectors were stacked one on the other to create a partition matrix. The defuzzification was applied to the partition matrix for classification. Pixel labeling has also been done at the same time to identify the classes.

In this study, two defuzzification approaches were employed in the FCM and the SGFCM classifiers. The maximum function was applied to the partition matrix generated from the standard FCM classifier. Thus the one of the three land cover classes having the highest membership values determined the class type of the pixel disregarding the differences of membership values between the classes. In the defuzzification, number codes were used to label the pixels to which class they belong. For example, “0” represents the class “Water”; “1” the class “Deciduous forest”; “2” the class “Wetland”, and “4” the class “Mixed stand”.

In contrast to the maximum function, the alpha ( $\alpha$ )-cuts method was applied to the partition matrix generated from the SGFCM classifier. The selection of  $\alpha$  was based on the number of predefined classes. Because the significant differences between the membership values of the three classes were taken into account in the defuzzification, pixels with membership values that were not significant enough to assign those pixels to one of the three classes were rejected to transition classes. Therefore new classes were created in the defuzzification procedure to allocate these in-between pixels. To label the pixels belonging to these transition classes, codes were added to represent a new class. For example, the transition class of “Deciduous forest” and “Wetland”, called “Transition Df-Wd”, was coded as “3”; the transition class of “Deciduous forest” and “Mixed stand”, i.e. “Transition Df-Ms”, as “5”; “Transition Wd-Ms”, the transition of “Wetland” and “Mixed stand”, as “6”. When pixels represented a mix of all three classes, they were coded as “7”. However, the occurrence of this last transition class is rare: it only happened when a pixel has the same membership value for each of the three classes.

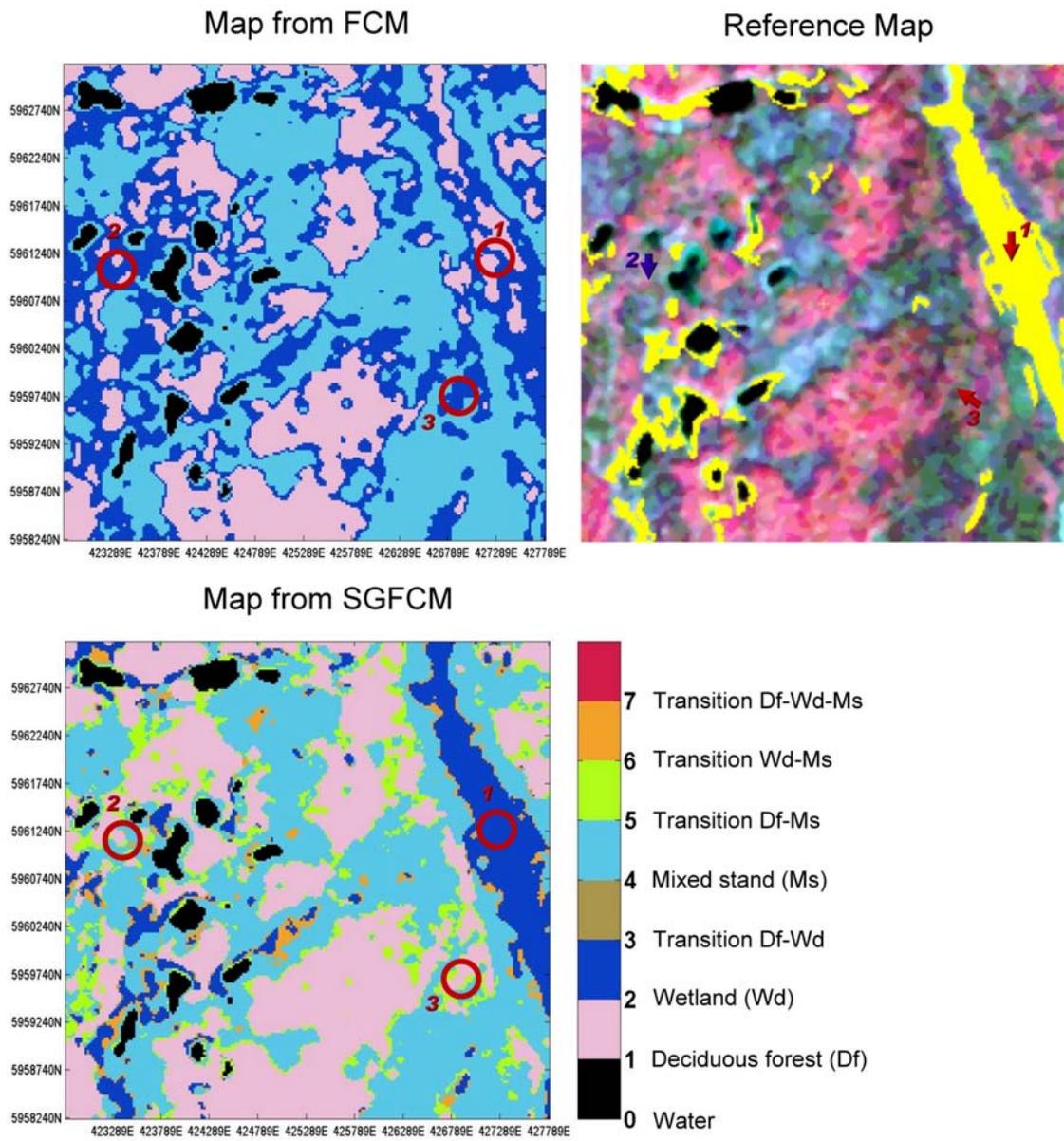


**Figure 5-6.** Reference data and two classification maps for test area A. Yellow mask obtained from the National Topographic Data-Base (NTDB) Canada, represents the wetland areas.

Figure 5-6 illustrates the mapping results for test area A. In this figure, the classification maps of the two fuzzy classifiers, the FCM classifier and the SGFCM classifier, are presented for comparison. The wetland data retrieved from the National Topography Data Base (NTDB) of Canada is also illustrated in the figure as reference. The yellow mask shown in the reference map indicates the wetland areas. The color legend shows the predefined classes. The class illustrated in dark blue is the class of interest in wetland mapping.

A visual evaluation is first used to examine the performance of the two classifiers in wetland mapping. To emphasize the differences found in the two classification maps, three areas are marked as examples. For the area noted 1, a wetland patch is shown in the reference map. However, the standard FCM classifier can detect only few pixels belonging to that wetland while the SGFCM classifier can extract the whole wetland patch. For the area noted 2, another wetland patch is masked in the reference map. Compared with the reference data, this patch can be correctly classified by neither the FCM classifier nor the SGFCM classifier. The FCM classifier misclassifies this patch as the class of “Mixed stand”, while the SGFCM classifier assigns the pixels within the patch as the transition class of “Wetland” and “Mixed stand”. Although the SGFCM classifier does not label the patch as the right class, it has the ability to denote the patch as a wetland related transition class. For the area noted 3, no wetland is indicated in the reference data. The SGFCM classifier, in contrast to the FCM classifier that misclassifies the area as a “Wetland” class, does not have misclassified pixels regarding to the wetland class. Although both fuzzy classifiers have problems with misclassification, the SGFCM classifier demonstrates a better performance than the FCM classifier.

Figure 5-7 illustrates the mapping result for test area B. Three areas are marked for a visual evaluation of the performances of the two classifiers. For the area noted 1, a large riparian area is shown in the reference data as wetland. The FCM classifier cannot detect the wetland area as a continuous landscape element, but it classifies the area as fragmental patches of deciduous forest and wetland.



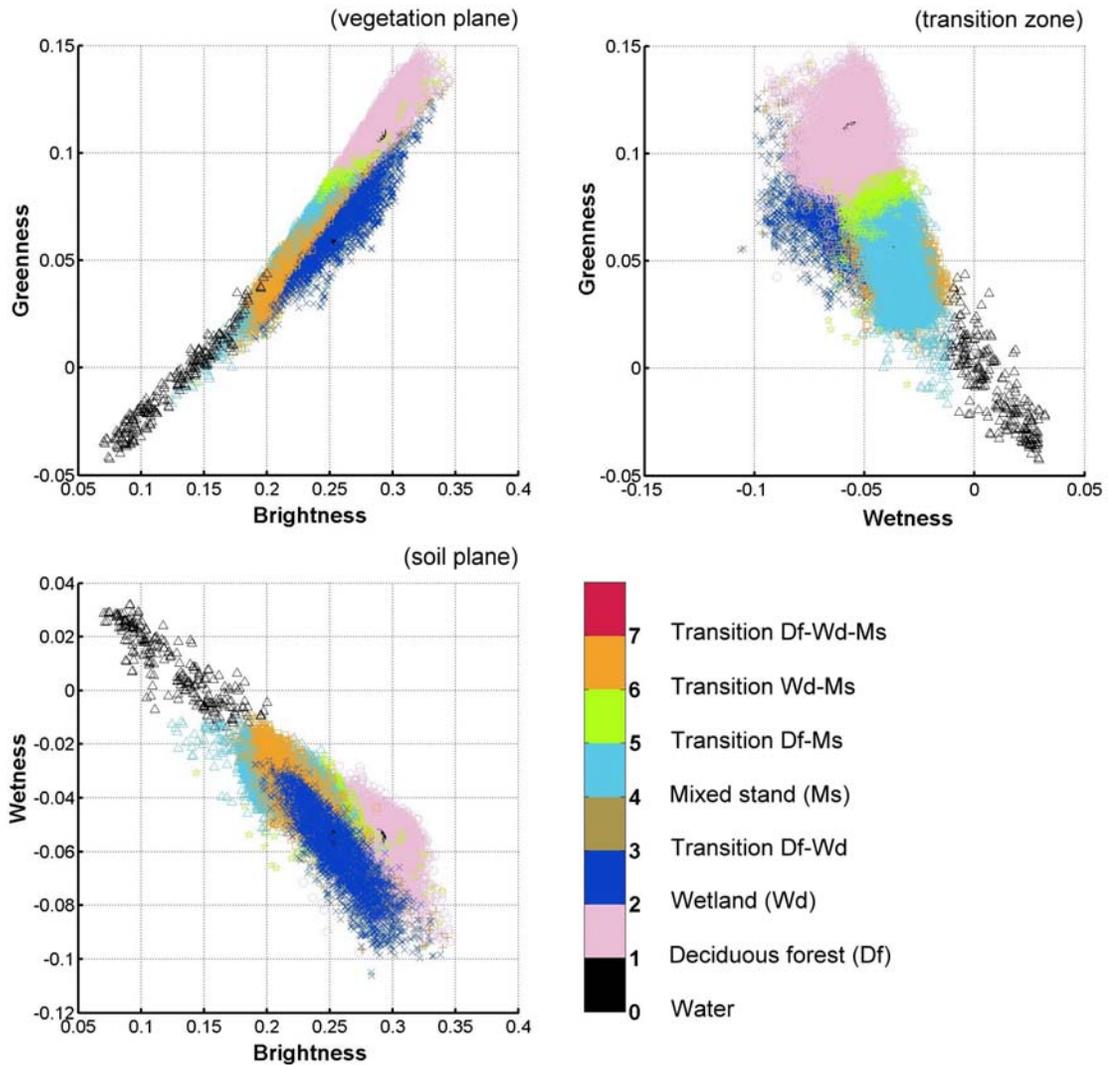
**Figure 5-7. Reference data and two classification maps for test area B. Yellow mask obtained from the National Topographic Data-Base (NTDB) Canada, represents the wetland areas.**

Meanwhile, the SGFCM classifier shows a significantly different mapping result. It extracts almost the whole patch of the riparian area and classifies the area as wetland. When compared to the reference data, the extent of the wetland class in this noted area is coincident with the area covered by the yellow mask. For the areas noted 2 and 3, the FCM classifier misclassifies these two areas as wetlands, whereas such misclassification does not happen on the classification map given by the SGFCM classifier. This visual evaluation again proves that the SGFCM classifier provides a better performance on wetland mapping than the FCM classifier.

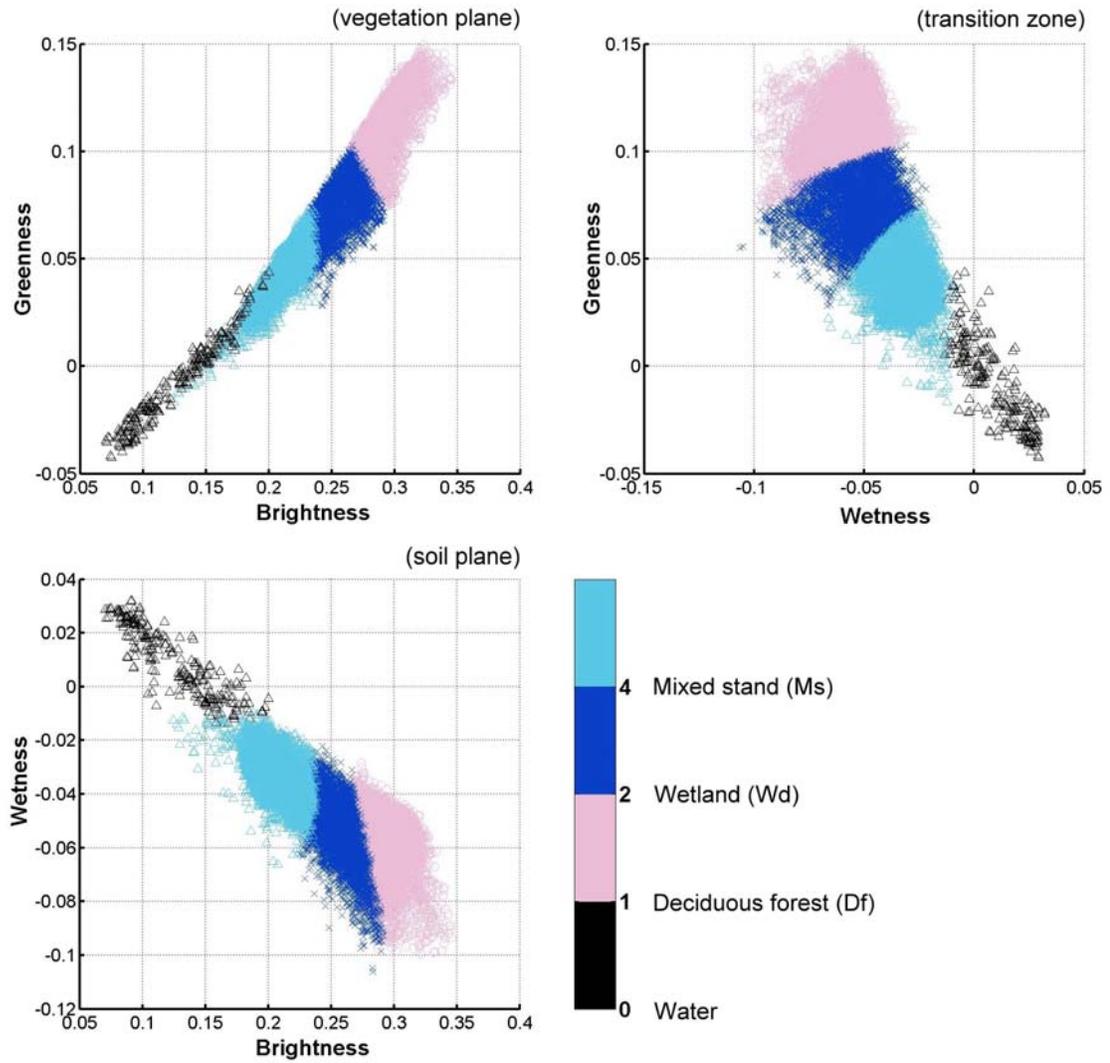
### **5.3.3 Analysis of the fuzzy class dispersion**

