Remote Sensing Techniques for Mangrove Mapping

Chaichoke Vaiphasa

Promotors:
Prof. Dr. A.K. Skidmore
Professor of Vegetation and Agricultural Land Use Survey
International Institute for Geo-information Science and Earth Observation (ITC), Enschede and Wageningen University, the Netherlands

Prof. Dr. H.H.T. Prins
Professor of Tropical Nature Conservation and Vertebrate Ecology
Wageningen University, the Netherlands

Co-promotor:
Dr. W.F. de Boer
Assistant Professor Resource Ecology Group
Wageningen University, the Netherlands

Examining Committee:
Prof. Dr. M.E. Schaepman
Wageningen University
Wageningen

Prof. Dr. S.M. de Jong
Utrecht University
Utrecht

Dr. Itthi Trisirisatayawong
Chulalongkorn University
Bangkok, Thailand

Prof. Dr. A.M. Cleef
University of Amsterdam
Amsterdam

Remote Sensing Techniques for Mangrove Mapping

Chaichoke Vaiphasa

Thesis

To fulfil the requirements for the degree of Doctor
on the authority of the Rector Magnificus of Wageningen University,
Prof. Dr. M.J. Kropff,
to be publicly defended on
Tuesday 31 January, 2006 at 15:00 hrs
in the auditorium of ITC, Enschede, The Netherlands.

ISBN: 90-8504-353-0
ITC Dissertation Number: 129
International Institute for Geo-information Science & Earth Observation,
Enschede, The Netherlands

Table of Contents

Abstract

Mangroves, important components of the world's coastal ecosystems, are threatened by the expansion of human settlements, the boom in commercial aquaculture, the impact of tidal waves and storm surges, etc. Such threats are leading to the increasing demand for detailed mangrove maps for the purpose of measuring the extent of the decline of mangrove ecosystems. Detailed mangrove maps at the community or species level are, however, not easy to produce, mainly because mangrove forests are very difficult to access. Without doubt, remote sensing is a serious alternative to traditional field-based methods for mangrove mapping, as it allows information to be gathered from the forbidding environment of mangrove forests, which otherwise, logistically and practically speaking, would be extremely difficult to survey. Remote sensing applications for mangrove mapping at the fundamental level are already well established but, surprisingly, a number of advanced remote sensing applications have remained unexplored for the purpose of mangrove mapping at a finer level. Consequently, the aim of this thesis is to unveil the potential of some of the unexplored remote sensing techniques for mangrove studies. Specifically, this thesis focuses on improving class separability between mangrove species or community types. It is based on two important ingredients:

(i) the use of narrow-band hyperspectral data, and
(ii) the integration of ecological knowledge of mangrove-environment relationships into the mapping process.

Overall, the results of this study reveal the potential of both ingredients. They show that delicate spectral details of hyperspectral data and the spatial relationships between mangroves and their surrounding environment help to improve mangrove class separability at the species level. Despite the optimism generated by the overall results, it was found that appropriate data treatments and analysis techniques such as spectral band selection and noise reduction were still required to harness essential information from both hyperspectral and ecological data. Thus, some aspects of these data treatments and analysis techniques are also presented in this thesis. Finally, it is hoped that the methodology presented in this thesis will prove useful and will be followed for producing mangrove maps at a finer level.

Samenvatting

Mangrove, een belangrijk onderdeel van ecosystemen in (tropische) kustgebieden wereldwijd, wordt bedreigd door de uitbreiding van woongebieden, de groei van commerciële aquacultuur, de invloed van getijden en stormvloeden, enzovoorts. Zulke bedreigingen leiden tot een toenemende vraag naar gedetailleerde mangrovekaarten om de mate waarin mangrove-ecosystemen afnemen te kwantificeren. Omdat mangrovebossen erg moeilijk toegankelijk zijn, is het moeilijk aan de vraag naar mangrovekaarten op gedetailleerd niveau of op soortniveau te voldoen. *Remote sensing* is een alternatief voor het in kaart brengen van mangrove met de traditionele veld-*survey*methode in gebieden die logistiek lastig zijn te onderzoeken.

Het toepassen van remote sensing voor het in kaart brengen van mangrove is op kleinschalig niveau al in gebruik, maar het is verbazingwekkend dat een aantal meer geavanceerde remote sensing-technieken om mangrove in kaart te brengen op een grootschaliger niveau niet zijn uitgeprobeerd. Dit proefschrift laat zien wat de mogelijkheden zijn van enige nog niet toegepaste technieken voor de bestudering/kartering van mangrove. Dit proefschrift concentreert zich op het verbeteren van het onderscheid tussen mangrovesoorten of plantengemeenschapstypen. Twee belangrijke bestanddelen vormen de basis:

(i) het gebruik van hyperspectrale data, en (ii) het integreren van ecologische kennis van mangrove-milieu relaties bij het proces van het in kaart brengen van mangrovegebieden.

De resultaten van deze studie tonen de mogelijkheden van beide ingrediënten. De samenhang van subtiele spectrale details en ruimtelijke relaties tussen mangrove en de omgeving versterken de mogelijkheid mangrove op soortniveau te onderscheiden. Ondanks deze positieve uitkomst blijkt het nodig te blijven de juiste datatransformaties en analysetechnieken toe te passen. Bijvoorbeeld bandselectie en *noise*-reductie zijn nog steeds nodig om onontbeerlijke informatie uit hyperspectrale en ecologische data af te leiden. Enige aspecten van deze technieken kunnen daarom niet achterwege blijven in dit proefschrift.

Tenslotte bestaat de hoop dat de methodologie die in dit proefschrift wordt beschreven bruikbaar blijkt te zijn en zal worden toegepast bij het karteren van mangrove op een gedetailleerder niveau.

Acknowledgements

In addition to the financial support from the Thai government lottery office, this thesis would be impossible without people who supported me and here they are…

Thanks must first of all go to Prof. Andrew Skidmore for his supervision. I have learnt a lot from you, Andrew. Next, I thank Henk van Oosten for his help on computer programming and, of course, his friendship that I will never forget. Then, thank Dr. Fred de Boer for his invaluable advice on mangrove ecology. I also would like to thank Dr. Suwit Ongsomwang for the data and his kind advice on many occasions. I thank my boss, Dr. Itthi Trisirisatayawong, for his understanding and support throughout my study at ITC and Prof. Sanit Aksronkoae for the inspiration he gave me to pursue this mangrove study. Many thanks to Khun Supawadee Dararucha for her great support in Pak Phanang.

Thanks Dr. Robert Duin of Delft University of Technology for the short course on pattern recognition, Prof. Herbert Prins, Dr. Jan de Leeuw, Dr. Chris Mannaerts, Dr. Yousif Hussin, Dr. Nopparat Bamrongrugsa, Karl Grabmaier, Sokhon Phem, Sangob Panitchart, Anuchit Rattanasuwan, Korn Manassrisuksi, and Wirote Teeratanatorn for their great advice on a number of occasions.

Thanks everyone at the forest research station in Pak Nakorn for their very kind supports. I will never forget the green boat used during the field works. It's been a great time there. Everyone at the Pak Phanang water gate has been excellent too. They are very supportive.

Thanks Loes Colenbrander for her invaluable help at the final stage and Esther Hondebrink for her great secretary skills that keep my PhD life running smoothly.

Thanks all my true friends who always support me and love from my family that helps me through.

And, Mon, you know, this thesis could have been nothing without you.

Chapter 1

General Introduction

1.1 Remote sensing for mangrove studies

Mangroves are reported to have covered up to 75% of the world's tropical coastlines (Spalding et al., 1997). The ecological value of these tropical mangroves is acknowledged in many respects, for example, in (i) protecting the coastline from tidal waves and storm surges, (ii) acting as biological filters in polluted coastal areas, (iii) supporting aquatic food chains, and (iv) shielding a large number of juvenile aquatic organisms (Lugo and Snedaker, 1974; Hogarth, 1999; Linneweber and de Lacerda, 2002; Barbier and Sathiratai, 2004). Unfortunately, mangroves in general are in serious decline owing to the expansion of human settlements, the boom in commercial aquaculture, the impact of tidal waves and storm surges, etc. (Linneweber and de Lacerda, 2002; Barbier and Sathiratai, 2004). International organizations and government agencies in several countries are, for this reason, urgently implementing mapping and monitoring programmes to measure the extent of the decline of these important ecosystems (Ramsar Convention, 1971; Green et al., 2000; Linneweber and de Lacerda, 2002; Barbier and Sathiratai, 2004). Without doubt, remote sensing is a serious alternative to the traditional field monitoring for large-scale tropical mangrove management (Blasco et al., 1998). This is mainly because remote sensing technology allows information to be gathered from the environment of mangrove forests, which otherwise, logistically and practically speaking, would be very difficult to survey. Remote sensing applications in mangrove management come in three categories and are used for three main purposes: (i) resource inventory, (ii) change detection, and (iii) the selection and inventory of aquaculture sites (Green et al., 2000). These applications are based on a number of instruments on both aeroplane and satellite platforms, including visible and infrared photographic cameras (Sulong et al., 2002; Verheyden et al., 2002), video recorders (Everett et al., 1996), synthetic aperture radar (Aschbacher et al., 1995; Held et al., 2003), and multispectral and hyperspectral sensors (Ramsey III and Jensen, 1996; Gao, 1999; Green et al., 2000; Demuro and Chisholm, 2003; Held et al., 2003).

1.2 Hyperspectral remote sensing for mangrove discrimination

Hyperspectral data are a form of spectral records that contain 100 bands or more throughout the visible, near-infrared, mid-infrared, and thermal infrared portions of the spectrum. Each band possesses a 10 nm, or

narrower, bandwidth. Hyperspectral sensors can be used for discriminating among earth surface features that have distinct diagnostic absorption and reflection characteristics over narrow wavelength intervals and which are lost within the relatively coarse bandwidths of conventional multispectral sensors (Lillesand and Kiefer, 2000).

Hyperspectral technology has already been successfully established in the field of vegetation research (Green et al., 1998; Asner et al., 2000; Cochrane, 2000; Green et al., 2000; Kruse et al., 2000; Curran et al., 2001; Soukupová et al., 2002; Goel et al., 2003; Hirano et al., 2003; Mutanga et al., 2003; Schmidt and Skidmore, 2003; Schuerger et al., 2003; Zarco-Tejada et al., 2004). This is because hyperspectral data contain information that is linked to important biochemical properties of plants (Gates et al., 1965; Hoffer, 1978; Peterson and Hubbard, 1992; Kokaly, 2001; McDonald, 2003). The study of the quality of tropical pastures for animal grazing (Mutanga et al., 2003), the use of hyperspectral sensors for detecting zinc stress in plants (Schuerger et al., 2003), a revised method for lignin detection (Soukupová et al., 2002), and the extraction of crop biophysical parameters (Goel et al., 2003) are just a few examples highlighting the recent application of hyperspectral technology. More importantly, recent reports confirm that hyperspectral data can be used for investigating vegetation dynamics and the ecology of species, as the data have the potential for discriminating terrestrial plants at the species level (Cochrane, 2000; Schmidt and Skidmore, 2003).