To further understand the reasons that result in the qualitative differences between the performances of the two fuzzy classifiers, the class dispersion in the TM tasseled cap feature space can give more information. Figure 5-8 illustrates the class dispersions of the seven classes given by the SGFCM classifier for test area A. The class dispersions are displayed in three tasseled cap planes. The major classes are distributed in elongated shapes and the transition classes lying between the classes are highlighted. In the TM tasseled cap feature space, the class of deciduous forest is the one that has the higher brightness and greenness values while the water class is located in the corner of the “cap”. This classification result is satisfactory in respect to the class properties. It fits with the expectations provided by the preliminary analysis of the tasseled cap transformation regarding these two classes. However, while we emphasize that the spectral and spatial vagueness in the transition areas should be taken into account, the classification generated from the FCM classifier is not acceptable, especially for the wetland class. Figure 5-9 shows the class distributions of the three major classes given by the FCM classifier in the TM tasseled cap feature space for test area A. Apparently the three classes are separated in layers. Figure 5-10 shows the class distribution when the alpha-cuts rule is applied in the defuzzification. Although some overlap areas are marked as transition classes, the transition class of deciduous forest and mixed stand, and the transition class that includes the three primary classes do not exist. Because the FCM

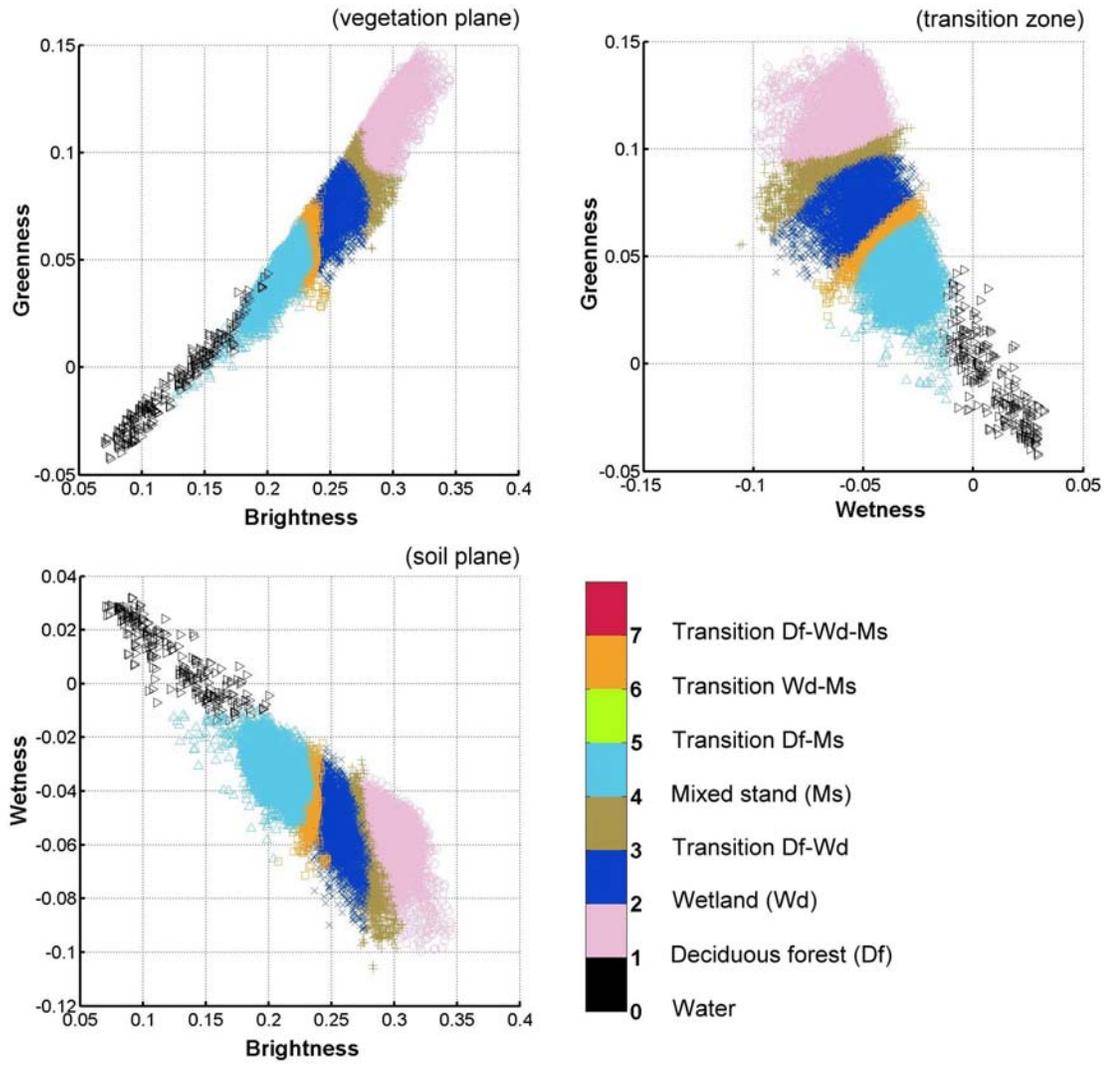
classifier partitions data into classes in layers, the transition classes are meaningless in this case.



**Figure 5-8. Class dispersion in the TM tasseled cap feature space for test area A using the SGFCM classifier**



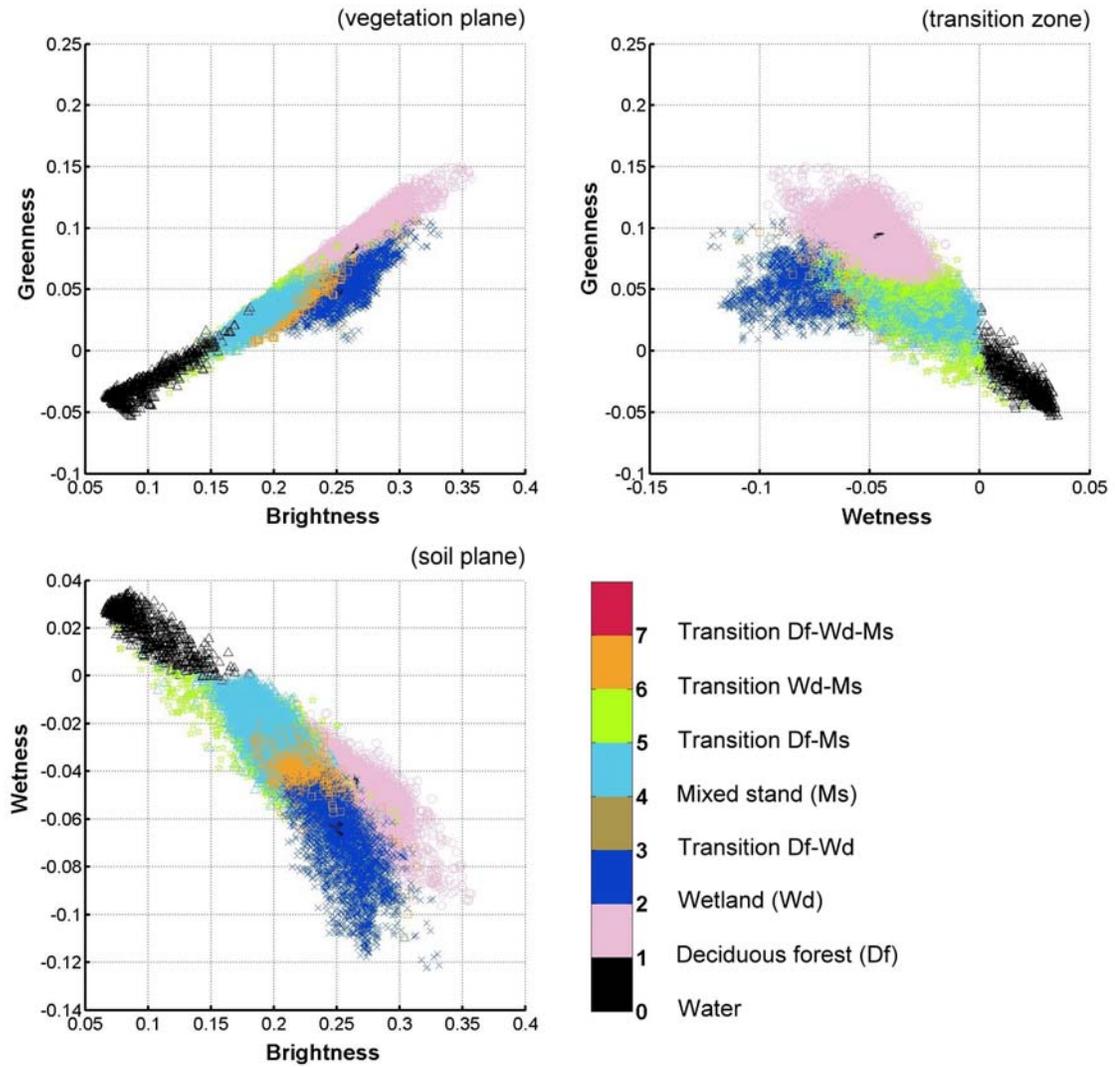
**Figure 5-9. Class dispersion in the TM tasseled cap feature space for test area A using the FCM classifier**



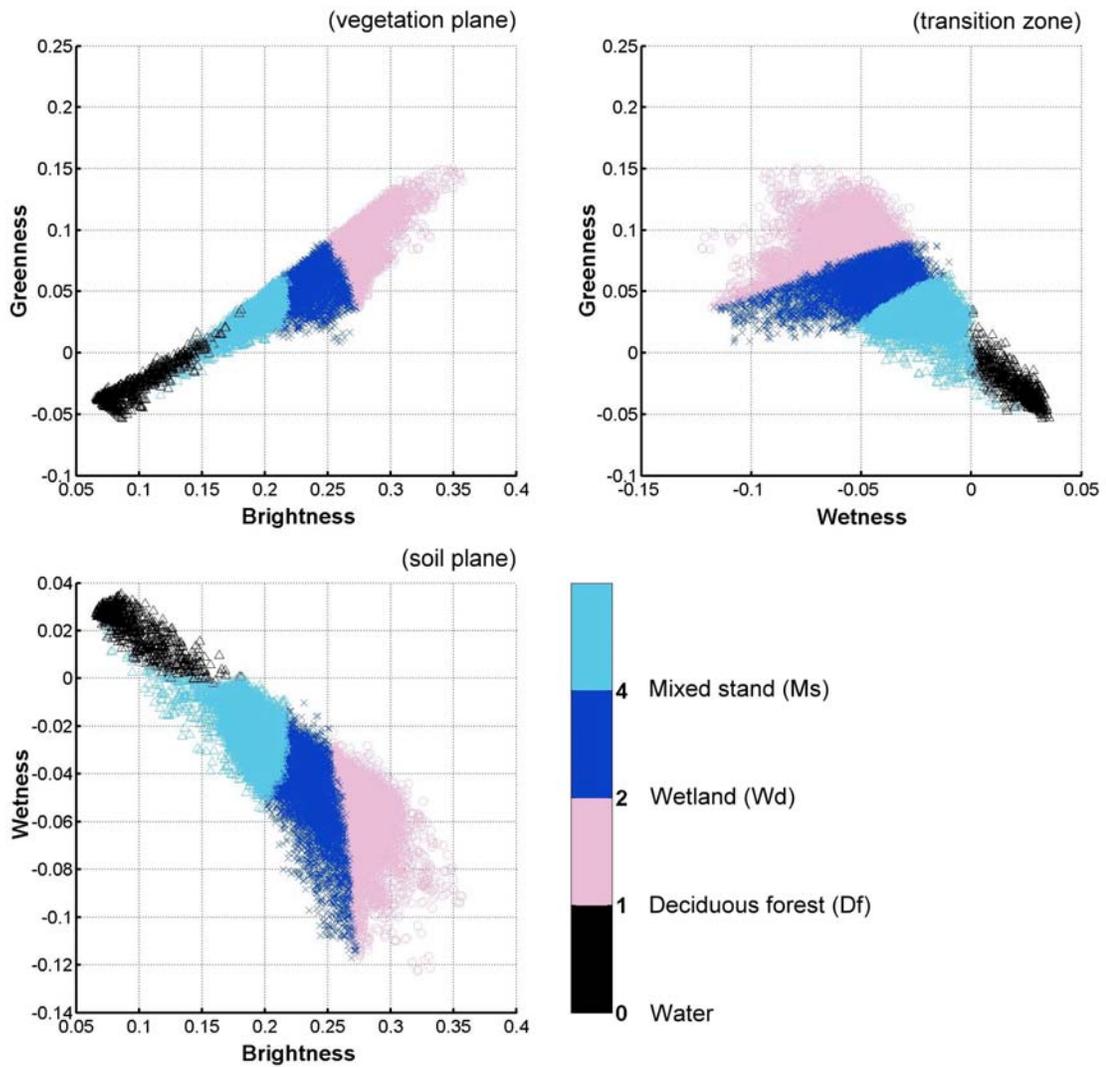
**Figure 5-10. Class dispersion in the TM tasseled cap feature space for test area A using the FCM classifier with alpha-cuts rule defuzzification.**

For test area A the class dispersions are significantly different in the feature space when the SGFCM classifier is compared to the FCM classifier. Since the SGFCM classifier uses the Mahalanobis distance in the clustering algorithm, the class dispersion demonstrates an ellipsoidal shape in the feature space. As the vegetation and the moisture conditions of wetland vary gradually in the space, an ellipsoidal shape in the feature space describes the data distribution better than a spherical shape. However, the FCM classifier, which does not incorporate information from the spatial domain, uses the Euclidean distance in the algorithm. Thus the three clusters are well separated with almost the same size and a spherical shape. The SGFCM classifier separates well wetland from deciduous forest and mixed stand from deciduous forest in test area A. The lower separability between wetland and mixed stand explains why in the visual evaluation some brighter pixels are commissioned as wetland. Because of the lower separability of wetland from other classes, it may result to a lower mapping accuracy.

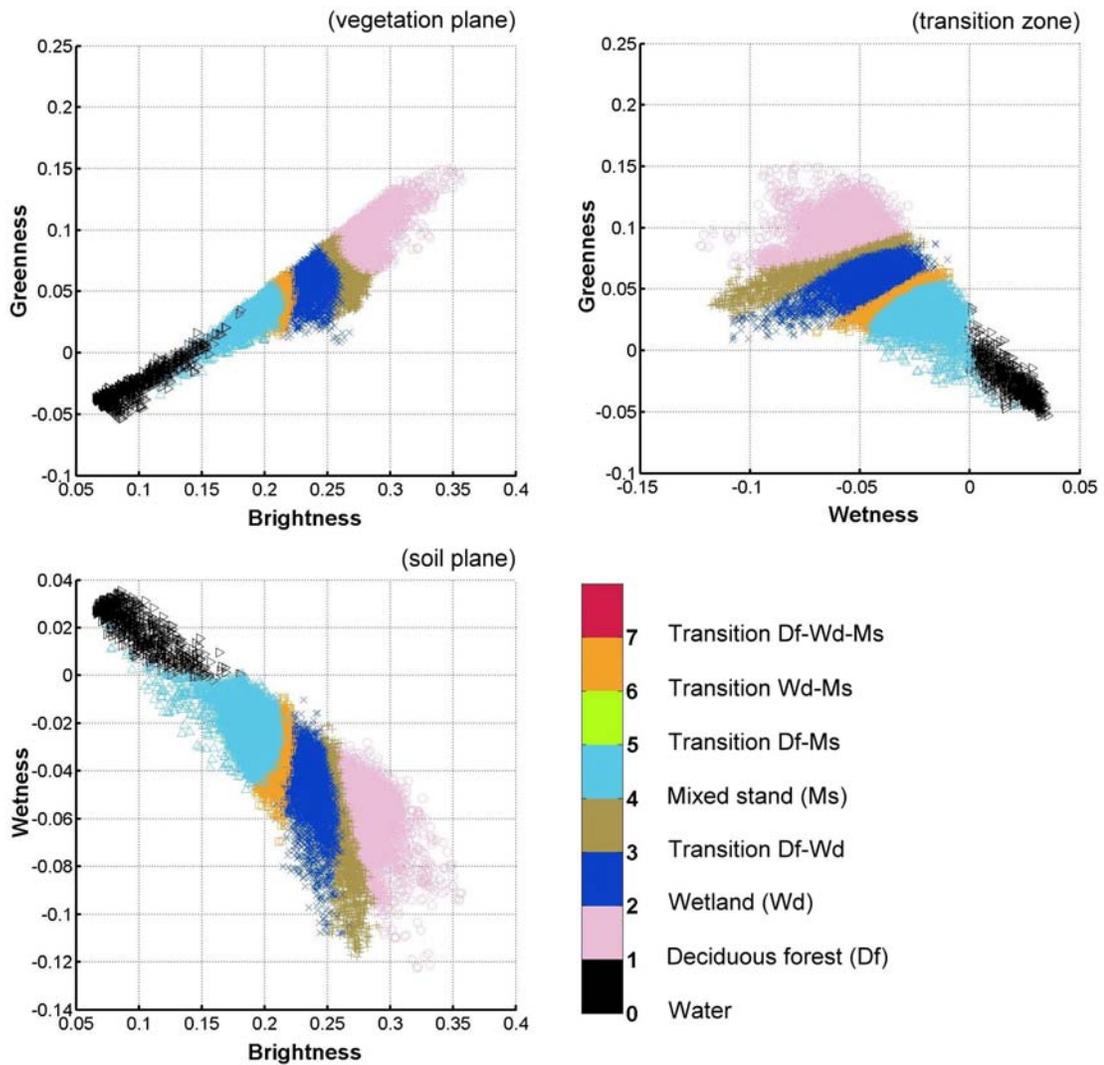
Figure 5-11 illustrates the class dispersions of the seven classes given by the SGFCM classifier in the TM tasseled cap feature space for test area B. The separability of deciduous forest and mixed stand is lower than the separability between the other two classes, i.e. wetland/mixed stand and wetland/deciduous forest. Because wetland is easier to separate from the other two classes in this data set, a higher wetland mapping accuracy can be expected. Furthermore, the SGFCM classifier allows the classes to be overlapped and gives a better mapping result. Figure 5-12 shows the class distributions of the three major classes given by the FCM classifier in the TM tasseled cap feature space for test area B. Classes are not allowed to overlap in the feature space for the FCM classifier because of maximum defuzzification. Since such restriction of the algorithm violates the natural phenomenon of the data, misclassification errors should increase when the FCM classifier is used with highest defuzzification for wetland mapping. This kind of errors has been noticed during the visual assessment of the mapping results. Figure 5-13 shows the class distributions of test area B when the partition matrix is defuzzified with the proposed alpha-cuts rule. Although the transition classes are shown as overlap areas in the feature space, they are presented in thin boundary layers between the primary classes.



**Figure 5-11.** Class dispersion in the TM tasseled cap feature space for test area B using the SGFCM classifier



**Figure 5-12.** Class dispersion in the TM tasseled cap feature space for test area B using the FCM classifier

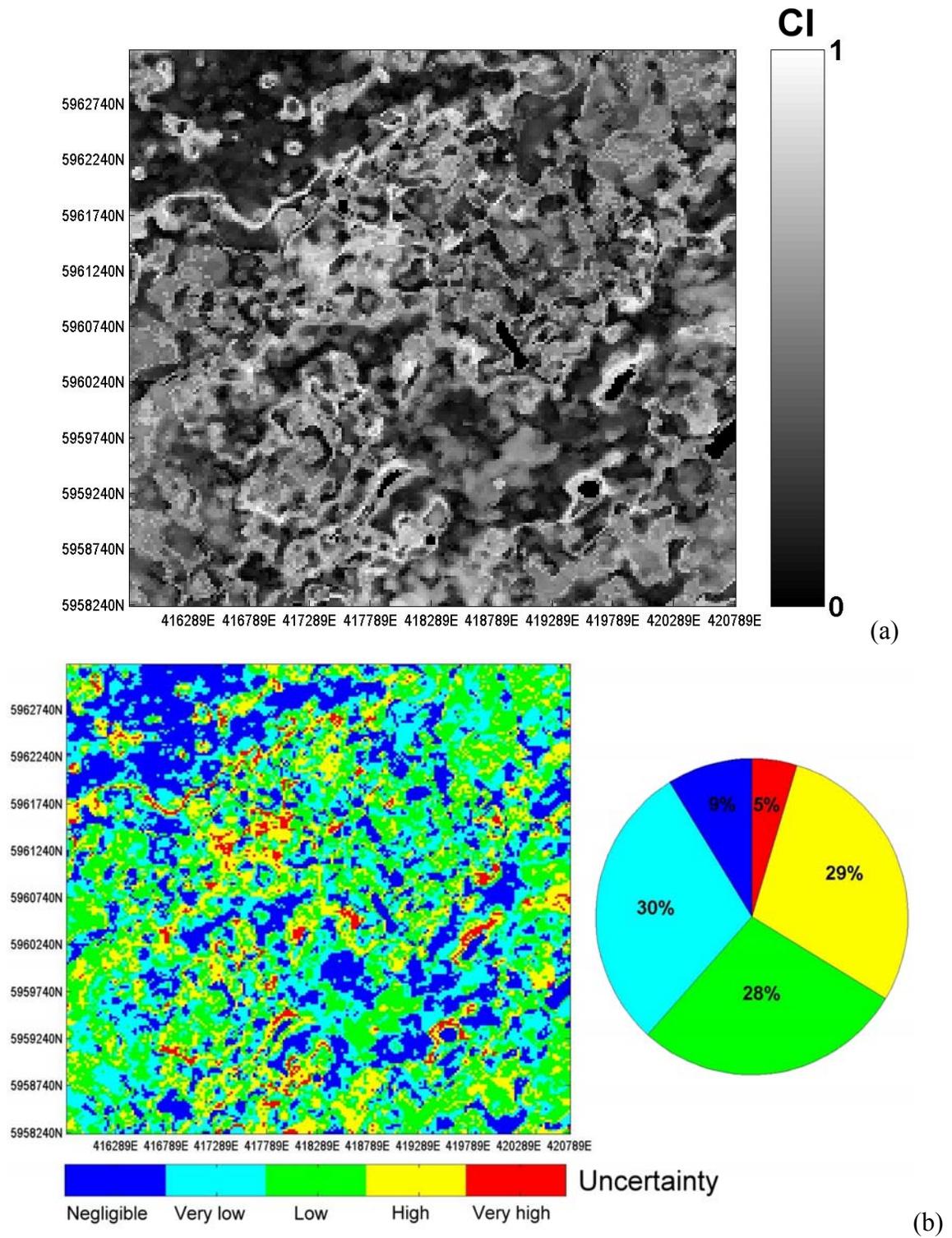


**Figure 5-13.** Class dispersion in the TM tasseled cap feature space for test area B using the FCM classifier with alpha-cuts defuzzification.

#### 5.3.4 Analysis of the confusion index and mixed pixels

The overlapped classes in the feature spaces represent the inherent ambiguity that may lead to the classification uncertainty of the maps. Pixels are classified to the “belonging” classes according to their membership values. Although pixels are assigned to the belonging classes based on highest membership values in the associated class, it does not promise that the maximum membership value of the belonging class is significantly higher than the second maximum membership value. The classification uncertainty thus increases when the differences between these two membership values are not large enough. A confusion index is used to describe the classification uncertainty of the mapping result.

Figure 5-14(a) shows the confusion index of the mapping result for test area A in gray level. The brighter the pixel is, the higher the confusion index is, i.e. the higher the uncertainty. Pixels located on the boundaries of landscape patches are found much brighter than the pixels within the patches. This finding is reasonable because boundary pixels are always spectrally mixed with two land cover types and thus more ambiguous. The confusion index is further divided into five levels to demonstrate the degree of classification uncertainty: negligible, very low, low, high, and very high. Figure 5-14(b) presents the confusion index with mixed pixel (i.e. mixel) information that illustrates the uncertainty level. Mixel information shows that around five percent of the classified pixels have a very high uncertainty and twenty-nine percent a high uncertainty in the mapping result for test area A. Only nine percent of the classified pixels have a negligible uncertainty level. Although the mixel information can be shown here, it is not possible to illustrate all the classes in a single pixel in a map. Assigning those highly mixed pixels (i.e. ambiguous pixels) to a single class could introduce some degree of classification errors, because each pixel can be assigned only to the dominant class. In the accuracy assessment, only the dominant class is taken into consideration as the reference to that pixel for the accuracy calculation. This is the explanation of why the results of the hardened fuzzy outputs are less accurate. However, most systems used in real applications are designed based on binary logic. Hardening the fuzzy outputs is necessary