Although hyperspectral remote sensing looks promising in the arena of vegetation applications, including species-level discrimination, the hyperspectral research on mangroves published to date (Green et al., 2000; Demuro and Chisholm, 2003; Held et al., 2003; Hirano et al., 2003) is inconclusive as to whether the hyperspectral technology could be used for studying tropical mangroves in finer detail, particularly at the species level. A most unfortunate case is the multi-sensor study carried out by Held et al. (2003). These authors should have been the first to ascertain whether an on-board hyperspectral sensor could be used for classifying mangrove species, but the hyperspectral images used were distorted by a high percentage of cloud cover. Furthermore, the other researchers (Green et al., 2000; Demuro and Chisholm, 2003; Hirano et al., 2003) could not explore thoroughly the capability of hyperspectral

data for discriminating mangroves at the species level because their study sites were dominated by a few mangrove species only.

As a result, the first goal of this thesis is to investigate further the potential of hyperspectral technology for discriminating mangroves at the species level. A laboratory experiment is set up to test whether hyperspectral data contain adequate information for mapping mangroves at the species level. This laboratory study is intended to be a prerequisite for the future application of airborne and satellite hyperspectral sensors.

1.3 Burdens of hyperspectral data

1.3.1 Dimensionality problems

Even though the application of hyperspectral remote sensing in vegetation studies is popular (Green et al., 1998; Asner et al., 2000; Cochrane, 2000; Green et al., 2000; Kruse et al., 2000; Curran et al., 2001; Soukupová et al., 2002; Goel et al., 2003; Hirano et al., 2003; Mutanga et al., 2003; Zarco-Tejada et al., 2004), extracting, analysing or classifying hyperspectral information effectively without appropriate image processing algorithms is not straightforward, owing to high dimensionality. With respect to image classification, high-dimensional data trigger the phenomenon known as "the curse of dimensionality". This means that the complexity caused by high dimensionality undermines the precision of the estimates of class distribution (e.g., mean and covariance) in the feature space. The imprecise class estimates then result in low output accuracy (Bellman, 1961; Hughes, 1968). Furthermore, this phenomenon results in the need for more training samples in order to construct better estimates of class models, thereby dramatically increasing the cost of the field survey. Thus, sophisticated methods, either feature extraction or feature selection, are inevitably required to handle the problems of hyper-dimensionality before performing image classification.

A number of feature extraction techniques have already been employed to ease the problems of high dimensionality. For example, Fisher's linear discriminant analysis (Duda and Hart, 1973), principal component analysis (PCA) (Anderson, 1984), canonical analysis (Richards, 1986), decision boundary feature extraction (Lee and Landgrebe, 1993), orthogonal subspace projection (OSP) (Harsanyi and Chang, 1994), and linear constrained distance-based discriminant analysis (LCDA) (Du and

Chang, 2001) are typical data extraction methods that aim to transform original hyperspectral data into a smaller feature space. Despite their three common weaknesses – (i) the loss of information during extraction, (ii) the requirement for a number of field samples to avoid singularity problems of covariance matrix inversion, and (iii) the assumption of data distribution – it has been proved that these techniques help to gain an acceptable level of classification accuracy by increasing the precision of class distribution estimates of the transformed data (Lee and Landgrebe, 1993; Harsanyi and Chang, 1994; Wu and Linders, 2000; Du and Chang, 2001; Flink et al., 2001; Gong et al., 2002; Metternicht and Zinck, 2003).

With respect to feature selection, it is possible to reduce significantly the number of data dimensions (spectral bands) in an *ad hoc* fashion when prior knowledge of specific spectral properties of the objects under study is explicit (Schmidt and Skidmore, 2003; van Niel et al., 2003). In this case, experienced analysts manually select a combination of useful bands that contain most of the relevant spectral information and leave out the rest. In many other cases, however, the spectral characteristics are not known beforehand. The analysts then have to select the best combination of bands by using separability indices as selection criteria (Kailath, 1967; Swain and King, 1973; Swain and Davis, 1978; Mausel et al., 1990; Bruzzone and Serpico, 2000). Specifically, the separability index is calculated for every possible band combination, and then the band combination that possesses the highest index score is selected. Thus, this method involves tedious and exhaustive, if not impossible, search efforts to find the best band combination. In practice, the quality of selection results is usually compromised by the adoption of sub-optimal (less exhaustive) search methods that reduce calculation time (Kavzoglu and Mather, 2002; Pekkarinen, 2002; Yu et al., 2002; Ulfarsson et al., 2003; Mutanga and Skidmore, 2004). Popular sub-optimal search schemes that are regularly used in remote sensing applications include (i) sequential forward selection algorithms (Pudil et al., 1994), (ii) branch and bound search (Narendra and Fukunaga, 1997), and (iii) genetic search algorithms (Holland, 1975; Goldberg, 1989).

With respect to the use of genetic search algorithms (GA), it has already been proved that GA-based band selectors perform better than many other popular band selection algorithms (e.g., exhaustive search, branch and bound search, and sequential forward selection) (Siedlecki and Sklansky, 1989). Rigorous comparison has been made, using a synthetic

error model instead of real remotely sensed data so as to eliminate the variables (e.g., sample size, the number of spectral bands, the number of classes of interest) that could bias the outcome. In addition, mounting evidence of the success of GA-based band selectors in real-life remote sensing applications can be found in recent literature with regard to (i) selecting a subset of multiple sensor/date data for image classification (Lofy and Sklansky, 2001; Kavzoglu and Mather, 2002; Ulfarsson et al., 2003), (ii) selecting spectral bands that relate to physio-chemical characteristics of plants and soils (Fang et al., 2003; Kooistra et al., 2003; Cogdill et al., 2004), and (iii) selecting a spectral subset of hyperspectral data for image classification (Yu et al., 2002).

Nevertheless, the class information used in the above studies (Lofy and Sklansky, 2001; Kavzoglu and Mather, 2002; Yu et al., 2002; Ulfarsson et al., 2003) for testing the performance of the GA-based band selector is broad (i.e., USGS level I or II (Anderson et al., 1976)). This means that each class possesses distinct spectral characteristics, and it is relatively easy for GA to find spectral bands that maintain high spectral separability between classes. In none of these studies has the band selector been tested on class information that possesses very similar spectral characteristics (e.g., species-level data). The question of whether GA can deal with such complexity therefore remains. As a result, the second goal of this thesis is to test whether the GA-based band selector can be used for selecting a meaningful subset of spectral bands that maintains spectral separability between species classes. The test data used consist of spectrometer records of very high dimensionality, comprising 2151 bands of leaf spectra of 16 tropical mangrove species.

1.3.2 Noise levels

Another important problem that arises when using hyperspectral data, and one which is treated in this thesis, is low signal-to-noise ratios. It is normally found that the noise level in hyperspectral data is high. This is because their narrow bandwidth can capture only very little energy and this is sometimes overcome by the self-generated noise inside the sensors. Physical disturbances such as the fluctuation of light illumination and atmospheric states can make the situation even worse, as the disturbances decrease the precision of spectral signals recorded by the sensor (Landgrebe, 1997; Lillesand and Kiefer, 2000).
Spectral smoothing and aggregating techniques, including both linear and non-linear methods, are the tools most popularly applied, mainly for

removing noise, in a large number of modern hyperspectral remote sensing studies (Tsai and Philpot, 1998; Gong et al., 2001; Ben-Dor et al., 2002; Strachan et al., 2002; Andréfouët et al., 2003; Hochberg and Atkinson, 2003; Schmidt and Skidmore, 2003; Vaughan et al., 2003; Yamano et al., 2003; Zarco-Tejada et al., 2003; Castro-Esau et al., 2004; Foody et al., 2004; Imanishi et al., 2004; le Maire et al., 2004; Meroni et al., 2004; Rees et al., 2004; Schmidt and Skidmore, 2004; Smith et al., 2004; Thenkabail et al., 2004; Whiting et al., 2004). All these studies, however, use subjective *ad hoc* inspections as their measure for selecting appropriate smoothing methods. In other words, they do not use any strict optimizing criteria for selecting suitable smoothing filters in their studies.

As a result, it is hypothesized that the subjective measure is not the most appropriate way of selecting smoothing criteria because it causes changes to some statistical properties (e.g., mean) of the original data, which could possibly affect the results of subsequent parametric analyses that utilize statistical properties of the data (e.g., Jeffries-Matusita distance, maximum likelihood classification). Instead, a more objective approach should be used as a measure for selecting an appropriate smoothing method in order to minimize the disturbance. As a result, the third goal of this study is to find a more appropriate approach to replace the *ad hoc* measure.

1.4 Utilizing mangrove-environment relationships

Although remote sensing applications for mangrove mapping at the fundamental level (e.g., regional areas) are well established (Aschbacher et al., 1995; Ramsey III and Jensen, 1996; Gao, 1999; Green et al., 2000; Sulong et al., 2002; Verheyden et al., 2002; Demuro and Chisholm, 2003; Held et al., 2003), there is an increasing demand for mangrove maps at a finer level. For example, detailed mangrove maps at the community or species level are needed for studying mangrove ecosystems and their diversity.

A number of advanced remote sensing techniques, however, remain unexplored for the purpose of mangrove mapping at a finer level. The use of ancillary data for mapping vegetation (Skidmore et al., 1997a; Skidmore et al., 1997b; Lehmann and Lenz, 1998; Berberoglu et al., 2004; Comber et al., 2004; Schmidt et al., 2004) is one such technique that has been successfully employed to map other plants but has never been tested on mangroves. In most cases, ancillary data are extracted

from plant-environment relationships, and then incorporated with remotely sensed data into the mapping model in order to improve the quality of the final map. Ancillary data can be incorporated at three different stages: before (pre-classification), during, and after (post-classification) the mapping process (Lillesand and Kiefer, 2000). Two of the most popular models used for integrating extra plant-environment information into the mapping process are the non-parametric model of artificial neural networks (Skidmore et al., 1997a; Lehmann and Lenz, 1998; Berberoglu et al., 2004) and the inference engine of expert systems (Skidmore et al., 1997b; Comber et al., 2004; Schmidt et al., 2004).