**Figure 5-14. Confusion index (CI) maps displayed in (a) gray level and (b) mixel level for test area A.**

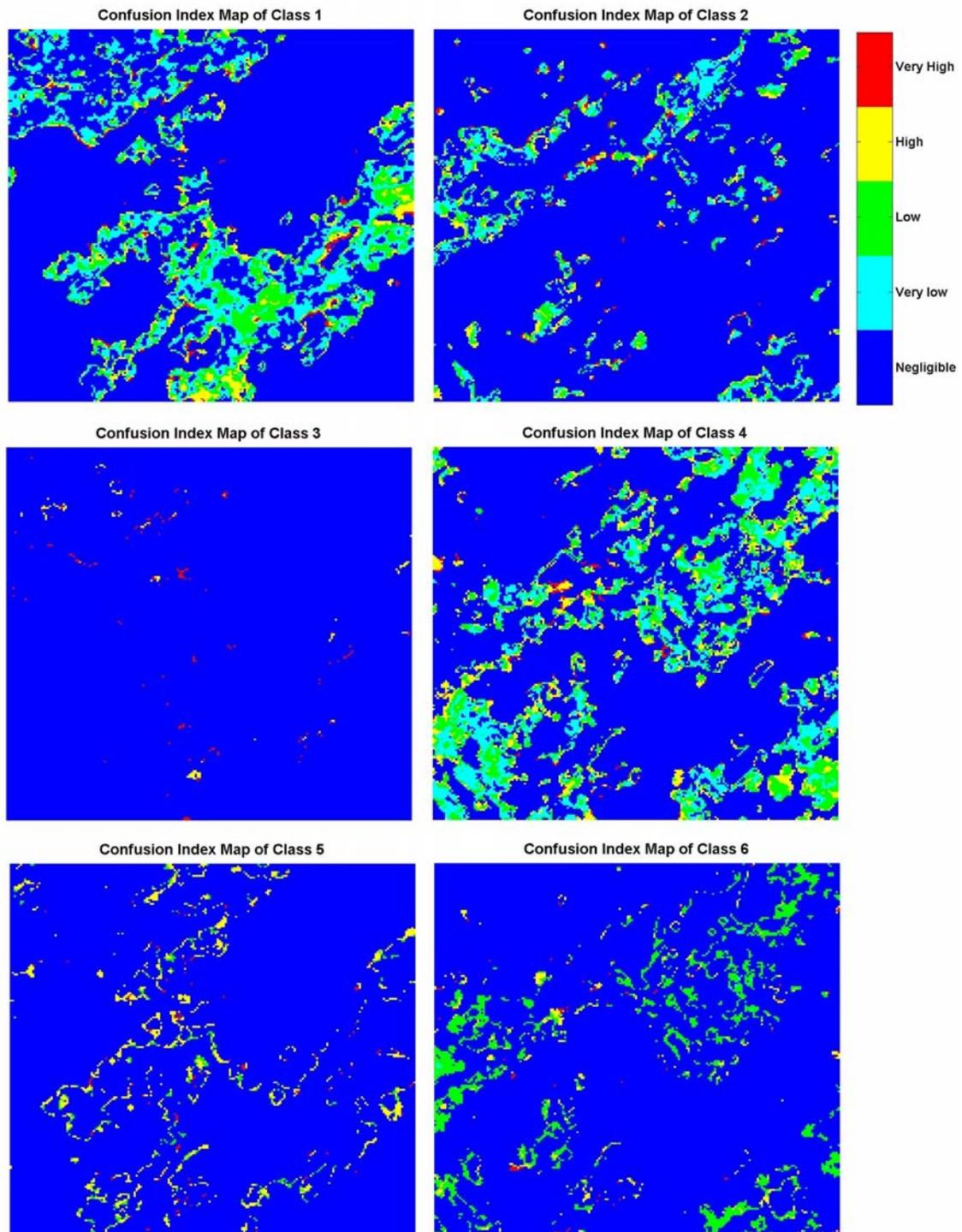
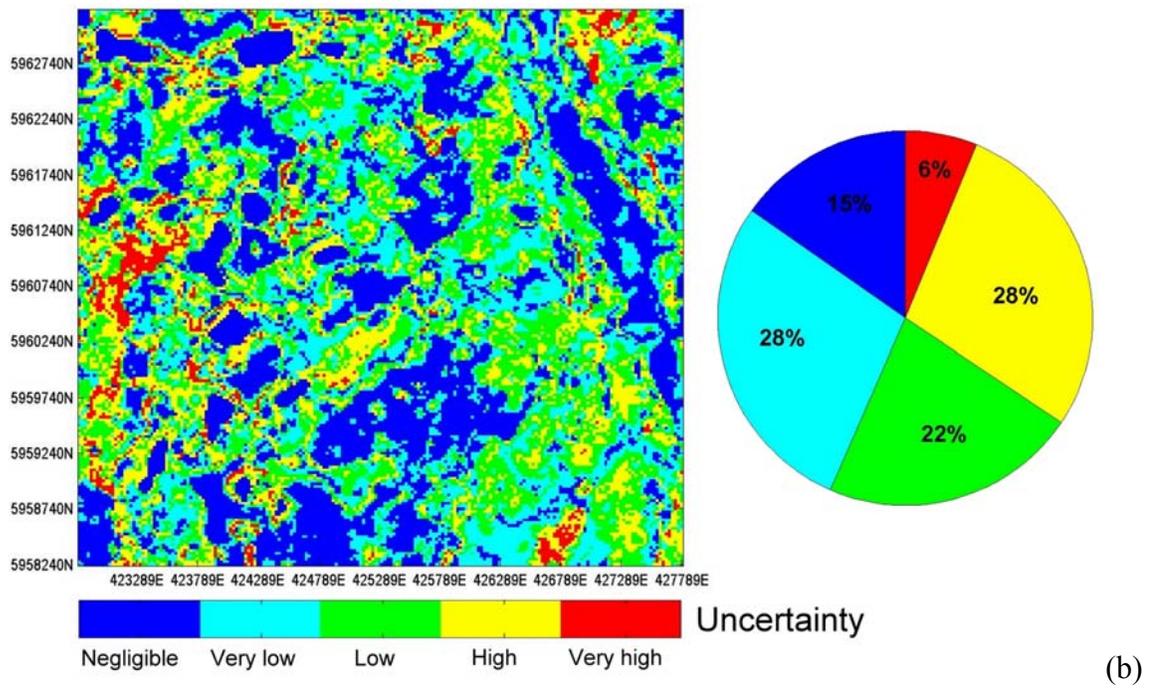
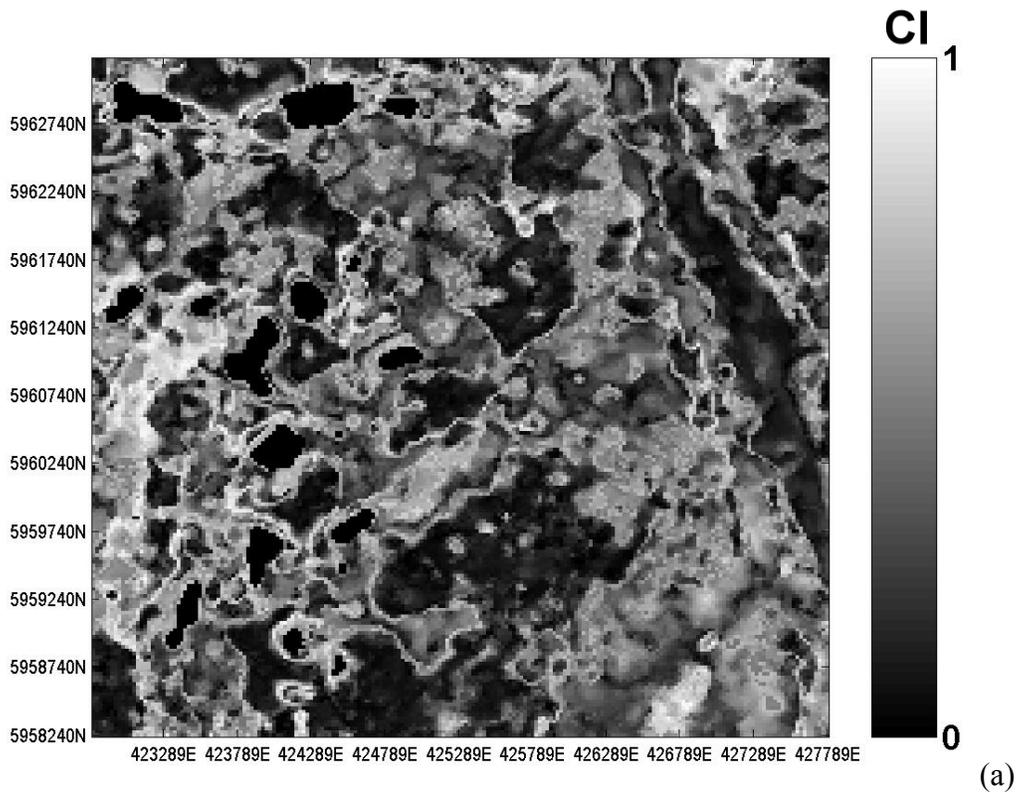


Figure 5-15. Class level uncertainty maps of test area A. (Code description: 1- Deciduous forest; 2-Wetland; 3-Transition Df-Wd; 4-Mixed stand; 5- Transition Df-Ms; 6- Transition Wd-Ms)

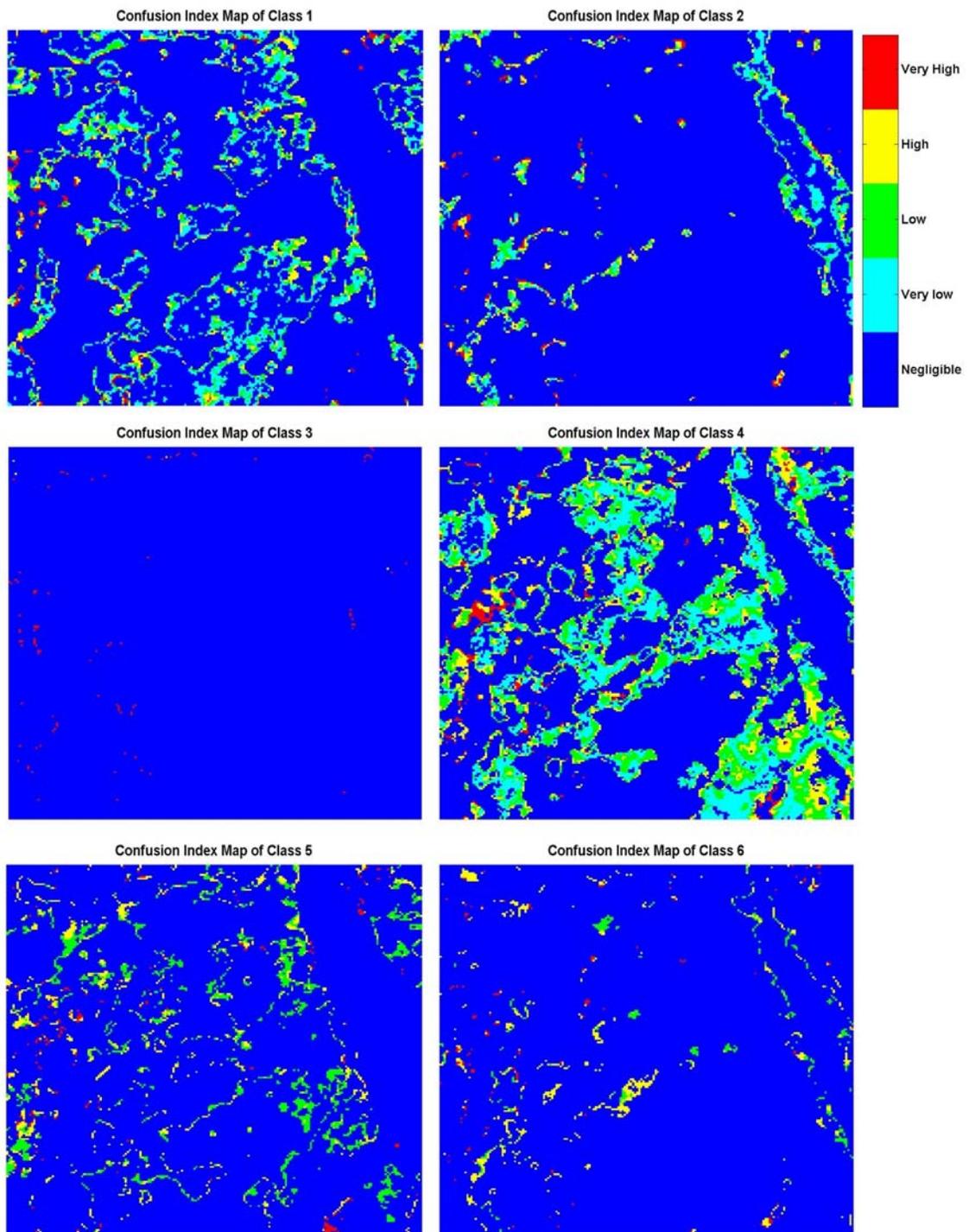
---

Being aware that not all the classes can be assigned to a single pixel when the uncertainty is high, we proposed to create new classes to allocate these ambiguous pixels. We break down the whole map into the different class levels to demonstrate the uncertainty inherent in each class. In Figure 5-15, only six classes are shown because of the empty class of the transition Df-Wd-Ms (i.e. class code “7”). Compared to the three major classes, all three transition classes show higher uncertainty. When these ambiguous pixels are highlighted on the map, the source of misclassification becomes apparent.

Figure 5-16(a) shows the confusion index in gray level for test area B. Mixel information in Figure 5-16(b) shows that fifteen percent of the classified pixels have negligible uncertainty level. Around six percent of the classified pixels have a very high uncertainty and twenty-eight percent a high uncertainty. Figure 5-17 represents the uncertainty level for each class. Most pixels belong to deciduous forest have uncertainty in negligible and lower level while the transition classes again are illustrated in the higher uncertainty level.



**Figure 5-16. Confusion index (CI) maps displayed in (a) gray level and (b) mixel level for test area B.**



**Figure 5-17. Class level uncertainty maps of test area B. (Code description: 1- Deciduous forest; 2-Wetland; 3-Transition Df-Wd; 4-Mixed stand; 5- Transition Df-Ms; 6- Transition Wd-Ms)**

## **5.4 Accuracy assessment**

To quantify the classification accuracy and errors, a confusion matrix is used. The confusion matrices of the two test areas are summarized in Table 5-1 and Table 5-2. Since wetland is the only land cover type that can be expected from the reference database, classes used in the accuracy assessment are simplified into two classes: wetland and non-wetland. Because water body is excluded in the classification, the producer's accuracy of the water class is one hundred percent.

The accuracy assessment of test area A (Table 5-1) shows that the FCM classifier provides an overall classification accuracy of 71 percent while the SGFCM classifier improves the overall accuracy to 87 percent. When examining the producer's accuracy given by the FCM classifier, we find that only 57 percent and 72 percent of accuracies are obtained for wetland and non-wetland, respectively. However, the SGFCM classifier can provide a producer's accuracy up to 65 percent for wetland and 90 percent for non-wetland. The commission error is also interesting in an accuracy assessment. Since the FCM classifier is used for wetland mapping, commission error of non-wetland is only 6 percent; the error is acceptable in contrast to the 81 percent commission error for the wetland class. Compared to the FCM classifier, the SGFCM classifier reduces the commission errors to 4 percent and 57 percent for non-wetland and wetland respectively. So half of the wetland classified pixels should be found to be wetland on the ground compared with 1/5 of them only for FCM.

The effectiveness of the SGFCM classifier is also examined through the accuracy assessment of test area B. In this example, an improvement of the overall accuracy and the commission error is also demonstrated in Table 5-2. When the FCM classifier is compared to the SGFCM classifier, the overall accuracy increases from 70 to 93 percent and the commission error decreases from 78 to only 26 percent for wetland, which means that with SGFCM 3/4 of the wetland classified pixels have a high probability to be wetland on the ground compared to 1/4 for FCM.