Similar to many other plants, mangroves have strong relationships with the surrounding environment. The occurrence of mangrove species at a certain location is related to surrounding ecological gradients such as elevation, tidal inundation, water salinity and soil pH (Macnae, 1968; Clough, 1982; Semeniuk, 1983; Tomlinson, 1994; Hogarth, 1999). In other words, mangrove species are likely to grow within their own niches. In many cases, this phenomenon causes strip-like patterns (i.e., mangrove zonations) parallel to the tide lines that are usually found in tropical mangrove forests (Tomlinson, 1994; Hogarth, 1999; Vilarrubia, 2000; Satyanarayana et al., 2002). Thus, it is hypothesized that these quantifiable spatial relationships between mangroves and the environment can be exploited for mangrove mapping at a finer level. Consequently, the fourth and final goal of this thesis is to test this hypothesis.

1.5. Objectives of the study

The main objective of this study is to unveil the potential of some of the unexplored techniques of remote sensing for mangrove studies. Specifically, this thesis focuses on improving the class separability between mangroves, based on two important ingredients: (i) the use of narrow-band hyperspectral data, and (ii) the integration of ecological knowledge of mangrove-environment relationships into the mapping process. The main objective can be divided into four sub-objectives:

(1) to demonstrate the potential of hyperspectral technology for discriminating mangroves at the species level
(2) to test whether a form of genetic algorithms can be used for selecting a meaningful subset of spectral bands that maintains spectral separability between mangrove species

(3) to investigate one of the most popular methods of reducing noise levels in hyperspectral data (i.e., spectral smoothing), as well as propose a technique for selecting an appropriate smoothing filter for the data at hand
(4) to test whether mangrove-environment relationships can be exploited in order to improve mapping accuracy.

1.6. Outline of the thesis

This thesis is composed mainly of two conference papers and five journal manuscripts. One of the manuscripts has already been published in an international peer-reviewed journal; two more have been accepted pending minor revisions; the last two are under review. The conference papers and manuscripts are organized into four separate chapters (Chapters 2 to 5), each of which stands alone and deals with one research objective. These four chapters, the core of the thesis, are accompanied by an introduction (Chapter 1) and a synthesis (Chapter 6). This format best suits the content of this thesis as each core chapter takes on one different aspect branching off from the main goal. Although some overlapping with regard to method description and illustrations is inevitable, this is considered justifiable as each chapter can be read individually without losing the context.

In Chapter 2, the study demonstrates the potential of hyperspectral technology for discriminating mangroves at the species level. This is a laboratory investigation to see whether hyperspectral data contain adequate spectral information for discriminating mangroves at the species level. This laboratory study is intended to be a prerequisite for the future application of airborne and satellite hyperspectral sensors.

Subsequently, in the next two chapters, two major technical concerns with respect to the use of hyperspectral data are addressed. Chapter 3 demonstrates the possibility of using genetic algorithms for selecting spectral band subsets of species-level data of very high dimensionality. Chapter 4 investigates the application of spectral smoothing techniques in existing hyperspectral remote sensing studies, which has normally been conducted in an *ad hoc* fashion. An alternative to the *ad hoc* approaches is then proposed.

Chapter 5 explores one of the techniques, an expert system, that has been successfully employed to map other vegetation but has never been tested

on mangroves. Specifically, a Bayesian expert system is used as a post-classifier to exploit spatial relationships between mangroves and environmental gradients in order to improve mapping accuracy.

Chapter 6 summarizes and synthesizes the main ideas of the thesis, and recommends some further studies.

Chapter 2*

Hyperspectral Data for Mangrove Discrimination

* Vaiphasa, C., Ongsomwang, S., Vaiphasa, T., Skidmore, A.K. (2005). Tropical mangrove species discrimination using hyperspectral data: a laboratory study. Estuarine, Coastal, and Shelf Science 65, 371-379.

Abstract

The aim of this study is to test whether canopy leaves spectra of various tropical mangrove species measured under laboratory conditions contain sufficient spectral information for discriminating mangroves at the species level. This laboratory-level study is one of the most important prerequisites for the future use of airborne and satellite hyperspectral sensors for mangrove studies. First, spectral responses of 16 Thai tropical mangrove species (2151 spectral bands between 350 nm and 2500 nm) were recorded from the leaves, using a spectrometer under laboratory conditions. Next, the mangrove spectra were statistically tested using one-way ANOVA to see whether they significantly differ at every spectral location. Finally, the spectral separability between each pair of mangrove species was quantified using the Jeffries-Matusita (J-M) distance measure. It turned out that the 16 mangrove species under study were statistically different at most spectral locations, with a 95% confidence level (p-value < 0.05). The total number of spectral bands that had p-values less than 0.05 was 1941, of which 477 bands had a 99% confidence level (p-value < 0.01). Moreover, the J-M distance indices calculated for all pairs of the mangrove species illustrated that the mangroves were spectrally separable except the pairs that comprised the members of Rhizophoraceae. Although the difficulties in discriminating the members of Rhizophoraceae are expected, the overall result encourages further investigations into the use of on-board hyperspectral sensors to see whether mangrove species can be separated when the difficulties of the field conditions are taken into account.

Keywords: data reduction; mangroves; signal processing; spectroscopic techniques; spectral discriminant analysis; Thailand, Chumporn, Sawi bay

2.1 Introduction

Remote sensing technology is a potentially fast and efficient approach to mangrove management, with many wetlands covering vast, mostly inaccessible areas, and where ground measurements become difficult and expensive (Held et al., 2003). This statement is strongly supported by a large number of successful applications of remote sensing for mangrove studies, particularly in mangrove resource inventory and change detection (see the review of Green et al., 2000). To highlight just a few, these applications include (i) mapping and monitoring mangroves by using multispectral and hyperspectral sensors (Rasolofoharinoro et al., 1998; Gao, 1999; Demuro and Chisholm, 2003; Held et al., 2003; Hirano et al., 2003), (ii) deriving principal foliage parameters such as Leaf Area Index (LAI) from the correlations between mangroves and spectral bands (Ramsey III and Jensen, 1996; Kovacs et al., 2005), and (iii) detecting change of mangrove ecosystems (Berlanga-Robles and Ruiz-Luna, 2002; Manson et al., 2003).

Multispectral sensors boarded on satellite platforms, including synthetic aperture radar (SAR), Landsat TM, and SPOT XS, are most popularly used for the abovementioned mangrove applications because of their cost-effective advantages (Aschbacher et al., 1995; Green et al., 2000; Held et al., 2003). Nevertheless, owing to their relatively coarse spatial and spectral resolutions, such satellite sensors are limitedly used at the regional scale. New types of sensors that provide higher spatial and spectral details are therefore needed for the study of mangroves at a finer level (Gao, 1999; Green et al., 2000; Sulong et al., 2002).

With respect to the lack of spectral details of multispectral sensors, the limited number of spectral bands of Landsat TM (7 bands in total), in which each band only covers a broad wavelength region of several tens of nanometres, offers a clear example of how opportunities to exploit spectral responses linked to the physico-chemical properties of plants are lost. The broad spectral information of Landsat TM cannot be used to resolve several key absorption pits as well as reflectance characteristics including the red edge (the unique feature of plant spectral responses between the wavelength of 690 nm and 720 nm that can be used for extracting important physico-chemical characteristics of plants including chlorophyll contents) (Elvidge, 1987; Himmelsbach et al., 1988; Curran, 1989; Elvidge, 1990; Kumar et al., 2001; Williams and Norris, 2001). In contrast, the report of Demuro and Chisholm (2003) demonstrates an

example of how more delicate tools such as the satellite-mounted HYPERION sensor (USGS EROS Data Centre (EDC), USA) that possesses 220 bands between 400 nm and 2500 nm handles the task of discriminating 8-class mangrove communities (i.e., broad mangrove classes) in Australia - a task considered difficult for any multispectral sensor (Green et al., 2000). Similarly, the 224-band Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) sensor with approximate 9.6 nm band width ranging between 400 nm and 2450 nm performs just as well in mapping the mangrove communities of the Everglades, Florida (Hirano et al., 2003). Since both HYPERION and AVIRIS sensors collect a contiguous range of narrow-band spectral data, they are technically termed "hyperspectral sensors".

Scientists in the field of vegetation research at large acknowledge that the narrow spectral bands of hyperspectral sensors are advantageous for their studies. Such information gathered from the relationships between plants and their spectral responses can be used for gauging biochemical contents inside the plants (Gates et al., 1965; Hoffer, 1978; Peterson and Hubbard, 1992; Kokaly, 2001, McDonald, 2003). The exploitation of these relationships are found in a large number of vegetation research (Green et al., 1998; Asner et al., 2000; Curran et al., 2001; Soukupová et al., 2002; Goel et al., 2003; Hirano et al., 2003; Mutanga et al., 2003; Schuerger et al., 2003; Zarco-Tejada et al., 2004), particularly in the area of plant species discrimination studies (Cochrane, 2000; Schmidt and Skidmore, 2003; Ramsey III et al, 2005; Clark et al., 2005). Nonetheless, the capability of hyperspectral technology for discriminating mangroves at the species level is still unconfirmed (Green et al., 2000; Demuro and Chisholm, 2003; Held et al., 2003; Hirano et al., 2003).

Consequently, this study is intended to move one step closer to the conclusion whether hyperspectral technology can be used for tropical mangrove species discrimination. Specifically, laboratory spectra of top canopy leaves of 16 tropical mangrove species are used for the spectral separability analysis to see whether they adequately contain useful spectral information for discriminating mangroves at the species level. Using laboratory data that leave out the difficulties of field conditions (e.g., the fluctuation of light source energy, the change of daily atmospheric states, the effect of canopy formations, the cost of accessibility, the coarser spatial and spectral resolutions of on-board hyperspectral sensors, the effect of seasonal changes, the effect of

background soils and water, the difference between the energy of artificial lamps used in the laboratory and the sun) means that the result of this study cannot be used to make any conclusion whether real-life hyperspectral sensors (e.g., HYPERION, AVIRIS, etc.) can be used for discriminating tropical mangrove species. Instead, this laboratory study is intended to be a cost-effective test to focus only on one of the most important prerequisites for the future application of on-board hyperspectral sensors: if the laboratory spectra of the mangrove species contain insufficient spectral information for discriminating mangroves at the species level, it is then not worthwhile to invest a lot of time and money to investigate further into the potential of the on-board hyperspectral sensors.