**Table 5-1. Confusion matrix for wetland mapping of the test area A (in pixels) for the:**

(a) FCM classifier

Class	Water	Wetland	Non-wetland			Producer's Accuracy (%)	Commission Error (%)
			Deciduous forest	Mixed stand	Non-wetland		
Water	294	0	0	0	100	0	
Wetlands	0	2286	574	1137	57.0	81.0	
Non-wetland	0	9779	12841	13089	72.6	6.2	

**Overall accuracy = 71.3 %**

(b) FCM classifier

Class	Water	Wetland	Non-wetland						Producer's Accuracy (%)	Commission Error (%)	
			Deciduous forest		Mixed stand		Transition				
			D-Wd	Wd-Ms	Df-Ms	Wd-Ms	Df-Wd-Ms	Wd-Ms	Df-Wd-Ms		
Water	294	0	0	0	0	0	0	0	0	100	0
Wetlands	0	2601	268	86	62	447	533	0	0	65.0	57.0
Non-wetland	0	3481	15012	152	1806	12238	3020	0	0	90.0	4.1

**Overall accuracy = 87.8 %**

Note: *Df* is the abbreviation of Deciduous forest; *Wd* is the abbreviation of Wetland; *Ms* is the abbreviation of Mixed stand

**Table 5-2. Confusion matrix for wetland mapping of the test area B (in pixels) for the:**

(a) FCM classifier				Non-wetland			Producer's Commission Error (%)	
Class	Water	Wetland	Deciduous forest			Mixed stand	Accuracy (%)	Error (%)
			D-Wd	Df-Ms	Wd-Ms			
Water	1275	0	0	0	0	100	0	
Wetlands	0	2710	1222	606	606	59.7	78.6	
Non-wetland	0	9974	8545	15668	15668	70.8	7.0	
<b>Overall accuracy= 70.5 %</b>								
(b) FCM classifier				Non-wetland			Producer's Commission Error (%)	
Class	Water	Wetland	Deciduous forest			Mixed stand	Accuracy (%)	Error (%)
			D-Wd	Df-Ms	Wd-Ms			
Water	1275	0	0	0	0	100	0	
Wetlands	0	3156	263	102	389	70.0	26.0	
Non-wetland	0	1128	13384	3432	935	96.7	4.0	
<b>Overall accuracy= 93.7 %</b>								

Note: *Df* is the abbreviation of Deciduous forest; *Wd* is the abbreviation of Wetland; *Ms* is the abbreviation of Mixed stand

Based on the reference data, the topographic maps obtained from NTDB, the accuracy assessments show that the SGFCM classifier can provide higher mapping accuracy than the FCM classifier. This result is satisfactory with the expectation of the visual evaluation of classification maps.

However, some weaknesses of the reference data should be noticed: the topographic maps are generated with some data structure errors. For example, errors may come from the aerial photographic interpretation. In consequence, some small wetland areas may be neglected by human interpretation depending on the scale of the imagery used. On the other hand, the reference data were generated decades ago and has not been updated. The change of the landscape should be considered when using the reference maps. Image georeferencing may also cause some shifting errors. When an accuracy assessment is conducted based on pixel unit, the shifting between a reference map and a classification map results in errors. For example, the accuracy of NTDB data is about 25m, which is equal to one pixel error of Landsat image used in the study.

Despite errors inherent to the reference data, such errors have limited effects on the comparison of the two algorithms because of the same reference data used in the accuracy assessment. Therefore according to both the qualitative and quantitative assessments, the conclusion can be drawn that the SGFCM classifier is better suited than the standard FCM classifier.

## **5.5 Summary**

This chapter has presented the findings of the study results and the relevant discussions. First, the preliminary examination of the data dispersion in the TM tasseled cap feature spaces has been discussed according to the three dimensions: the plane of vegetation, the transition zone, and the plane of soil. The result showed that both the original data sets of test area A and B are dominated by forested type ground objects and water. The forest cluster could be further partitioned into two to three subclasses. Second, the

semivariogram behaviors of each training land cover class were analyzed. Water body showed a flat curve; deciduous forest illustrated a stable variation of variances in all three TM tasseled cap features; wetland had larger variances in semivariogram with a wave shape; mixed stand exhibited a typical semivariogram curve type of dry conifers. Furthermore, the classification results were given and discussed regarding to membership functions, visual evaluations, class dispersions, and classification uncertainty. Spectrally mixed pixels that were rejected to the transition class showed a higher-level uncertainty in the classification. Finally, an accuracy assessment was conducted to quantify the findings of the visual evaluation. The results showed that the SGFCM classifier is more effective than the FCM classifier for wetland mapping because it has demonstrated its ability to improving the producer's and overall accuracy, and to reduce the commission errors.

## CHAPTER 6

# CONCLUSIONS AND FUTURE SCOPE

*“Everything is deducible, everything is linked. The cause allows one to guess the effect, just as each effect allows one to reconstruct a cause. The scientist can resuscitate in this manner even the warts of ancient times. From this comes without doubt the prodigious interest that an architectural description can inspire when the writer's fantasy is faithful to its basic elements. Cannot each person reattach it to its past by rigorous deductions?”(In The Search for the Absolute)*

***Honoré de Balzac, 1799-1850***

*French writer*

### **6.1 Conclusions**

To identify wetlands from a multispectral satellite image, a robust classification algorithm is one of the critical keys to derive a reliable mapping outcome. However, a traditional classification algorithm, a “hard” classifier, is built based on binary logic, which cannot give good descriptions of mixed and imprecise data since pixels are assumed to be pure. Although the fuzzy concept is introduced in the Fuzzy C-Means (FCM) clustering algorithm, a “soft” classifier, to describe data attributes with fuzzy membership functions, the spatial variability of the data attributes is completely ignored in the algorithm. To compensate these weaknesses, this thesis has presented a variance

---

involved fuzzy classifier, the Semivariogram Guided Fuzzy C-Means (SGFCM) clustering algorithm, by modifying the standard FCM classifier.

The main idea of the SGFCM classifier is the incorporation of the spatial variability of the data set itself into the clustering algorithm. Because wetland mapping is the purpose of the application, the vagueness of wetland's attributes should be taken into account by classifiers; especially such inherent vagueness that leads to the spectral mixture and the classification uncertainty that reflects on an image spectrally. This claim has been demonstrated in the analysis of the class dispersions in the TM tasseled cap features. Crist *et al.* (1986) stated, "Neither the TM tasseled cap transformation nor any other transformation can create information that was not present in the original data". In this study when the image data is transformed into the TM tasseled cap features, the data dispersion exhibits some emerged clusters in the feature space. The data dispersion also demonstrates the data fuzziness inherent in the cluster boundaries. The TM tasseled cap transformation greatly facilitates the extraction of the information contained in the multispectral data. However, without a proper classifier the information may still be misinterpreted and lead to improper applications. This thesis has shown that the inherent vagueness of natural objects is revealed by the SGFCM classifier but undiscovered by the FCM classifier when comparing the data dispersion in the preliminary examination to the class dispersion of the final classification results.

Since the tasseled cap features (i.e. brightness, greenness, and wetness) can make direct association between the feature response and the physical characteristics of the scene classes, the semivariograms derived from these features can represent the spatial variations of the physical characteristics of a landscape element. This thesis has proved this hypothesis, while the scene classes illustrated the different semivariogram patterns when their behaviors were analyzed. This evidence gives the possibility of treating semivariogram as a kind of texture index to be employed in the classifier. However, the selection of training sites for deriving semivariogram may be a critical issue just as other supervised classification approaches. This is because the sizes of landscape patches vary

---

in the nature environment: even same type forest stands may have different stand size, unless the land cover belongs to a man-made object or a cultivated crop. Therefore the window size, i.e. the maximum lag distance, used for deriving a semivariogram pattern should be large enough to generate a stable and a representative semivariogram pattern for the associated class. Furthermore, the computation time should be of concern for deriving semivariogram texture features if the window size is too large. This thesis has also demonstrated that the SGFCM classifier has an ability to handle a classification based on multi-scale texture features.

The new developed SGFCM classifier shows its effectiveness on wetland mapping with the incorporation of semivariogram texture features in the classification algorithm. Compared to the standard FCM classifier, the SGFCM not only increases the producer's accuracy but also reduces the commission errors quite dramatically in this study. The improvement of the overall accuracy shows an increase from 70 to 93 percent. Two things make a major contribution to this improvement. The first one is the consideration of the fuzzy covariance, i.e. replacing the Euclidean distance measure with the Mahalanobis distance. The second one is the premeditation of the semivariogram attributes in the clustering algorithm, i.e. adding a texture-typed weighting factor in the objective function. This thesis demonstrates that the spatial variability inherent in the image data set can provide extra information beyond a digital number itself. A robust classifier is required to have a capability of discovering the spatial variability in land cover mapping, especially when dealing with a mixed and imprecise image data set.

For a mixed and imprecise data set, a fuzzy classifier provides a better description of the data than a "hard" classifier does. However, hardening the fuzzy output is unavoidable in the defuzzification because some real applications need to be a single scalar quantity as opposed to a fuzzy set. Although the maximum function has been widely used in classification applications to defuzzify the fuzzy outputs of the FCM classifier, it neglects the ambiguity inherent in the membership values. The maximum function does not consider whether a significant difference appears between the membership values of two

---

classes or not; pixels are assigned to the one predefined class that has the highest fuzzy membership value. The uncertainty thus remains in the classification and results in lower accuracy. By contrast, the alpha-cuts defuzzification method with newly created transition classes has successfully extract out the ambiguous pixels resulting to misclassification. In an image, these ambiguous pixels are always located in the boundaries of the landscape elements. Reflected back to the nature environment, the boundary areas of two ecosystems are always weak and sensitive to external disturbances. By highlighting the ambiguous pixels in the boundary areas, a map provides an additional information for users to notice the “status and trend” of the associated land cover.

## **6.2 Future Scope**

The partition matrix consisting of fuzzy membership values is an important outcome of a fuzzy clustering algorithm. However, a question remains here: how to interpret this partition matrix in the following step? Further applications of these membership values, either the defuzzification or the mathematical combination, are fields that need to be discovered. For example, setting different thresholds for the alpha-cuts defuzzification method inspires an idea for delineating a buffer zone for a sensitive ecosystem, or for formatting a fuzzy ground object. In addition, a mixed pixel can be further decomposed into sub-pixels according to the fuzzy membership values to examine the portion of each class, i.e. to provide pixel unmixing or sub-pixel classification. If another higher spatial resolution image is available, the locations of these sub-pixels can even be more accurately defined. The weakness of uncertainty of the ambiguous pixels may be the strength of any inspiration for a future study.

---

**REFERENCES**

- Ahmed, M.N., S.M. Yamany, N. Mohamed, A.A. Farag, and T. Moriarty. 2002. A modified fuzzy c-means algorithm for bias field estimation and segmentation of MRI data. *IEEE Transactions on Medical Imaging*. **21**(3): 193-199.
- Anderson MC, Neale CCM, Li F, Norman J, Kustas WP, Jayanthi H, Chavez JL. 2004. Upscaling ground observations of vegetation water content, canopy height, and leaf area index during Smex 02 using aircraft and Landsat imagery. *Remote Sensing Of Environment* **92**: 447-464.
- Arzandeh S and Wang J. 2002. Texture evaluation of RADARSAT imagery for wetland mapping. *Canadian Journal of Remote Sensing* **28**(5): 653-666.
- Atkinson PM and Lewis P. 2000. Geostatistical classification for remote sensing: an introduction. *Computers and Geosciences* **26**(3): 361-371.
- Bezdek, J.C., Ehrlich, R. and Full, W (1984) FCM: The fuzzy c-means clustering algorithm. *Computational Geoscience*, **10**: 191–203.
- Bezdek JC and Pal SK. 1992. *Fuzzy Models for Pattern Recognition: Methods that Search for Structures in Data*. IEEE Press.
- Blaschke T and Strobl J. 2001. What's wrong with pixels? Some recent developments interfacing remote sensing and GIS. *GIS—Zeitschrift für Geoinformationssysteme* **6**: 12-17.
- Bruzzone L, Cossu R. Vernazza G. 2002. Combination parametric and non-parametric algorithms for a partially unsupervised classification of multitemporal remote-sensing images. *Information Fusion* **3**(4): 289-297.
- Burrough PA. 1996. Natural objects with indeterminate boundaries. In *Geographic Objects with Indeterminate Boundaries*. eds PA Burrough and AU Frank. London Taylor & Francis. pp. 3-28.
- Burrough PA and Frank AU. 1996. *Geographic Objects with Indeterminate Boundaries*. *GISDATA* **2**. London Taylor & Francis. pp. 345.
- Carr JR. 1996. Spectral and textural classification of single and multiple band digital images. *Computers and Geosciences* **22**(8): 849-865.
- Carr JR and Miranda FP. 1998. The semivariogram in comparison to the co-occurrence matrix for classification of image texture. *IEEE Transactions on Geoscience and Remote Sensing* **36**(6): 1945-1952.

- 
- Carter V, Gammon PT, and Garrett MK. 1994. Ecotone dynamics and boundary determination in the Great Dismal Swamp. *Ecological Applications* **4**(1): 189-203.
- Chamorro-Martinez, Sanchez D, Prados-Suarez B, Galan-Peraales E, and Vila MA. 2003. A hierarchical approach to fuzzy segmentation of color images. *The IEEE International Conference on Fuzzy Systems* 966-970.
- Cheng T and Molenaar M. 1999. Objects with fuzzy spatial extent. *Photogrammetric Engineering & remote Sensing* **65**(7): 797-801.
- Chica-Olmo M and Abarca-Hernandez F. 2000. Computing geostatistical image texture for remotely sensed data classification. *Computers and Geosciences* **26**(3): 373-383.
- Chiu WY and Couloigner I. 2004a. Evaluation of incorporating texture into wetland mapping from multispectral images. *EARSeL eProceeding* **3**(3): 363-371.
- Chiu WY and Couloigner I. 2004b. The use of radar data incorporating with texture for extracting wetland patches. *Proceeding of the 25th Anniversary Meeting of the Society of Wetland Scientists*, July 18-23 2004, Society of Wetland Scientists, Seattle, WA. pp. 153-154.
- Chumsamrong W, Thitimajshima P, and Rangsanseri Y. 2000. Synthetic Aperture Radar (SAR) image segmentation using a new modified fuzzy c-means algorithm. *Proceedings of Geoscience and Remote Sensing Symposium* **2**: 624-626.
- Civco DL and Hurd JD. 1999. A hierarchical approach to land use and land cover mapping using multiple image types. *Proceeding of ASPRS Annual Convention*, pp. 687-698.
- Cohen WB, Spies TA, and Fiorella M. 1995. Estimating the age and structure of forests in a multi-ownership landscape of western Oregon, U.S.A. *International Journal of Remote Sensing* **16**: 721-746.
- Coucleli H. 1996. A typology of geographic entities with ill-defined boundaries. In *Geographic Objects with Indeterminate Boundaries*. eds Burrough PA and Frank AU. Taylor and Francis, pp. 230-241.
- Cowardin LM, Carter V, Golet FC, and LaRoe ET. 1979. *Classification of Wetlands and Deepwater Habitats of the United States*. US Department of the Interior, Fish and Wildlife Service, Washington DC, 131 pp.
- Crist EP and Cicone RC. 1984. Application of the tasseled cap concept to simulated thematic mapper data. *Photogrammetric Engineering And Remote Sensing* **50**(3): 343-352.

- Crist EP, Laurin R, and Cicone RC. 1986. Vegetation and soils information contained in transformed thematic mapper data. In *Proceedings of IGARSS' 86 Symposium*, pp. 1465-1470.
- Cuevas, E., D. Zaldivar, and R. Rojas. 2004. Fuzzy segmentation in image processing. XXVI International Congress on Electrical Engineering Electro.
- Curran PJ. 1988. The semivariogram in remote sensing: an introduction. *Remote Sensing of Environment* **24**: 493-507.
- Daily GG. 1997. *Nature's Services: Societal Dependence on Natural Ecosystems*. Island Press, Washington, 329 pp.
- Davidson I, Vanderkam R, and Padilla M. 1999. Review of wetland inventory information in North America. In *Global Review of Wetland Resources and Priorities for Wetland Inventory*, eds CM Finlayson & AG Spiers. [online]. Supervising Scientist Report 144, Canberra, Australia, 38 pp. Available from: <[http://www.wetlands.org/inventory&/GroWI/report\\_list.html](http://www.wetlands.org/inventory&/GroWI/report_list.html)> [June 6, 2005]
- Day JW, Jr. Psuty NP, and Perez BC. 2000. The role of pulsing events in the functioning of coastal barriers and wetlands: implications for human impact, management and the response to sea level rise. In *Concept and Controversies in Tidal Marsh Ecology*. eds MP Weinstein and DA Kreeger. Kluwer Academic Publishers, pp. 633-659.
- De Bruin S and Stein A. 1998. Soil-landscape modelling using fuzzy c-means clustering of attribute data derived from a digital elevation model (DEM). *Geoderma* **83**: 17-33.
- Dechka, JA, Franklin SE, Watmough MD, Bennett RP, and Ingstrup DW. 2002. Classification of wetland habitat and vegetation communities using multitemporal Ikonos imagery in southern Saskatchewan, *Canadian Journal of Remote Sensing* **28**(5): 679-685.
- Fisher PF and Pathirana S. 1990. The evaluation of fuzzy membership of land cover classes in the suburban zone. *Remote Sensing of Environment* **34**: 121-132.
- Foody GM and Cox DP. 1994. Sub-pixel land cover composition estimation using a linear mixture model and fuzzy membership functions. *International Journal of Remote Sensing* **15**(3): 619-631.
- Foody GM and Boyd DS. 1999. Detection of partial land cover change associated with the migration of inner-class transitional zones. *International Journal of Remote Sensing* **20**(14): 2723-2740.
- Franklin SE, Wulder MA, and Lavigne MB. 1996. Automated derivation of geographic window sizes for use in remote sensing digital texture analysis. *Computers and Geosciences* **22**(6): 665-673.

- Franklin SE, Hall RJ, Mosdal LM, Maudie AJ, and Lavigne MB. 2000. Incorporating texture into classification of forest species composition from airborne multispectral images. *International Journal of Remote Sensing* **21**(1): 61-79.
- Gao BC. 1996. NDWI – A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment* **58**(3): 257-266.
- Gluck M, Rempel R, Uhlig PWC. 1996. An evaluation of remote sensing for regional wetland mapping applications. *Forest Research Report* 137. Ontario Forest Research Institute, Sault Ste Marie, Ontario, Canada, 33 pp.
- Gordan M, Kotropoulos C, Georgakis A, and Pitas I. 2002. A new fuzzy c-means based segmentation strategy applications to lip region identification. IEEE-TTTC International Conference on Automation, Quality and Testing, Robotics. May 23-25. Cluj-Napoca, Romania.
- Gustafson, D.E. and W.C. Kessel. 1979. Fuzzy clustering with a fuzzy covariance matrix. In Proceeding of IEEE Conference on Decision Control. San Diego, CA. 761-766.
- Haralick, RM and Shanmugan K. 1974. Combined spectral and spatial processing of ERTS imagery data. *Remote Sensing of Environment* **3**: 3-13.
- Haralick, RM. 1979. Statistical and structural approaches to texture. *Proceeding of IEEE* **67**(5): 786-803.
- Huang C, Wylie B, Yang L, Homer C, and Zylstra G. 2001. *Derivation of a tasseled cap transformation based on Landsat 7 at-satellite reflectance*. USGS EROS Data Center.
- Irish RR. 2000. *Landsat 7 science data user's handbook*. Report 430-15-01-003-0. National Aeronautics and Space Administration. Available from: [http://ftpwww.gsfc.nasa.gov/IAS/handbook/handbook\\_toc.html](http://ftpwww.gsfc.nasa.gov/IAS/handbook/handbook_toc.html) [April 30, 2005]
- Jensen JR. 1996. *Introductory Digital Image Processing*. 2<sup>nd</sup> ed. Upper Saddle River, NJ, Prentice Hall, 184 pp.
- Jeon B and Landgrebe DA. 1999. Partially supervised classification using weighted unsupervised clustering. *IEEE Transactions on Image Processing* **37**(2): 1073-1079.
- Johnson RA and Wichern DW. 1988. *Applied Multivariate Statistical Analysis*. 2<sup>nd</sup> ed. Prentices-Hall, New Jersey, 607 pp.
- Kauth RJ and Thomas GS. 1976. The tasseled cap – graphic description of the spectral – temporal development of agricultural crops as seen in Landsat. *Proceedings on the Symposium on Machine Processing of Remotely Sensed Data*, pp. 41-51.

- 
- Kiil AD, Lieskovshy RJ, and Grigel JE. 1973. *Fire hazard classification for Prince Albert National Park, Saskatchewan*. Information Report Nor-x-58. Forestry Service of Environment Canada.
- Krishnapuram R. and J. Kim. 1999. A note on the Gustafson-Kessel and Adaptive Fuzzy Clustering Algorithms. *IEEE Transactions on Fuzzy Systems*. **7**(4): 453-461.
- Kulik L. 2003. Spatial vagueness and second-order vagueness. *Spatial cognition & Computation* **3**: 157-183.
- Leung SH, Wang SL, and Lau WH. 2004. Lip image segmentation using fuzzy clustering incorporating an elliptic shape function. *IEEE Transactions on Image Processing* **13**(1): 51-62.
- Lucas, N.S., Shanmugam, S., and Barnsley, M. (2002) Sub-pixel habitat mapping of a coastal dune ecosystem. *Applied Geography*, **22**: 253–270.
- Lyon JG and McCarthy J. 1995. Introduction to Wetlands and Environmental Applications of GIS. In *Wetland and Environmental Applications of GIS*. eds. Lyon JG and McCarthy J. CRC Press, pp. 3-4.
- Marceau DJ, Howarth PJ, Dubois JM, and Gratton DJ. 1990. Evaluation of the Grey-Level Co-Occurrence Matrix method for land-cover classification using SPOT imagery. *IEEE Transactions on Geoscience and Remote Sensing* **28**(4): 513-518.
- Mäenpää T. 2003. The local binary pattern approach to texture analysis—extensions and applications. University of Oulu. Ph. D. Dissertation. 78 pp.
- McBratney AB and Moore AW. 1992. Application of fuzzy sets to climatic classification. *Agriculture and Forest Meteorology* **35**: 165-185.
- Milton GR, Bélanger L, Crevier Y, Hélie R, Hurley J, and Kazmerik BH. 2003. Development of a remote-sensed wetland inventory and classification system for Canada. *Backscatter* **14**(1): 32-34.
- Milton GR and Hélie R. 2003. Wetland inventory and monitoring: partnering to provide a national coverage. In *Wetland Stewardship in Canada*. North American Wetlands Conservation Council Report 03-2.
- Miranda FP, Fonseca LEN, and Carr JR. 1998. Semivariogram textural classification of JERS-1 (Fuyo-1) SAR data obtained over a flooded area of the Amazon rainforest. *International Journal of Remote Sensing* **19**(3): 549-556.
- Mitsch WJ and Gosselink JG. 1993. *Wetlands*. 2<sup>nd</sup> ed. Van Nostrand Reinhold, New York, 722 pp.

- 
- National Research Council. 1995. *Wetlands: Characteristics and Boundaries*. [online]. National Academy Press, Washington DC, 268 pp. Available from: <<http://www.nap.edu/books/0309051347/html>> [June 6, 2005]
- Natural Resource Canada. 1986. Wetland regions. The National Atlas of Canada 5<sup>th</sup> ed. Available from: <<http://atlas.gc.ca/site/english/maps/archives/5thedition/environment/ecology/mcr4108>> [June 9, 2005]
- Natural Resource Canada. 2003. National Topographic Data Base Edition 3.1 Simplified User's Guide. Available from: <<http://www.lib.unb.ca/gddm/maps/NRCan/ntdguid2.pdf>> [September 14, 2005]
- Noordam JC, van den Broek WHAM, and Buydens LMC. 2000. Geometrically guided fuzzy c-means clustering for multivariate image segmentation. *Proceedings of 15<sup>th</sup> International conference pattern recognition*. Barcelona, Spain. pp. 463-465
- Oak Ridge National Laboratory Distributed Active Archive Center (ORNL DAAC). 2001. BOREAS. Web page. Available from: <[http://www-eosdis.ornl.gov/BOREAS/boreas\\_home\\_page.html](http://www-eosdis.ornl.gov/BOREAS/boreas_home_page.html)> [April 30, 2005]
- Ozesmi SL and Bauer ME. 2002. Satellite remote sensing of wetlands. *Wetlands Ecology and Management* **10**: 381-402.
- Parmuchi MG, Karszenbaum H, and Kandus P. 2002. Mapping wetlands using multi-temporal RADARSAT-1 data and a decision-based classifier. *Canadian Journal of Remote Sensing* **28**(2): 175-186.
- Pedrycz W and Waletzky J. 1997. Fuzzy clustering with partial supervision. *IEEE Transactions on Systems, Man and Cybernetics –Part B: Cybernetics* **27**(5):787-795.
- Peng X, Wang J, Raed M, and Gari J. 2003. Land cover mapping from RADARSAT stereo images in a mountainous area of southern Argentina. *Canadian Journal of Remote Sensing* **29**(1): 75-87.
- Pham D. 2001. Spatial Models for fuzzy clustering. *Computer vision and image Understanding* **84**: 285-297.
- Richards JA. 1993. *Remote Sensing Digital Image Analysis: An Introduction*. 2<sup>nd</sup>. Springer-Verlag, 340 pp.
- Ross TJ. 1995. *Fuzzy Logic with Engineering Applications*. McGraw-Hill, 600 pp.
- Schmidt M and Schoettker B. 2004. Sub-pixel analysis in combination with knowledge based decision rules to optimize a land cover classification. In *Remote Sensing in Transition*. ed Goossens G. Milpress, Rotterdam, pp.53-59.

- Schneider M. 1996. Modelling spatial objects with undetermined boundaries using the realm/ROSE approach. In *Geographic Objects with Indeterminate Boundaries*. eds PA Burrough and AU Frank. London Taylor & Francis. pp. 141-152.
- Soper, J.D. 1952. The birds of Prince Albert National Park, Saskatchewan. *Wildlife management bulletin* 4. Canadian Wildlife Service. 83 pp.
- Tammi CE. 1994. Offsite identification of wetlands. In *Applied Wetlands Science and Technology*. ed. D.M. Kent. Boca Raton, FL, Lewis Publishers, pp. 13-34.
- Thitimajshima P. A new modified fuzzy c-means algorithm for multispectral satellite images segmentation. *Proceedings of Geoscience and Remote Sensing Symposium 4*: 1684-1686.
- Tiner RW. 1990. Use of high-altitude aerial photography for inventorying forested wetlands in the United States, *Forest Ecology and Management* 33: 593-604.
- Tolias YA and Panas SM. 1998. Image segmentation by a fuzzy clustering algorithm using adaptive spatially constrained membership functions. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans* 28(3): 359-369.
- Töyrä J, Pietroniro A, and Martz LW. 2001. Multi-sensor hydrologic assessment of a freshwater wetland. *Remote Sensing of Environment* 75: 162-173.
- Treitz PM, Howarth PJ, Filho OR, and Soulis ED. 2000. Agricultural crop classification using SAR tone and texture statistics. *Canadian Journal of Remote Sensing* 26(1): 18-29.
- Turner MG, Gardner RH, and O'Neill RV. 2001. *Landscape Ecology in Theory and Practice*. Springer-Verlag, New York, 401 pp.
- Zadeh L. 1965. Fuzzy sets. *Information and Control* 8: 338-353.
- Zadeh LA. 1973. Outline of a New Approach to the Analysis of Complex Systems and Decision Processes. *IEEE Transactions on Systems, Man, and Cybernetics* 3: 28-44.
- Zhang J and Foody GM. 1998. A fuzzy classification of sub-urban land cover from remotely sensed imagery. *International Journal of Remote Sensing* 19: 2271-2738.
- Zhang L, Liu C, Davis CJ, Solomon DS, Brann TB, and Caldwell LE. 2004. Fuzzy classification of ecological habitats from FIA data. *Forest Science* 50(1): 117-127.

**APPENDIX A:**

**EVALUATION OF THE DATA-DRIVEN WINDOW SIZE TO INCORPORATE  
TEXTURE FEATURES INTO WETLAND MAPPING**

## EVALUATION OF INCORPORATING TEXTURE INTO WETLAND MAPPING FROM MULTISPECTRAL IMAGE

*Wen-Ya Chiu and Isabelle Couloigner*

University of Calgary, Department of Geomatics Engineering, Calgary, Canada  
wchiu@ucalgary.ca, couloigner@geomatrics.ucalgary.ca

### ABSTRACT

Multispectral images have been transformed into Tasseled Cap features to characterize the wetland properties for mapping purpose. The texture derivatives were applied to the brightness, greenness, and wetness using three texture measures based on grey-level co-occurrence matrix method. In this study, the data-driven window size over which texture measures are derived will be determined based on the experimental semivariograms instead of a trial-and-error method. Eight combinations of window sizes have been analyzed to evaluate the benefit of the proposed strategy. A supervised classification based on the maximum likelihood algorithm was applied to the three Tasseled Cap features and to their combination with each texture inputs under different window sizes. Classification accuracy is measured by the overall accuracy for the whole set of classification. User's accuracy and kappa coefficient are used to estimate individual class accuracy. The combination of multiple window sizes from the Tasseled Cap features to derive texture measures for classification purposes is proposed according to the semivariograms. The overall accuracy of the spectral-textural classification shows a 95.5% accuracy higher, than the multispectral classification alone. For the purpose of wetland mapping of the study site, the proposed combinations of multiple window sizes provide wetland class 92.6% accuracy higher than randomly selected identical window sizes.