2.2 Methods

2.2.1 Acquisition of hyperspectral data

2.2.1.1 Mangrove leaf preparation

In the morning of February, 6, 2001, top tree canopies of 16 tropical mangrove species (Table 2.1) were collected using a line-transect method in the natural mangrove forest of Ao Sawi (Sawi Bay), the province of Chumporn, the south of Thailand (10˚ 15'N, 99˚ 7'E) (Figure 2.1). Species identification was carried out by the staffs of Royal Thai Forestry Department; taxonomy follows Tomlinson (1994) and Teeratanatorn (2000). There were ten transects randomly placed throughout the area so as to collect tree samples from every mangrove zone (e.g., pioneer, intermediate, upper zones). Only the trees that are higher than 2.5 m were considered for the sampling campaign. The canopies of the sampled trees were cut off and transported to the laboratory where leaves were picked off for the spectral measurement. The whole process was done within four hours so as to preserve the quality of the leaves.

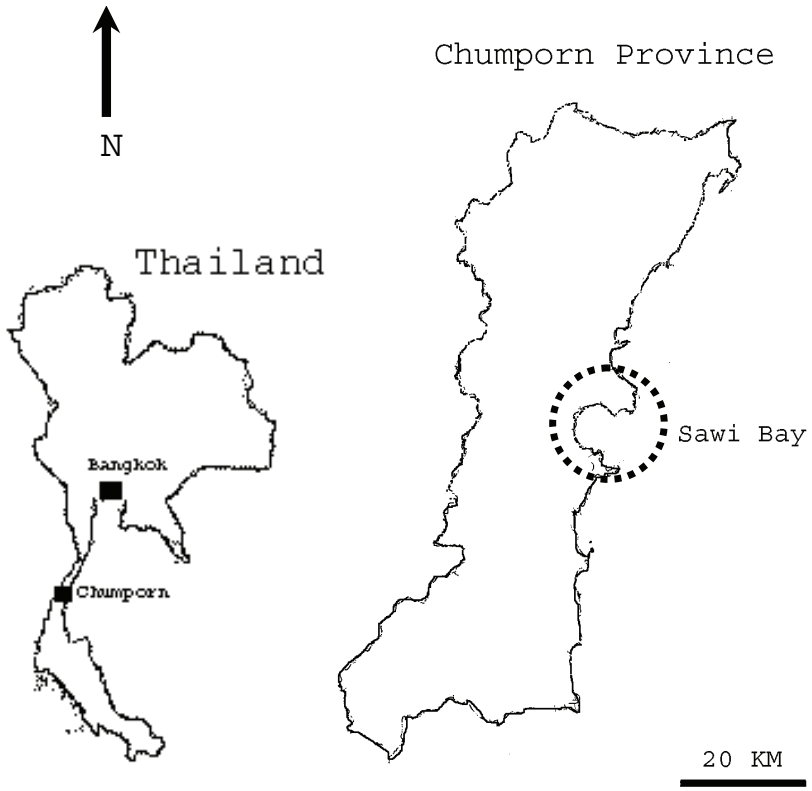


***Figure 2.1**: Sawi Bay, Chumporn, Thailand (10˚ 15'N, 99˚ 7'E)*

Table 2.1: *Sixteen tropical mangrove species of different mangrove zones collected from Sawi Bay, Chumporn, Thailand used for the laboratory reflectance measurement (See also: Appendix II for spectral representations of these mangroves)*

Mangrove species	Species code
Avicennia alba	AVA
Acrostichum aureum	ACA
Bruguiera cylindrica	BC
Bruguiera gymnorrhiza	BG
Bruguiera parviflora	BP
Ceriops tagal	CT
Excoecaria agallocha	EA
Heritiera littoralis	HL
Lumnitzera littorea	LL
Lumnitzera racemosa	LR
Nypa fruticans	NF
Pluchea indica	PI
Rhizophora apiculata	RA
Rhizophora mucronata	RM
Sonneratia ovata	SO
Xylocarpus granatum	XG

2.2.1.2 Leaf spectral measurements

Freshly-picked leaves were randomly divided into 30 piles of the same size (20 to 30 leaves) per mangrove species. First, each pile of leaves (top side up) was spread on top of a black metal plate painted with ultra-flat black paint until the background metal plate could not be seen. Second, the spectral response of each leaf plate was recorded 20 times. Each plate was rotated 45˚ horizontally after every fifth record in order to correct for the bi-directional reflectance distribution function (BRDF). Third, the 20 records were averaged to construct a radiance curve. Fourth, the radiance was converted to a reflectance curve, using a Spectralon reference panel as well as the correction of the spectrometer internal current (dark current). The steps above were repeated for all the leaf plates. As a result, 30 reflectance curves were constructed for each mangrove species (Table 2.1). Please note that the whole operation was conducted under laboratory conditions (i.e., dark room, 25˚C) in order to avoid ambient light sources unrelated to the true spectral signal of the leaves.

The whole process was conducted using a spectroradiometer (FieldSpec Pro FR, Analytical Spectral Device, Inc.). This spectroradiometer was equipped with three spectrometers (i.e., VNIR, SWIR1, and SWIR2), covering 350 nm to 2500 nm, with sampling intervals of 1.4 nm between 350 nm and 1000 nm, and 2 nm between 1000 nm and 2500 nm. The spectral resolution of the spectrometers was 3 nm for the wavelength interval 350 nm to 1000 nm, and 10 nm for the wavelength interval 1000 nm to 2500 nm. The sensor, equipped with a field of view of 25˚, was mounted on a tripod and positioned 0.5 m above the leaf plate at the nadir position.

Since this laboratory study was intended to be a prerequisite for the future use of real hyperspectral sensors, the energy source in use should at least provide the same energy range that real hyperspectral sensors capture. In this study, a halogen lamp was selected to provide stable electro-magnetic energy between 400 nm and 1800 nm. This energy range was reconciled with most of the hyperspectral sensors (Lillesand and Kiefer, 2000). As a result, a halogen lamp fixed on the tripod at the same position as the sensor of the spectrometer was used to illuminate the sample plate.

2.2.2 Data treatments

2.2.2.1 Statistical test

A statistical test was used to compare between the spectral responses of the 16 individual tropical mangrove species (Table 2.1) whether at least one pair of them were statistically different at every spectral band, that is to say, the null hypothesis H_o: $\mu_1=\mu_2=\ldots=\mu_{16}$ versus the alternative hypothesis H_a: $\mu_1\neq\mu_2\neq\ldots\neq\mu_{16}$, where μ_i was the mean reflectance value of the i^{th} species (i = 1, 2,…, 16). Before conducting the test, the distribution of the spectral responses at every spectral band was assumed to be normal under the central limit theorem (N spectra≥30) as well as the equality of statistical variances (homoscedasticity) was verified for every spectral location. Then, the hypothesis test was carried out using one-way ANOVA at every spectral location between 350 nm and 2500 nm (a total of 2151 spectral bands) with 95% and 99% confidence limits (α=0.05 and 0.01).

The aim of the ANOVA test was mainly to visualise the spectral differences between the 16 mangrove species. The test was chosen as a

replacement for the direct graphical presentation of the mangrove spectral responses because the direct visualization was not an effective visualization tool for comparing as many as 16 mangrove species. In other words, the spectral variations within an individual species (i.e., intra-species spectral variations) caused spectral overlaps that made it very difficult to spot the spectral differences between the 16 mangrove species (i.e., inter-species variations) with the naked eye. The reader is recommended to find more details on the direct visualization versus spectral variations in Landgrebe (1997, p.8). Unlike the direct display, applying the ANOVA test helped highlight poor spectral locations at which p-values were greater than α (e.g., α=0.05 or α=0.01). P-values higher than α at some spectral locations indicated that the spectra of different mangrove species were very similar, as none of them was statistically separable from the group. On the other hand, the p-value less than the α threshold indicated that there was at least one pair of mangrove spectra that was statistically different. The ANOVA test was therefore a rapid way of visualising spectral differences. It helped demonstrate that separating spectral responses of 16 different mangrove species was likely at certain spectral positions.

2.2.2.2 Spectral separability

Even though the ANOVA test was a practical data exploration tool, the result of the test may not be independently interpreted without additional treatments. One of the major reasons was the increasing chance of the TYPE I error that usually happened when conducting multiple hypothesis testing (Hsu, 1996; Rothman, 1990; Perneger, 1998; Feise, 2002). In this case, the TYPE I error could lead the reader to feel too positive about the capability of hyperspectral technology for separating mangroves at the species level. As a result, the spectral separability index of every mangrove pair needed to be calculated to guarantee the actual differences between the mangrove spectra. The quantification of spectral separability indices for every mangrove pair was not only minimise the chance of the TYPE I error found in the ANOVA test, but it also was the main contribution of this study (i.e., proving whether the laboratory spectra adequately contain useful spectral information for discriminating mangroves at the species level).

The separability index used in this study was the square of Jeffries-Matusita (J-M) distance analysis. The J-M distance method delivers a value between 0 and $\sqrt{2}$ ($\approx$ 1.414), so the squared distance gives a

number between 0 and 2. We follow the common practice in remote sensing (Thomas et al., 2003; ENVI software's user guide, RSI Inc.) of using a squared J-M distance threshold of ≥ 1.90 to indicate whether any two mangrove species were spectrally separable. The calculation of the J-M distance in this study was based on the following equation (Eq.1). The reader was recommended to consult (Richards, 1993) for further details in separability analyses.

$$JM_{ij} = \sqrt{2(1 - e^{-a})}$$

where $a = \frac{1}{8}(\mu_i - \mu_j)^T \left(\frac{C_i + C_j}{2} \right)^{-1} (\mu_i - \mu_j) + \frac{1}{2} \ln \left(\frac{\frac{1}{2}|C_i + C_j|}{\sqrt{|C_i| \times |C_j|}} \right)$ [Eq.1]

Note: i and j are the spectral responses of two mangrove species being compared. C is the covariance matrix of the spectral response. μ is the mean vector of the spectral response. ln is the natural logarithm function. T is the transposition function. $|C|$ is the determinant of C.

Because the J-M distance measure in use was a parametric method, it was necessary to reduce the number of spectral features (bands) prior to the calculation. In other words, it was not possible to calculate the J-M distance by using all 2151 bands because of the singularity problem of matrix inversion (i.e., the number of spectral samples per mangrove species was too small). A wrapper feature selection approach (Siedlecki and Sklansky, 1989; John et al., 1994; Kohavi and John, 1997; Kavzoglu and Mather, 2002; Yu et al., 2002; Vaiphasa, 2003) was therefore applied in this study to reduce the number of spectral features. The wrapper approach is generally a kind of feature selection algorithms that combines the strength of a traditional search algorithm (e.g., sequential forward selection, branch and bound technique, genetic search) with the capability of a classifier (e.g., nearest neighbour classifier, maximum likelihood classifier). In this case study, the search mechanism of the wrapper tool was based on a genetic algorithm, and its classifier was a nearest neighbour classifier. The algorithm was applied to select the best band combination out of the total of 2151 bands. The algorithm was initialised with the following genetic search parameters: crossover rate = 50%; mutation rate = 1%; and the maximum number of iterations = 1000. The estimated classification accuracy was chosen at an 80% level as an

optimising criterion. Following the USGS guideline (Anderson et al., 1976), the optimising criterion chosen at the 80% level was adequate for the difficulties in discriminating mangroves at the species level (i.e., Level III or IV of the USGS classification standard).