**Keywords:** Texture analysis, Semivariograms, Grey-level co-occurrence matrix (GLCM), Tasseled Cap features, Wetland, Multispectral image.

### INTRODUCTION

Information about landcover is essential for environmental monitoring. Remotely sensed data supply a current and important source of data for wetland mapping. Image texture quantifies the spatial variation of tone that is related to the distributions of different landcover types on the ground surface. However classical classification algorithms, which applied on a pixel-by-pixel basis, ignore the potential of the spatial information existing between a pixel and its neighbours. To achieve reliable and accurate results in mapping applications, image attributes within a landcover type over its neighbourhood should be characterized. Texture, the intrinsic spatial variability of radiometric data, is a valuable feature to discriminate the different landcover types.

Many approaches were developed for texture analysis. According to the processing algorithms, three major categories, namely, structural, spectral, and statistical methods, are common ways for texture analysis. Grey-level co-occurrence matrix (GLCM), one of the most widely used methods, contains the relative frequencies of the two neighbouring pixels separated by a distance on the image. Several statistical measures (1) such as homogeneity, contrast, and entropy can be computed from the matrix to describe specific textural characteristics. Each texture measure can create a new channel that can be incorporated with spectral features for classification purposes. However a certain number of parameters directly associated with the GLCM method should be considered before computing texture measures. Two important factors, the combinations of texture features and the window size selection, have been examined according to their benefits on the classification accuracy.

Various combinations of texture measures have been tested for different applications such as crop classification in agriculture (2) and forest species classification (3) in nature resources

management. Results showed that incorporating texture features in classification was superior to the classification of the original image. A combination of three or four texture features performs better than the combinations of one or two texture features. But no rules have been recommended for the texture measures selection. The most appropriate combination of texture features depends strongly on the surface properties of the landcover types of interest. Since unique texture patterns were hypothesized to discriminate different landcover types, a proper window size that matches the patch size can extract the textural pattern of this particular landscape. Large window size can capture the spatial patterns of each landcover type better, but may contain more than one land category, which could introduce systematic error. The window should be then small enough to keep the variance low and to maximize the potential for class separability. Previous studies have tried examining several different window sizes (4,5). These trail-and-error methods are time intensive and window size strongly depends on the attributes of the radiometric data for each particular case.

Geospatial techniques utilize spatial information that considers the spectral dependence existing between a pixel and its neighbour. Radiometric data that are highly correlated within a range can be indicated through the semivariogram function (6). The digital number (DN) value of each pixel can be interpreted as a regionalized variable. Meanwhile a data-driven semivariogram provides a method of measuring the spatial dependency of continuously varying phenomena. Recently some techniques have involved geostatistical parameters deduced from the semivariogram function for image classification (7, 8, 9). Although suggestions have been made that the window size should be defined for each particular case, identical windows as fixed square pixel arrays were used for all input channels. The approach of this paper intends to analyze the spatial dependence of radiometric data by geostatistical methods to obtain the suitable window size for the landcover type of interest from data-driven semivariograms. For this purpose multiple window sizes will be used to derive texture measurements from the Tasseled Cap features – brightness, greenness, and wetness - for wetland mapping. The objective of this paper is to assess the benefit of incorporating texture for classification by the proposed methodology.

## **METHODS**

The study site is located within the boundaries of Prince Albert National Park in Northern Saskatchewan, Canada. Approximate coordinates of the study are as follows: 53°45'00"N to 54°00'00"N and 106°00'00"W to 106°25'00"W. The elevation in the area generally decreases from west to east, with elevation varying from 501 to 747 m above sea level. The lowest elevation is Waskesiu Lake (elevation 501 m) while the highest (about 747 m) is in the western part of the site. According to the 7-year Meteorological Service of Canada (MSC) normals for 1996-2002, the mean monthly temperature range from approximately -17.2°C in January to 17.5°C in July and the mean monthly precipitation vary significantly from 80.2 mm in July to only 14.7 mm in November.

Multispectral data was obtained from the Landsat ETM+ sensor. The multispectral image was acquired in August 1999 and processed at level 1G (standard geocoded image resampled to UTM projection). The scene was resampled to 25 m resolution by cubic convolution and a 1086×1086 pixels sub-image was extracted for this study (Figure 1).

### **Image pre-processing**

According to the definition given by the National Wetlands Working Group (1988), wetlands are characterized by three components: soil, vegetation, and water. A Tasseled Cap transformation utilizes a canonical component analysis to decompose multispectral image into three-dimensions: brightness, greenness, and wetness. Wetland pixels can be extracted by using Tasseled Cap transformed images (10) since the brightness channel highlights areas of high reflectance; the greenness channel represents vegetated areas and the wetness channel marks areas that have a high water or moisture content. A Tasseled Cap transformation based on at-satellite reflectance is more appropriate for regional applications where atmospheric correction is not feasible (11). Thus the six cloud-free multispectral bands were chosen to not use atmospheric correction due to the

lack of atmospheric data necessary for running atmospheric correction algorithm. Raw digital numbers were converted to radiance and at-satellite reflectances were calculated according to Landsat 7 Science Data Users Handbook (12).

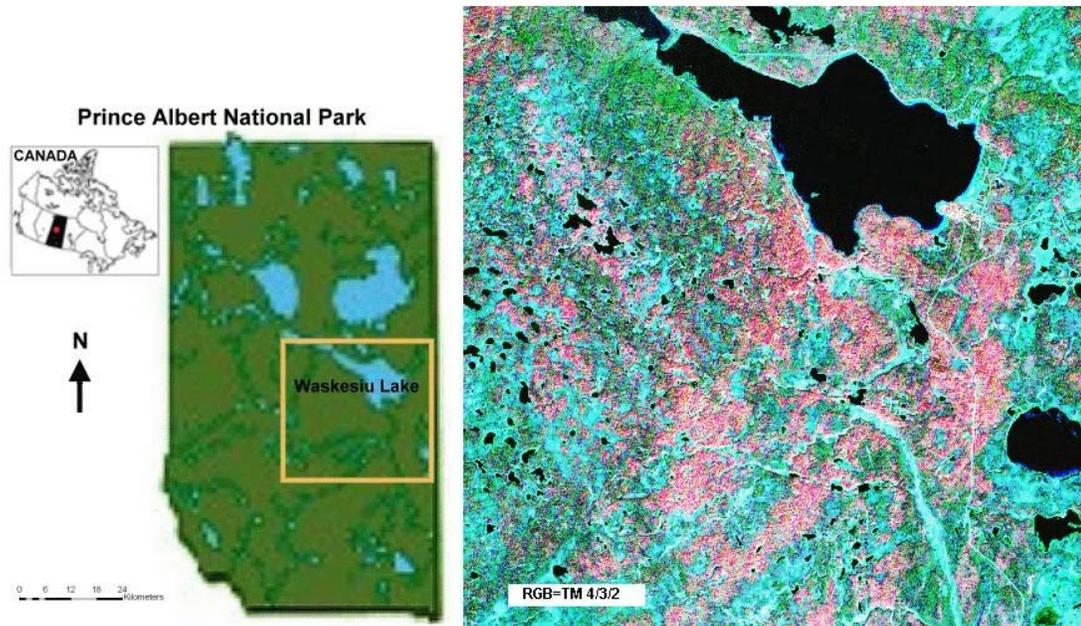


Figure 1: Location map of study area and Landsat-7 composite image (RGB=TM 4/3/2)

### Semivariogram

The semivariogram was employed as a tool to model the spatially varying phenomenon of natural objects. The average change of a property is illustrated by a changing lag and the classical equation can be expressed as follow:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x_i) - Z(x_{i+h})]^2 \quad (1)$$

The experimental semivariance  $\gamma(h)$  is defined as half the average squared difference between values separated by a given lag  $h$ , where  $h$  is a vector in both distance and direction. While  $Z(x_i)$  represents the DN value at a pixel location  $x_i$ ,  $N(h)$  means the total number of pairs. Semivariogram interpretation is usually focused on relating nugget, sill, and range parameters (Figure 2). In this study, lag  $h$  increased by one pixel instead of a real measurement in length unit. Pixels separated within the range are highly correlated with each other. Range can be used as a measure of homogeneity. Automatic fitting of models to semivariograms is the main problem (13) with variogram model-based approaches for texture classification. Since the choice of model may be restricted to certain regions or classes, the coefficient of the model fitting the local variogram may be misleading and unreliable. Modelling was not used to fit the semivariance curves in this study; only experimental values of the semivariograms were used. Semivariograms of four landcover types, wetland, water, dense vegetation, and open vegetation, were examined.

### Image textural channels and classification

Texture analysis, which provides a complementary tool to multispectral studies, has received great attention in image processing. The grey level is assumed to be not a randomly distribution within an image, but associated with structures of landcover types. Texture reflects the local variability of grey levels in the spatial domain and reveals the information about the object

structures in the natural environment. In this study, texture features are computed over a moving window determined by semivariograms. Odd numbers of pixels from 5 to 11 were employed as window size for the three Tasseled Cap features to derive texture measures. In addition combinations of multiple window sizes were also evaluated. The following texture measures were computed from the Tasseled Cap features: mean, variance, and angular second moment (ASM). To evaluate the effects of the proposed window sizes, data were subjected to a maximum likelihood classification algorithm. Accuracy was assessed for wetland mapping.

Although open water can be classified very accurately from the image, some misclassified errors were still resulted from water pixels. Water bodies in the study area varied from few pixels to thousand of pixels due to the natural geographic condition. A pixel-based image classification algorithm may eliminate small size water bodies. To minimize the errors from misclassification of water body, the normalized difference vegetation index (NDVI) was employed to develop an upper threshold, which would identify pixels likely to be open water. One binary map highlighting all pixels within the image being considered open water was created according to this threshold. The map masked out the water-likely pixels to eliminate those pixels during the classification procedure. Therefore, only three classes were considered in the classification process: dense vegetation, open vegetation, and wetland.

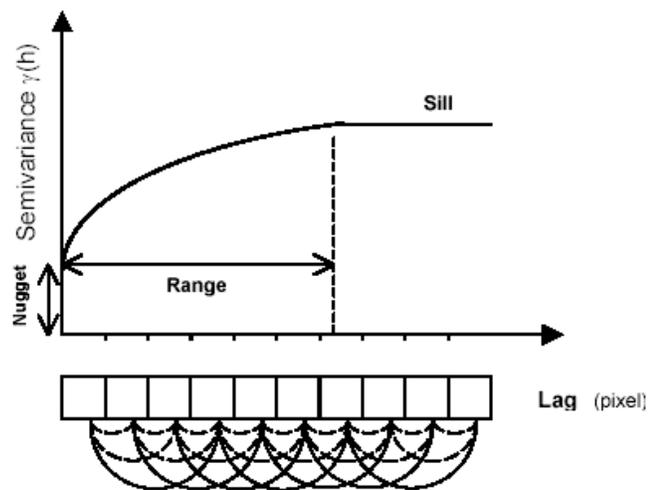


Figure 2: Example semivariogram showing nugget, sill, and range in image application.

## RESULTS

### Analysis of semivariogram behavior

Semivariogram behaviours of four classes, water, wetland, dense vegetation, and open vegetation were examined in the study. An arbitrary size (34x34 pixels) was selected for each training site (Figure 3). The DN statistics (mean  $\pm$  standard deviation) within the geometric size are presented in Table 1.

While dense vegetation has higher brightness and greenness values, its wetness value is lower than for any of the other three classes. On the other hand wetland and open vegetation have similar values in brightness and greenness. Although open vegetation has slightly higher vegetation density than wetland, these similarities are factors that decrease the signature separability of the two landcover types. In this study of spatially autocorrelation, experimental semivariograms of the three Tasseled Cap features were computed within the selected training areas for four directions (*i.e.* NS, EW, NNE, SSE) and for one isotropic curve. The semivariograms have different behaviours due to variations in the correlation patterns of the DN values. Only omnidirectional semivariograms are analyzed to extract the optimum lag distance for deriving texture features in the study.

Table 1: DN statistics (mean  $\pm$  standard deviation) of Tasseled Cap features for four training sites.

Class	Brightness	Greenness	Wetness
Water	50.41 $\pm$ 1.64	-38.30 $\pm$ 2.39	-65.95 $\pm$ 5.22
Wetland	155.95 $\pm$ 12.78	-50.50 $\pm$ 2.77	-271.82 $\pm$ 25.19
Dense Vegetation	180.56 $\pm$ 4.13	-24.79 $\pm$ 2.61	-283.03 $\pm$ 7.62
Open Vegetation	134.47 $\pm$ 13.62	-47.60 $\pm$ 4.90	-230.665 $\pm$ 23.66

Semivariograms computed for each class are unique (Figure 4) and have the following characteristics.

- (1) Water: the semivariograms calculated from brightness, greenness, and wetness are essentially flat, exhibiting little or any spatial correlation for lag distances greater than one pixel. Although the nugget and sill values may varied with the DN data, the analogous curve behaviours can be observed from the semivariograms for the three different data features.
- (2) Wetland: directional and isotropic semivariograms have similar behaviours either for brightness or wetness features. They rose smoothly and reached the sill at a lag of 7 pixels. Semivariogram of greenness feature showed the less of variances among the training classes of the study area, but greatest variance in wetness. The curve of greenness rose steadily upwards up to a local peak at a lag distance of 5 pixels and waved a little bit until it reached the sill at a lag of 11 pixels.
- (3) Dense vegetation: the semivariograms of dense vegetation calculated either from brightness, greenness or wetness features showed periodic forms in four directions. For the isotropic curves of the three Tasseled Cap features, the semivariograms reached a limiting value at a lag of 5, 9, and 5 pixels respectively.
- (4) Open vegetation: although open vegetation showed the greatest of variances among the training classes in brightness and wetness features, the range was slightly higher than that of the wetland class. The semivariograms for both brightness and wetness features rose upwards to a lag distance of 11 pixels, curving to a flat level fairly coincident to the DN variance of the training site. One significant difference should be noticed: although wetland and open vegetation have similar spectral DN values in greenness feature, their semivariograms showed the difference in variance.

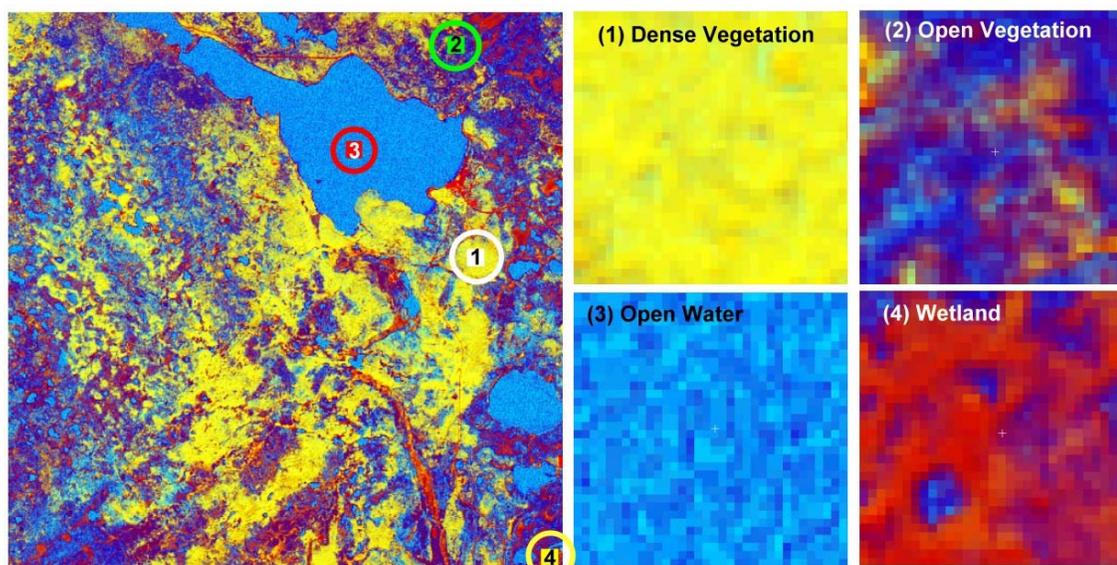


Figure 3: Selected training sites for semivariance calculation. (1) Dense vegetation; (2) open vegetation; (3) open water; and (4) wetland. Composite image is illustrated by Tasseled Cap features in RGB=Brightness/ Greenness/ Wetness.

The semivariograms of the three Tasseled Cap features are used as criteria to determine the optimal window size for deriving texture measurements. A window size for each brightness, greenness, and wetness feature was determined according to the experimental semivariogram signatures of wetland class presented in Figures 4. Therefore, the window size used to derive texture features from the three Tasseled Cap features were 7×7, 11×11, and 7×7 pixels respectively.

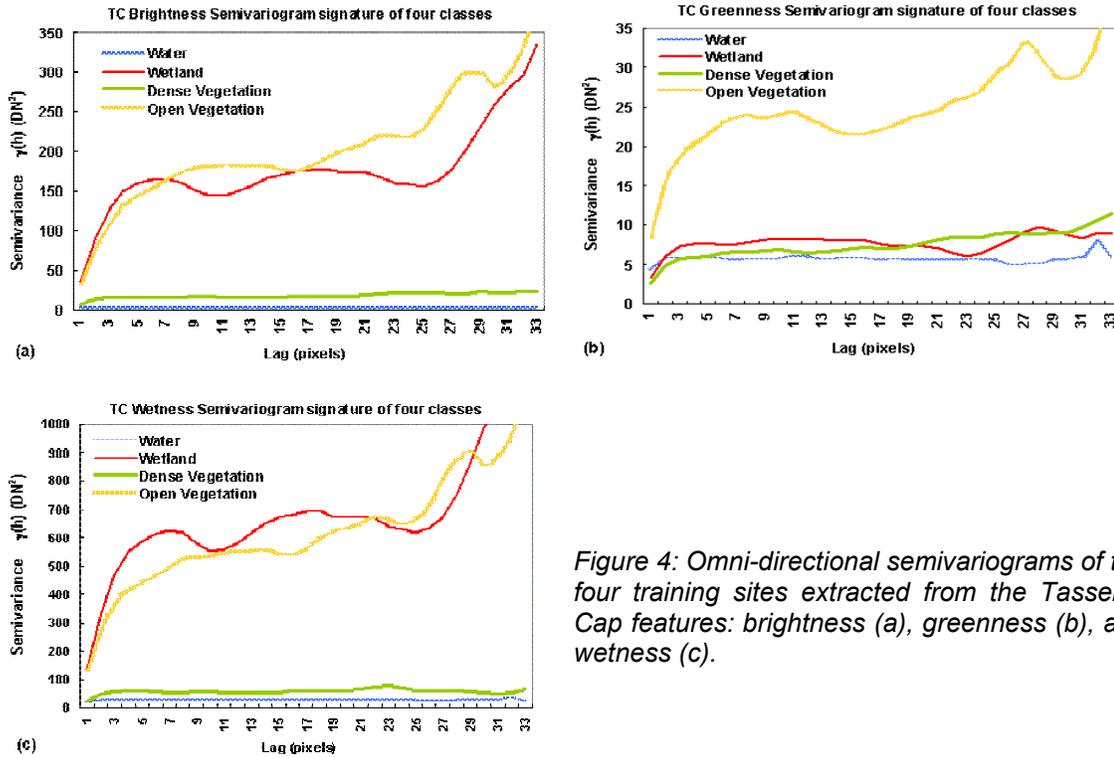


Figure 4: Omni-directional semivariograms of the four training sites extracted from the Tasseled Cap features: brightness (a), greenness (b), and wetness (c).

**Classification**

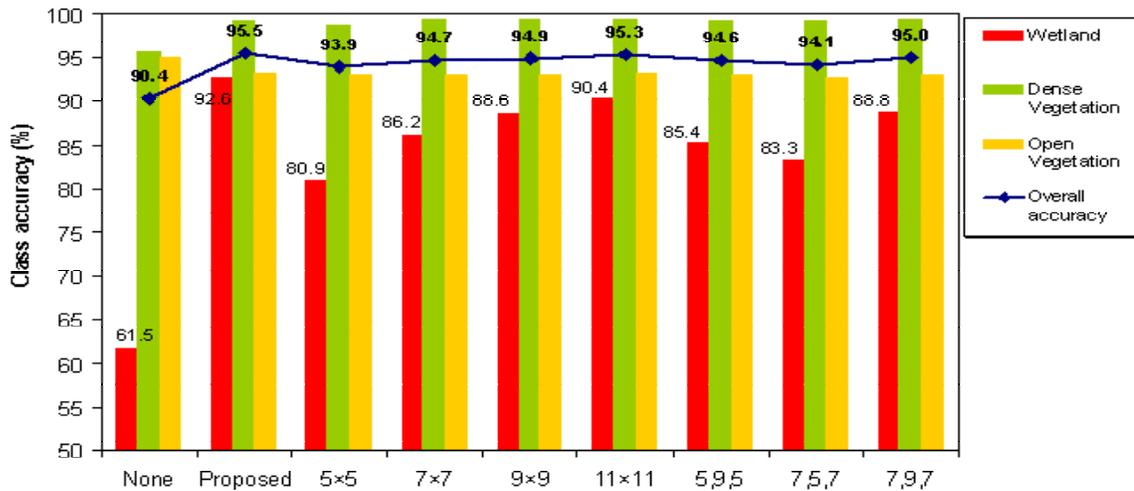


Figure 5: Graph illustrating comparison for classification accuracy for three classes based on the Tasseled Cap features and on different window sizes for texture measures.

The overall accuracy and accuracies of the three classes are illustrated in Figure 5. The three Tasseled Cap features were always used as the input channels for the classification. Texture features derived from different window sizes were compared to assess their influence on wetland mapping. Consequently the strategy utilizing semivariograms to determine the optimal window size was also evaluated. For this purpose, four identical window sizes from 5×5 to 11×11 and four multiple sizes combinations were investigated.

The comparison of spectral and spectral-textural classification accuracies indicates that introducing texture features into classification could provide a better result than spectral data alone. The overall accuracy increases by 4% at least (Figure 5). The proposed method predicting the preferred window sizes for deriving texture features as 7×7 for brightness, 11×11 for greenness, and 7×7 for wetness shows a highest overall accuracy of 95.5%. Incorporation of the texture features into the classification of the Landsat TM data improved the accuracy of the wetland class. The accuracy of wetland class improved from 61.5% using spectral bands to 92.6% using a combination of spectral bands and texture features. Wetlands in the study area are fragmentary and distributed around small water bodies or in the river riparian. Nearly identical spectral reflectances between vegetation types cause the low signature separability between the wetland and open vegetation classes. Texture features therefore provide additional information to distinguish the insignificant differences in the spectral signature.

Table 2: Summary results of accuracy assessment of spectral and textural classification.

Band combination <sup>a</sup>	Overall accuracy	User's accuracy (%)			Kappa coefficient		
		Wetland	Dense Vegetation	Open Vegetation	Wetland	Dense Vegetation	Open Vegetation
Spectral alone	90.4	61.5	95.7	95.0	0.56	0.93	0.90
Spectral-textural							
Proposed size	95.5	92.6	99.4	93.3	0.92	0.99	0.86
5×5	93.9	80.9	98.8	93.1	0.78	0.98	0.86
7×7	94.7	86.2	99.4	92.9	0.84	0.99	0.86
9×9	94.9	88.6	99.4	92.9	0.87	0.99	0.86
11×11	95.3	90.4	99.4	93.3	0.89	0.99	0.86
5,9,5	94.6	85.4	99.2	93.1	0.83	0.99	0.86
7,5,7	94.1	83.3	99.2	92.7	0.81	0.99	0.85
7,9,7	95.0	88.8	99.4	93.0	0.87	0.99	0.86

<sup>a</sup>The window sizes used to derive the texture features from Tasseled Cap transformations are represented by numbers.

User's accuracy and kappa coefficients for spectral and spectral-textural classifications are computed to estimate the accuracy of individual class in Table 2. The table also illustrates the comparison between random selected window size and the one predicted by the semivariogram analysis. The window size is responsible for most of the variability in the classification because a significantly correlation between class accuracy and selected window sizes used for deriving texture features is observed. However this trend was not found in dense vegetation class and open vegetation class. The accuracies of these two classes do not show much variation between different combinations of textural classification. Since the wetlands in the study area are fragmentary and vary in different sizes, semivariogram captures the spatial correlation by predicting an appropriate lag distance for deriving texture measures. The kappa coefficient of the wetland class evaluated by adding proposed texture channels is 0.92, which is higher than the other randomly selected 5×5 window to 11×11 window size. When examining the semivariogram of the greenness feature (Figure 4.b), the variance reached a local peak at a lag of 5 pixels. By using a 5×5 window for greenness, it shows a lower kappa coefficient for the wetland class. The result indicates that a small window size may lose some spatial information of the specific class. However template window size at 11 pixels, which is the range value related to the sill of the

semivariogram, can provide a better classification result. The utilization of multiple window sizes (*i.e.* 7×7 for brightness, 11×11 for greenness, and 7×7 for wetness) is proposed in the classification. Multiple window sizes can retain the integrity of the small windows while reducing the effects of noise encountered with large windows. The result also illustrated the capability of improving the accuracy by applying this concept.

## CONCLUSIONS

The overall accuracy indicated that the incorporation of texture measures into multispectral data could improve the classification result by 5% for this case study. Window size for deriving texture features is a factor contributing to classification accuracy. The study addresses the need to determine the data-driven window size predicted by the range of semivariogram for specific class inspection. According to the semivariograms of the target class, the resulting range parameter can provide superior discrimination and correlation results compared to those obtained using randomly selected identical windows. The proposed method shows the capability in increasing wetland class discrimination from 61.5% to 92.6%. This is a time-effective strategy that can be used to optimize texture derivations of remotely sensed imagery. Future study will examine if the pixels of these fragmentary classes can be grouped as segments and then take the advantage of texture analysis for identification of landcover units.

## ACKNOWLEDGEMENTS

The authors gratefully acknowledge the assistance of Natural Resources Canada and the University of Calgary for supplying the Landsat 7 images and the topographic data used in the research. We further appreciate critical comments provided by two anonymous reviewers and the editor.

## REFERENCES

- 1 Haralick, R.M. and K. Shanmugan. 1974. Combined spectral and spatial processing of ERTS imagery data. Remote Sensing of Environment. 3: 3-13.
- 2 Treitz, P.M., P.J. Howarth, O.R. Filho, and E.D. Soulis. 2000. Agricultural crop classification using SAR tone and texture statistics. Canadian Journal of Remote Sensing. 26(1): 18–29.
- 3 Franklin, S.E., R.J. Hall, L.M. Mosdal, A.J. Maudie, and M.B. Lavigne. 2000. Incorporating texture into classification of forest species composition from airborne multispectral images. International Journal of Remote Sensing. 21(1): 61-79.
- 4 Marceau, D.J., P.J. Howarth, J.M. Dubois, and D.J. Gratton. 1990. Evaluation of the Grey-Level Co-Occurrence Matrix method for land-cover classification using SPOT imagery. IEEE Transactions on Geoscience and Remote Sensing. 28(4): 513-518.
- 5 Arzandeh, S. and J. Wang. 2002. Texture evaluation of RADARSAT imagery for wetland mapping. Canadian Journal of Remote Sensing. 28(5): 653-666.
- 6 Curran, P.. 1988. The semivariogram in remote sensing: an introduction. Remote Sensing of Environment. 24: 493-507.
- 7 Miranda, F.P. and J.R. Carr. 1998. Semivariogram textural classification of JERS-1 (Fuyo-1) SAR data obtained over a flooded area of the Amazon rainforest. International Journal of Remote Sensing. 19(3): 549-556.
- 8 Atkinson, P.M. and P. Lewis. 2000. Geostatistical classification for remote sensing: an introduction. Computers and Geosciences. 26(3): 361-371.
- 9 Chica-Olmo, M. and F. Abarca-Hernandez. 2000. Computing geostatistical image texture for remotely sensed data classification. Computers and Geosciences. 26(3): 373-383.

- 10 Civco, D.L. and J.D. Hurd. 1999. A hierarchical approach to land use and land cover mapping using multiple image types. Proc. 1999 ASPRS Annual Convention, Portland, OR. p. 687-698.
- 11 Huang, C., B. Wylie, L. Yang, C. Homer, and G. Zylstra. 2001. Derivation of a tasseled cap transformation based on Landsat 7 at-satellite reflectance. USGS EROS Data Center.
- 12 Irish, R. R., 2000, Landsat 7 science data user's handbook, Report 430-15-01-003-0, National Aeronautics and Space Administration.
- 13 Franklin, S.E., M.A. Wulder, and M.B. Lavigne. 1996. Automated derivation of geographic window sizes for use in remote sensing digital image texture analysis. Computers and Geosciences. 22(6): 665-673.