2.3 Results

2.3.1 ANOVA test

The results of 2151 ANOVA tests (p-values) for all spectral bands were plotted in Figure 2.2. A reflectance of *Rhizophora apiculata* measured in the laboratory was also drawn in the figure to give an impression of the actual mangrove spectral continuum collected by the spectrometer. The 16 mangrove species under study were statistically different at most spectral locations, with a 95% confidence level (p-value < 0.05). The total number of spectral bands that had p-values less than 0.05 was 1941, of which 477 bands had a 99% confidence level (p-value < 0.01). The exceptions were at the ultraviolet region (350-400 nm) and shortwave infrared region (1800-2500 nm) where the halogen lamp did not radiate strong energy.

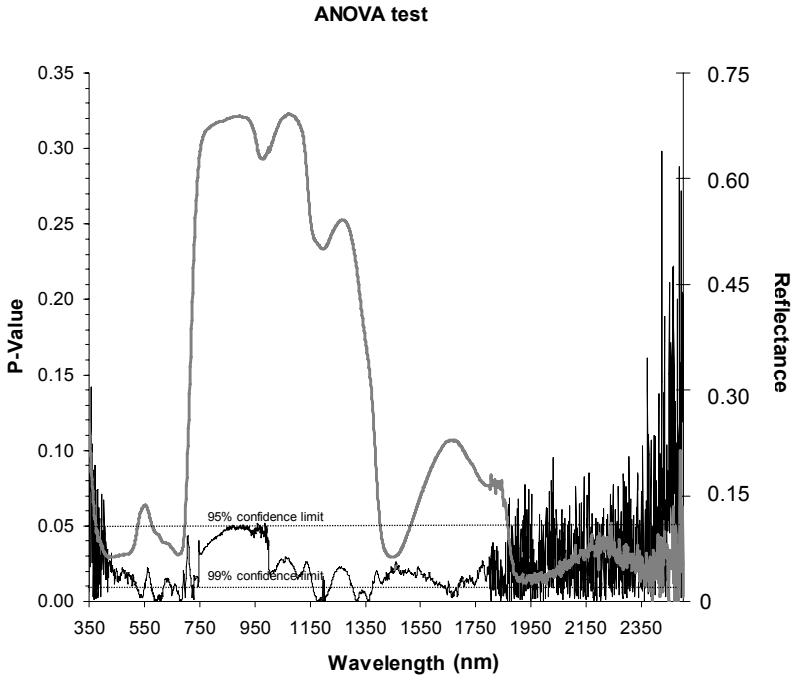


***Figure 2.2**: The plot of p-values of the ANOVA test (black line) shown against a laboratory reflectance of Rhizophora apiculata (grey line)*

2.3.2 Wrapper feature selection

The feature selection algorithm was applied to search for the sub-optimal spectral band combination out of the total of 2151 bands. The best combination found by the wrapper tool comprised four spectral members at 720 nm, 1277 nm, 1415 nm, and 1644 nm. These four spectral bands guaranteed an 80% level of estimated classification accuracy. In Figure 2.3, the selected bands were shown against a reflectance of *R. apiculata*. These four selected bands were then used for the calculation of J-M distances in the next section.

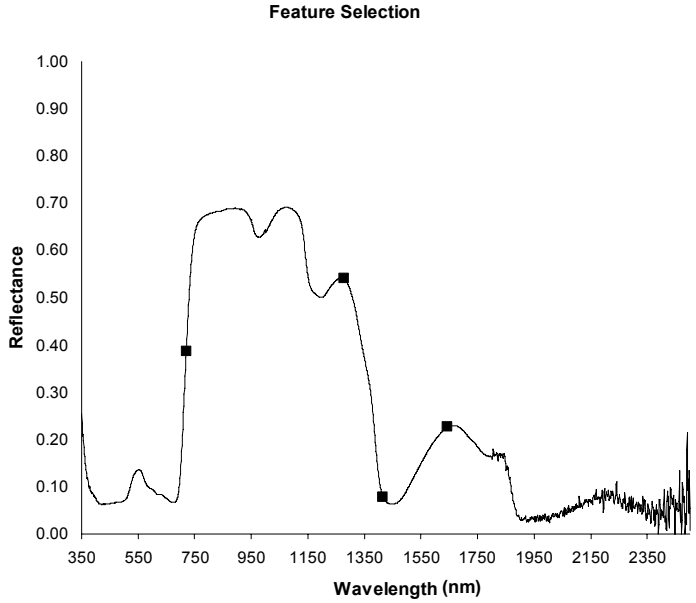


***Figure 2.3**: Four locations of spectral bands selected by the feature selection tool at 720 nm, 1277 nm, 1415 nm, and 1644 nm, respectively.*

2.3.3 J-M distance

The J-M distance measure was applied to reveal the spectral separability between each pair of mangrove species (Table 2.2), using the four spectral bands selected in section 3.2. Please note that the mangrove species of Table 2.2 were grouped by their family name. The overall spectral separability between the pairs of mangrove species was high, since most of them acquired levels of separability higher than the selected threshold (namely, 1.90). Only ten instances where the J-M distances were lower than 1.90 were found. These instances were

highlighted in Table 2.2. One should, however, note that the members of the Rhizophoraceae family (*B. cylindrica*, *B. Gymnorrhiza*, *B. parviflora*, *C. tagal*, *R. apiculata*, and *R. mucronata*) were spectrally similar. Five out of ten highlighted instances were found among them. Moreover, the Rhizophoraceae family was also similar to other mangrove families, as each of the other 5 highlighted pairs contained at least one mangrove species of Rhizophoraceae.

Table 2.2: *The J–M distances between all pairs of 16 mangrove species (120 pairs in total). The species names are coded in Table 1. The pairs that possess lower than 1.90 separability levels are highlighted in grey colour. Mangrove species were grouped by their family names*

		Avicennia-ceae	Pterida-ceae	Rhizophoraceae					Euphorbia-ceae	Sterculia-ceae	Combreta-ceae		Wurmb	Astera-ceae	Sonneratia-ceae	Melia-ceae
		AVA	ACA	BC	BG	BP	CT	RA	RM	EA	HL	LL	LR	NF	PI	SO	XG
Avicennia-ceae	AVA																
Pteridaceae	ACA	1.99															
Rhizophora-ceae	BC	1.99	2.00														
	BG	2.00	1.56	1.99													
	BP	1.99	1.99	1.99	1.82												
	CT	2.00	1.99	1.97	1.99	1.90											
	RA	1.99	1.99	1.93	1.99	1.99	1.99										
	RM	2.00	1.99	1.72	1.99	1.73	1.85	1.86									
Euphorbia-ceae	EA	1.94	1.99	2.00	1.99	1.99	2.00	1.99	2.00								
Sterculia-ceae	HL	1.99	1.99	2.00	1.94	1.99	1.99	1.99	1.99	1.98							
Combreta-ceae	LL	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00						
	LR	2.00	2.00	2.00	2.00	2.00	2.00	1.99	2.00	2.00	2.00	1.99					
Wurmb	NF	1.99	1.99	2.00	1.99	1.99	2.00	1.99	1.99	1.99	1.99	2.00	2.00				
Asteraceae	PI	2.00	2.00	1.98	1.99	1.99	1.87	1.99	1.89	2.00	2.00	1.99	2.00	2.00			
Sonneratia-ceae	SO	1.99	1.99	1.99	1.95	1.84	1.99	1.98	1.99	1.99	1.99	1.99	2.00	1.99	1.99		
Melia-ceae	XG	1.97	1.99	1.99	1.82	1.96	2.00	1.99	1.99	1.98	1.99	2.00	2.00	1.99	2.00	1.99	

2.4 Discussion & conclusion

A laboratory-scale test of spectral separability between various tropical mangrove species, which is one of the most important prerequisites for the future use of airborne and satellite hyperspectral sensors, has been completed in this study. Overall, the results confirm that discriminating spectral responses of different tropical mangroves at the species level is possible in the laboratory. First, the result of the ANOVA test in Figure 2.2 helps visualise the possibility of separating the mangrove species at many spectral locations. Then, the report on pair-wise spectral separability between the 16 mangrove species in Table 2.2 guarantees the result of the ANOVA test, as most mangrove pairs possess high separability indices (≈2.00). The results therefore encourage further investigation into the capability of using airborne and satellite hyperspectral sensors for mapping mangrove species when taking field conditions into account (e.g., the fluctuation of solar energy, the change of daily atmospheric states, the effect of canopy formations, the cost of accessibility, the coarser spatial and spectral resolutions of on-board hyperspectral sensors, the effect of seasonal changes, the effect of background soils and water, the difference between the energy of artificial lamps used in the laboratory and the sun).

Despite the optimism of the overall outcome, one should not ignore the minority (10 instances) of Table 2.2 where the separability indices are lower than 1.90. The locations of these instances on the table imply that the members of the Rhizophoraceae family are probably the most problematic: the members of this mangrove family are spectrally similar to the other mangroves as well as among themselves. Such results reflect the similarity between their spectral responses; hence the closeness between the leaf physico-chemical properties of these mangroves. Since the mangroves of Rhizophoraceae dominate the study area, the difficulties in discriminating these mangroves are expected when implementing the on-board hyperspectral sensors. This statement could also be true for other areas that shared similar floristic conditions (i.e., dominated by the Rhizophoraceae family).

Lastly, the result of the wrapper tool may not be overlooked. Even if the four bands selected by the wrapper tool guarantee an 80% level of estimate classification accuracy (i.e., complied with the USGS standard (Anderson et al., 1976)), only one out of four is reconciled with the locations of the spectral responses of mangrove leaf pigments (e.g.,

chlorophylls, carotenoids) between 380 nm and 750 nm (Menon and Neelakantan, 1992; Basak et al., 1996; Das et al., 2002). This could lead us to hypothesise that the spectral responses of mangrove pigments may contain less important spectral information for mangrove species discrimination than the information from the spectral responses of the other leaf components that interacts with electro-magnetic energy at longer wavelengths. This may be because mangroves generally possessed similar amounts of pigment substances across the species but the differences in other leaf components (salt, sugar, water, protein, oil, lignin, starch, cellulose, and leaf structure) that normally interact with energy at longer wavelengths are more marked. Even if a number of studies on the physico-chemical properties of leaves of different mangrove species are available (Menon and Neelakantan, 1992; Tomlinson, 1994; Basak et al., 1996; Das et al., 2002), it is unfortunate that they can not be readily compared in order to draw any conclusion. This is because these studies are not standardised (i.e., the mangrove leaves used in different reports were collected from different field conditions). A non-bias comparative study is therefore recommended so as to confirm this part of the findings. Then, the four spectral locations selected by the wrapper tool could be seen as a guideline for selecting appropriate spectral locations for the future use of the on-board hyperspectral sensor.

In summary, one of the most important prerequisites for the future investment of the airborne and satellite hyperspectral sensors for mangrove studies was investigated in this study. Laboratory spectra of top canopy leaves of 16 tropical mangrove species were analyzed to see whether they adequately contained useful spectral information for discriminating mangroves at the species level. Overall, the results from the statistical test and the spectral distance analysis provide optimistic evidence that encourages a full-scale investigation into the capability of on-board hyperspectral sensors for mangrove species discrimination, but the doubt of discriminating some members of the Rhizophoraceae family still remains.

Chapter 3[ab]

Dimensionality Problems

[a] Vaiphasa, C., van Oosten, H., Skidmore, A.K., de Boer, W.F. (in review). A genetic algorithm for hyperspectral feature selection. Photogrammetric Engineering and Remote Sensing.
[b] Vaiphasa, C. (in review). A hyperspectral band selector for plant species discrimination. Photogrammetric Engineering and Remote Sensing.

Abstract

Using genetic search algorithms (GA) as spectral band selectors is popular in the field of remote sensing. Nevertheless, class information used in the existing research for testing the performance of the GA-based band selector is broad (i.e., Anderson's level I or II). This means that each class possesses distinct spectral characteristics, and it is relatively easy for the band selector to find spectral bands that maintain high spectral separability between classes. None of the existing studies has tested the band selector on class information that possesses very similar spectral characteristics (e.g., species-level data). The question therefore remains if the band selector can deal with such complexity. As a result, the key hypothesis of this research is that the GA-based band selector can be used for selecting a meaningful subset of spectral bands that maintains spectral separability between species classes. The test data in use are spectrometer records of high dimensionality, comprising 2151 bands of leaf spectra of 16 tropical mangrove species. Overall, it turned out that the GA-based band selector was able to cope with spectral similarity at the species level. It selected spectral bands that related to principal physico-chemical properties of plants, and, simultaneously, maintained the separability between species classes at a high level.

Keywords: Artificial Intelligence; Classification; Hyper spectral; Remote sensing; Vegetation

3.1 Introduction

Since the first introduction by Holland (1975), many forms of genetic algorithms (GA) have been developed for remote sensing applications, for example, (i) image segmentation and classification (Tseng and Lai, 1999; Pal et al., 2001; Harvey et al., 2002; Liu et al., 2004; Bandyopadhyay, 2005), (ii) sub-pixel classification (Mertens et al., 2003), (iii) model optimization (Jin and Wang, 2001; Chen, 2003; Fang et al., 2003), (iv) image registration (Jones et al., 2000; Chalermwat et al., 2001); (v) pixel aggregation (Lu and Eriksson, 2000), and (vi) image band selection (Siedlecki and Sklansky, 1989; Lofy and Sklansky, 2001; Kavzoglu and Mather, 2002; Yu et al., 2002; Fang et al., 2003; Kooistra et al., 2003; Luo et al., 2003; Ulfarsson et al., 2003; Cogdill et al., 2004). Ranking by the number of publications, the use of GA as band selectors is the most popular.

In general, band selectors alleviate the problem of high-dimensional complexity (Bellman, 1961; Kendall, 1961; Hughes, 1968; Fukunaga, 1990; Shahshahani and Landgrebe, 1994) that usually affects the outcome of analysing multiple band data (e.g., multi-sensors, multi-temporal, or hyperspectral images). In most cases, a large number of image bands (i.e., > 20 bands) are too complex for familiar parametric tools (e.g., Jeffries-Matusita distance, Bhattacharyya distance, maximum likelihood classifier). Mathematically, the complexity when using such a large number of bands does not only undermine the precision of parametric model estimation (Bellman, 1961; Hughes, 1968), but it also causes the singularity of covariance matrix inversion (Fukunaga, 1990). Furthermore, this high-dimensional complexity also leads to an excessive demand for field samples, which is not feasible in practice owing to the time and budget limitations (Shahshahani and Landgrebe, 1994).

The superiority of the GA-based band selectors over the other popular band selection algorithms (e.g., branch and bound search, exhaustive search, sequential forward selection) is acknowledged (Siedlecki and Sklansky, 1989). For example, the successful applications include (i) selecting a subset of multiple sensor/date data for image classification (Lofy and Sklansky, 2001; Kavzoglu and Mather, 2002; Ulfarsson et al., 2003), (ii) selecting spectral bands that relate to physico-chemical characteristics of plants and soils (Fang et al., 2003; Kooistra et al., 2003, Cogdill et al., 2004, and (iii) selecting a spectral subset of hyperspectral data for image classification (Yu et al., 2002).

Nonetheless, the existing GA-based band selectors (Lofy and Sklansky, 2001; Kavzoglu and Mather, 2002; Yu et al., 2002; Ulfarsson et al., 2003) have only been tested on broad class information at the USGS level I or II (Anderson et al., 1976). In other words, it is relatively easy for GA to select spectral bands from these broad class data and maintain high spectral separability between classes. Therefore, it is still unclear whether GA can deal with higher spectral complexity. In this study, we then hypothesise that the GA-based band selector can be used for selecting a meaningful spectral subset from complex test data that maintains spectral separability. The test data in use are spectrometer records of high dimensional complexity that comprise 2151 bands of leaf spectra of 16 tropical mangrove species.

3.2 Data & Methods

3.2.1 Species-level hyperspectral data

3.2.1.1 Mangrove leaf preparation

Top-level canopies of 16 tropical mangrove species (Table 3.1) were collected using a line-transect method in the natural mangrove forest of Ao Sawi (Sawi Bay), the province of Chumporn, the south of Thailand (10˚ 15'N, 99˚ 7'E) on February, 6, 2001. This line-transect method enabled us to collect the mangrove canopies from pioneer, intermediate, and landward zones. The canopies were only sampled from fully-grown trees (i.e., > 2.5 m tall). During the sampling campaign, species identification was carried out by the staffs of Royal Thai Forestry Department; taxonomy follows Tomlinson (1994) and Teeratanatorn (2000). At the laboratory, the leaves were then picked off the canopies for the spectral measurement.

Table 3.1: *Thirty spectra of mangrove leaves were collected per mangrove species, using a 2151-band spectroradiometer*

Mangrove species	Species code	Number of spectra
Avicennia alba	1	30
Acrostichum aureum	2	30
Bruguiera cylindrica	3	30
Bruguiera gymnorrhiza	4	30
Bruguiera parviflora	5	30
Ceriops tagal	6	30
Excoecaria agallocha	7	30
Heritiera littoralis	8	30
Lumnitzera littorea	9	30
Lumnitzera racemosa	10	30
Nypa fruticans	11	30
Pluchea indica	12	30
Rhizophora apiculata	13	30
Rhizophora mucronata	14	30
Sonneratia ovata	15	30
Xylocarpus granatum	16	30

3.2.1.2 Leaf spectral measurements

The leaves were randomly shuffled and separated evenly into 30 piles per mangrove species. Each pile of leaves (top side up) was placed on top of a black metal plate painted with ultra-flat black paint until the background metal plate could not be seen. Next, the spectral response of each leaf plate was recorded 20 times. Each plate was rotated 45˚ horizontally after every fifth record to compensate for the bi-directional reflectance distribution function (BRDF). Then, the mean of the 20 records were calculated to construct a radiance curve. Finally, the radiance was converted to a reflectance curve by using a reference panel as well as the correction of the spectrometer internal current (dark current). The steps above were followed for all other leaf plates. As a result, we have 30 reflectance curves per each mangrove species (Table 3.1).

The spectral measurement was conducted under laboratory conditions by using a spectroradiometer (FieldSpec Pro FR, Analytical Spectral Device, Inc.). This spectroradiometer was equipped with three

spectrometers (i.e., VNIR, SWIR1, and SWIR2), covering 350 nm to 2500 nm, with sampling intervals of 1.4 nm between 350 nm and 1000 nm, and 2 nm between 1000 nm and 2500 nm. The spectral resolution of the spectrometers was 3 nm for the wavelength interval 350 nm to 1000 nm, and 10 nm for the wavelength interval 1000 nm to 2500 nm. The sensor, equipped with a field of view of 25˚, was mounted on a tripod and positioned 0.5 m above the leaf plate at the nadir position. A halogen lamp fixed on the tripod at the same position as the sensor of the spectrometer was used to illuminate the sample plate. The room was conditioned to be dark with 25˚C in order to avoid unwanted external energy sources.

3.2.2 Genetic search algorithms (GA)

The theory of GA was first introduced by Holland (1975). The elaboration of its practical side including a basic computer source code can be found in Goldberg (1989). In this chapter, guidelines of Goldberg (1989) are strictly followed. More details of the algorithm are added in Appendix I. Only three major connections between the concept of GA and remote sensing applications are emphasised in this chapter (i.e., gene encoding scheme, reproduction mechanism, and fitness criterion). Additionally, the code of GA used in this study has been developed in the IDL language at the International Institute for Geo-Information Science and Earth Observation (ITC) (Vaiphasa, 2003).

3.2.2.1 Gene encoding

The gene encoding scheme in use is a direct method instead of binary encoding that is popularly used in related studies (Siedlecki and Sklansky, 1989; Kavzoglu and Mather, 2002; Yu et al., 2002). The key reasons for choosing direct encoding are that it is transparent for tracking the process of evolution as well as straightforward for reproducing the population (i.e., crossover and mutation) (Vaiphasa, 2003).

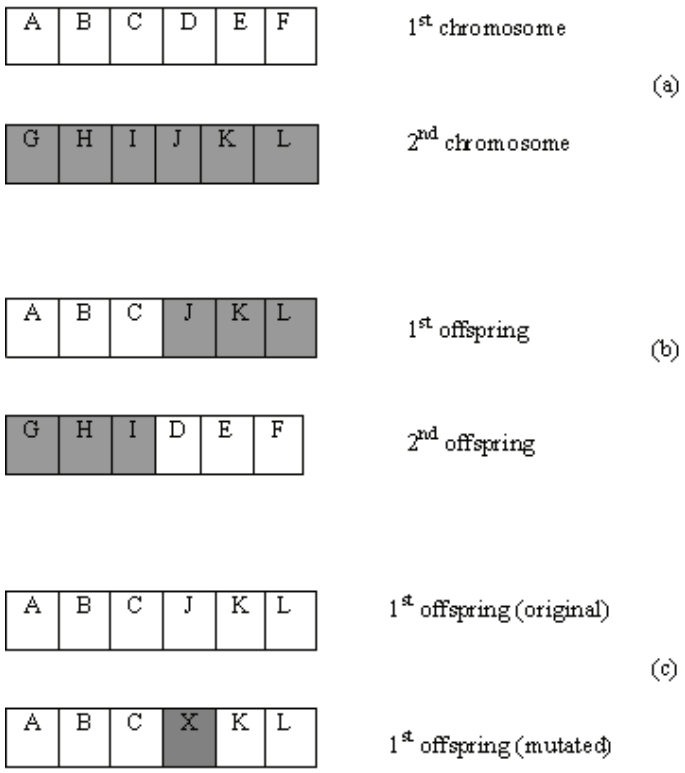


Figure 3.1*: (a) Two parent chromosomes, (b) Two offspring chromosomes, and (c) An example of random mutation*

Figure 3.1a illustrates an example of two chromosomes (e.g., chromosome size = 6). Following the direct encoding scheme, the 1st chromosome comprises 6 different genes flagged by the letter A to F, and the 2nd one comprises gene G to L. Each gene can be assigned with a band label. For example, the 1st chromosome, {A, B, C, D, E, F}, in Figure 3.1a can be assigned with an array of band names, {Band 2, Band 8, Band 37, Band 59, Band 97, Band 99}, and, similarly, the {G, H, I, J, K, L} can be set to {Band 3, Band 5, Band 38, Band 55, Band 83, Band 100}.

3.2.2.2 Reproduction mechanism

The mechanism of cross over and mutation is illustrated in Figure 3.1. By mating the two chromosomes in Figure 3.1a, the offspring that they produce share, in this example, half the characters of the first parent and

the other half from the second parent. The two offspring are shown in Figure 3.1b.

Occasionally, some of the genes in any newly produced chromosome are randomly altered by mutation. This phenomenon causes a change in the character of the offspring, independent from the chromosome composition of the parents. The illustration of the mutation effect is shown in Figure 3.1c. The "J" gene is mutated to the "X" gene through random mutation. In remote sensing context, this is equal to a random flip of a band label inside a chromosome.

3.2.2.3 Fitness criterion

The fitness function in use is a well-known spectral angle mapper based nearest neighbour classifier (SAM) (see Kruse et al. (1993) and Keshava (2004) for full details). This means that the evolution is guided by the classification accuracy reported by SAM. Chromosomes (i.e., a subset of spectral bands) that possess higher classification accuracy are likely to have more chance to "mate" and produce "young" than chromosomes that possess lower classification accuracy (Goldberg, 1989).

3.3 Experiments & Results

3.3.1 Initialising the genetic search algorithm

GA was initialised with the following parameters: population size = 1000, crossover rate = 100%, and mutation rate = 1%. The maximum number of generations was 500. The fitness function (i.e., SAM) was trained with half of the mangrove spectra of Table 3.1 (15 spectra per class), and the other half was used for calculating on-line fitness progress.

3.3.2 Choosing an appropriate chromosome size

Since the genetic algorithm in use was an unconstrained combinatorial optimization search (i.e., search without any constraint or penalty on the size of a chromosome), preliminary runs of GA had to be carried out to look for an appropriate chromosome size (i.e., chromosome size = the number of genes in a chromosome) that maintained high class separability. In this study, the 80% level of classification accuracy was chosen as the target, as it was appropriate for separating very similar spectra of 16 mangrove species (USGS level III or IV (Anderson et al.,

1976)). As a result, it was found that a minimum chromosome size that maintained class separability above the chosen threshold was four. A comparison between the performance of three different chromosome sizes (i.e., chromosome size=2, 3, and 4) is shown in Figure 3.2.

3.3.3 Running the genetic search algorithm

GA with chromosome size four was repeatedly run 30 times to check the consistency of the results. The spectra were randomly rotated at the start of every run (i.e., data rotation) to avoid the bias. The real-time progress was plotted for each run in Figure 3.3. The highest fitness score of each run was marked with a cross. Overall, the genetic algorithm quickly reached an averaged fitness score level of 80% at about the 100^{th} generation.

An example of the evolution process of a single run is shown in Figure 3.4 to give an impression of how GA worked. The horizontal axis represented band labels (or genes) from B1 to B2151. The vertical axis showed the number of genes. In general, the gene distribution pattern converged from originally 2151 types of genes at the 1^{st} generation (Figure 3.4a) to only a few kinds of genes at the 500^{th} generation (Figure 3.4f). The convergence quickly happened as early as the 100^{th} generation (Figure 3.4b) where most genes were already extinct. This convergent evolution from Figure 3.4a to 4b directly connected to Figure 3.3 where the majority of the progress lines levelled off at the 100^{th} generation, as the convergence happened. Genes that dominated the evolution were individually labelled in the plots. In this example, at the last generation (the 500^{th} generation), the gene pool was dominated by the following genes: B369, B915, B1050, B1262, and B1297.

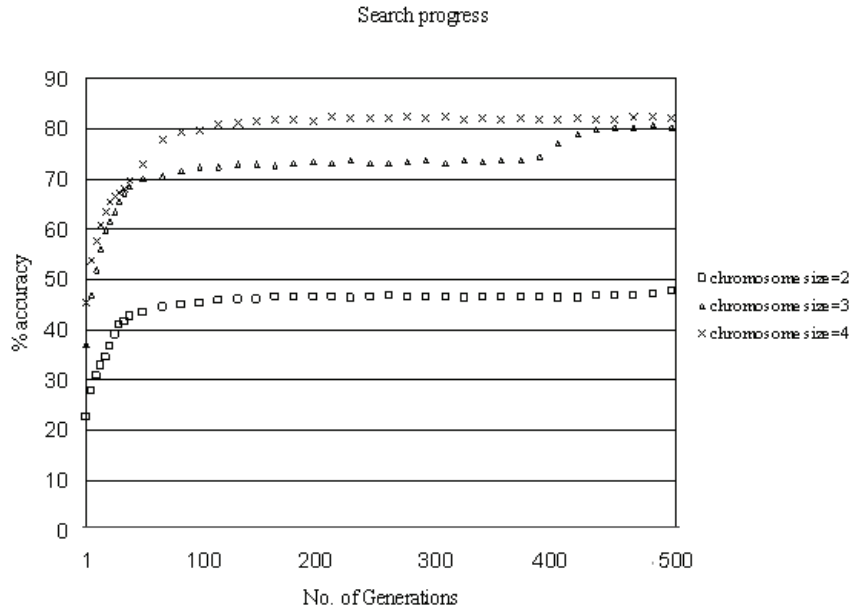


***Figure 3.2**: A comparison between the performances of three different chromosome sizes*

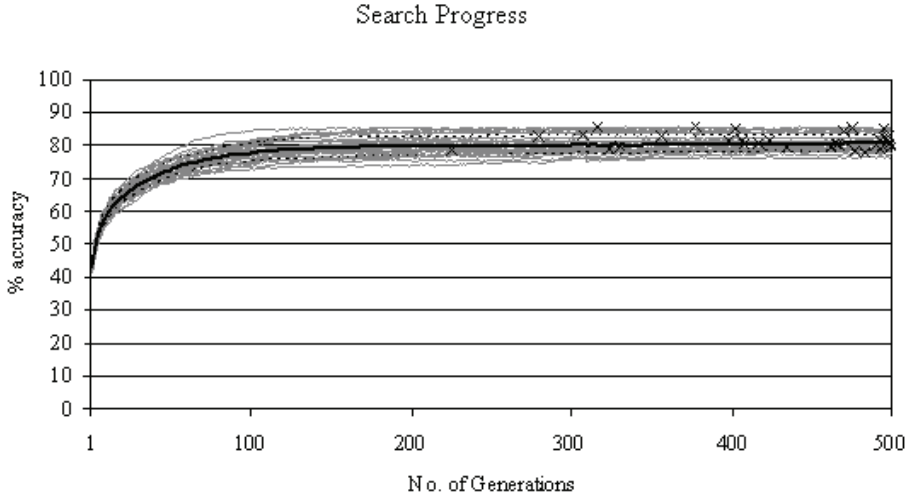


***Figure 3.3**: The real-time progress of 30 runs (grey lines) with their peaks (crosses), mean (black line), and standard deviation limits (dashed lines)*

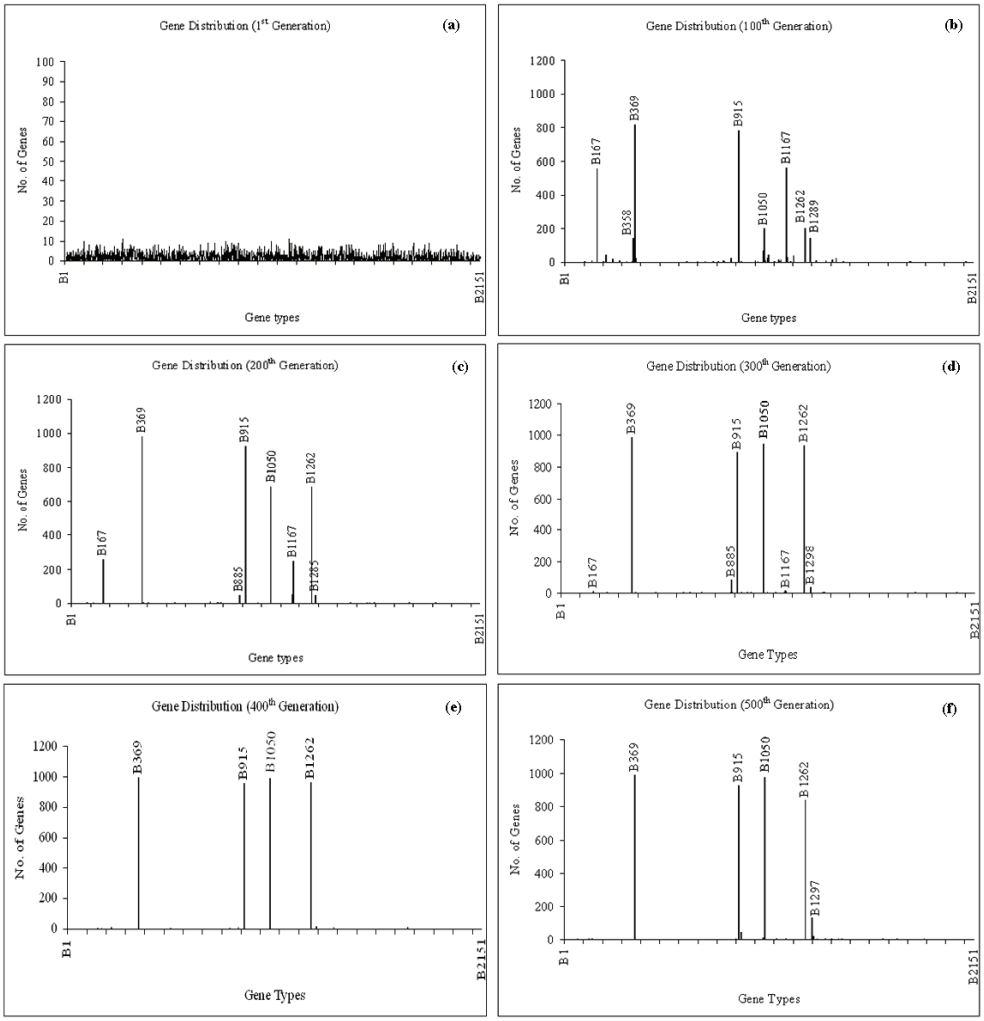


Figure 3.4*: An example of the convergence of gene distribution patterns from (a) the 1st generation to (f) the 500th generation*

The winning chromosomes from every run were reported in Table 3.2 along with their fitness scores (i.e., SAM classification accuracy). The best of all were chromosome No.2 and No.10. Both possessed an 86% level of classification accuracy. Then, all the genes of the 30 winning chromosomes (i.e., 120 genes in total) are grouped by minimising their variance. The results were illustrated in a plot against a hypothetical mangrove reflectance (Figure 3.5). It was found that the genes (spectral bands) can be grouped at 6 different spectral positions (mean ± standard deviation): visible area (21 genes at 513±19nm); red edge (15 genes at

717±16nm); near-infrared region (9 genes at 1263±23nm); infrared slope (44 genes at 1385±27nm); mid-infrared absorption pitch (5 genes at 1489±21nm), and mid-infrared peak (26 genes at 1669±25nm).

Table 3.2: *Thirty winning chromosomes with their encapsulated genes*

Chromosome No.	Genes (nanometre)				Fitness Scores (%)
1	523	1358	1385	1710	83
2	518	1381	1393	1685	86
3	517	708	1436	1639	82
4	523	1372	1418	1671	81
5	524	1375	1496	1681	82
6	521	1378	1398	1665	79
7	549	1333	1390	1681	81
8	679	1316	1389	1673	79
9	535	1384	1506	1667	85
10	716	1246	1409	1607	86
11	722	758	1392	1436	80
12	534	1364	1385	1685	84
13	528	1363	1408	1661	81
14	725	1264	1402	1682	79
15	533	1369	1507	1660	79
16	548	1335	1458	1644	80
17	546	714	1403	1626	82
18	515	1380	1409	1674	78
19	593	1388	1480	1667	80
20	711	1234	1381	1699	83
21	536	717	1230	1397	85
22	526	725	1253	1395	83
23	495	705	1355	1398	81
24	726	1282	1381	1692	83
25	532	1350	1435	1676	79
26	523	1368	1418	1671	85
27	717	1290	1389	1668	80
28	717	1276	1405	1721	82
29	713	1289	1393	1658	78
30	528	1337	1386	1628	81

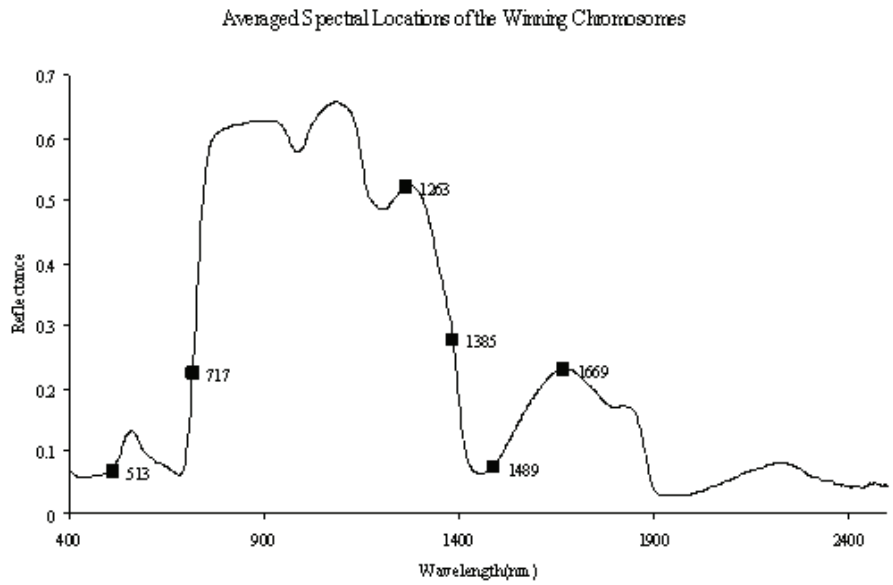


***Figure 3.5**: Six averaged spectral positions of the winning chromosomes*

***Table 3.3:** A statistical comparison between the class separability of band combinations selected by the genetic algorithm and the class separability of band combinations selected by chance (please see Table 3.1 for the class information): (a) genetic search results; (b) random search results; and (c) p-values*

(a)

Class	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1																
2	2.00															
3	2.00	2.00														
4	1.99	1.74	2.00													
5	1.98	2.00	2.00	1.87												
6	2.00	2.00	2.00	2.00	1.84											
7	1.98	2.00	2.00	1.98	2.00	2.00										
8	2.00	1.95	2.00	1.90	2.00	2.00	1.99									
9	2.00	2.00	2.00	2.00	2.00	1.99	2.00	2.00								
10	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00							
11	2.00	1.97	2.00	1.87	1.95	2.00	1.97	1.99	2.00	2.00						
12	2.00	2.00	1.99	2.00	2.00	1.99	2.00	2.00	2.00	2.00	2.00					
13	1.97	2.00	1.94	2.00	1.99	2.00	2.00	2.00	2.00	2.00	2.00	2.00				
14	1.99	2.00	1.96	2.00	1.97	1.96	2.00	2.00	2.00	2.00	2.00	1.98	1.97			
15	1.98	2.00	2.00	1.99	1.82	1.97	2.00	2.00	2.00	2.00	1.98	2.00	1.97	1.98		
16	1.91	2.00	2.00	1.96	1.99	2.00	1.98	2.00	2.00	2.00	2.00	2.00	1.99	2.00	1.98	

(b)

Class	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1																
2	1.99															
3	1.84	2.00														
4	1.89	1.41	1.98													
5	1.73	1.98	1.97	1.77												
6	1.56	1.99	1.91	1.95	1.59											
7	1.88	1.86	1.99	1.55	1.64	1.88										
8	1.99	1.64	2.00	1.73	1.98	1.98	1.88									
9	1.96	1.99	1.98	1.96	1.95	1.92	1.98	1.97								
10	1.95	2.00	1.95	1.99	1.98	1.98	1.98	2.00	1.94							
11	1.81	1.97	1.97	1.68	1.67	1.92	1.67	1.98	1.95	1.96						
12	1.82	1.98	1.95	1.91	1.90	1.82	1.88	1.97	1.92	1.92	1.86					
13	1.75	2.00	1.22	1.98	1.97	1.83	1.99	2.00	1.97	1.94	1.96	1.96				
14	1.51	1.99	1.68	1.97	1.79	1.46	1.93	2.00	1.94	1.91	1.93	1.81	1.55			
15	1.79	2.00	1.98	1.94	1.75	1.77	1.97	2.00	1.93	1.99	1.81	1.86	1.96	1.82		
16	1.58	1.82	1.95	1.36	1.47	1.78	1.38	1.90	1.97	1.97	1.51	1.77	1.94	1.83	1.77	

(c)

Class	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1																
2	0.01															
3	0.00	0.11														
4	0.00	0.00	0.05													
5	0.00	0.03	0.03	0.02												
6	0.00	0.06	0.00	0.02	0.00											
7	0.00	0.00	0.02	0.00	0.00	0.00										
8	0.06	0.00	0.16	0.00	0.12	0.13	0.01									
9	0.00	0.07	0.03	0.05	0.05	0.01	0.10	0.04								
10	0.04	0.10	0.04	0.06	0.08	0.02	0.10	0.15	0.01							
11	0.00	0.50	0.03	0.00	0.00	0.00	0.00	0.06	0.02	0.06						
12	0.00	0.09	0.02	0.04	0.02	0.00	0.01	0.14	0.00	0.01	0.01					
13	0.00	0.07	0.00	0.05	0.03	0.00	0.04	0.08	0.10	0.02	0.02	0.04				
14	0.00	0.11	0.00	0.02	0.00	0.00	0.02	0.13	0.00	0.02	0.00	0.00	0.00			
15	0.00	0.03	0.02	0.01	0.12	0.00	0.01	0.03	0.01	0.12	0.00	0.00	0.22	0.00		
16	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.03	0.03	0.00	0.00	0.01	0.00	0.00	

3.3.4 Testing the key hypothesis

The key hypothesis of this study was tested to see whether the results of band selection carried out by GA were meaningful (i.e., maintaining spectral separability between mangroves). Specifically, the results of GA

were statistically compared against the results of random selection using t-tests. The Jeffries-Matusita (J-M) distance was chosen as an evaluation tool. For each of the 30 winning chromosomes in Table 3.2, its 4 encapsulated spectral bands were used for calculating J-M distances between all mangrove classes. The averaged J-M distances of the 30 winning chromosomes were demonstrated in Table 3.3a. Next, the J-M distances were calculated for 30 sets of randomly generated band combinations, and their averaged J-M distances were reported in Table 3.3b. Subsequently, the t-test results between the two cases were demonstrated in Table 3.3c in terms of p-values. It was clear that the class separability of band combinations selected by the genetic algorithm was significantly higher than the class separability of randomly selected band combinations with a 95% level of confidence (α=0.05), as most of the p-values (94/120 $\approx$ 78%) in Table 3.3c were < 0.05.

3.4 Discussion & Conclusion

In this study, a form of GA-based band selectors was challenged to select spectral subsets of very high-dimensional, species-level data. Unlike the broad-level data (i.e., Anderson's level I or II (Anderson et al., 1976)) used in the existing studies (Lofy and Sklansky, 2001; Kavzoglu and Mather, 2002; Yu et al., 2002; Ulfarsson et al., 2003), spectral profiles of the species-level data were very similar. In other words, spectral profiles of one species were overlapped with the others at most wavelengths (Figure 3.6). The results in Table 3.2 and Figure 3.3 demonstrated that the GA-based band selector overcame such spectral similarity of the species-level data. The band selector was able to select spectral subsets that maintained class separability at an acceptable level (i.e., $\approx$ an 80% level of classification accuracy).

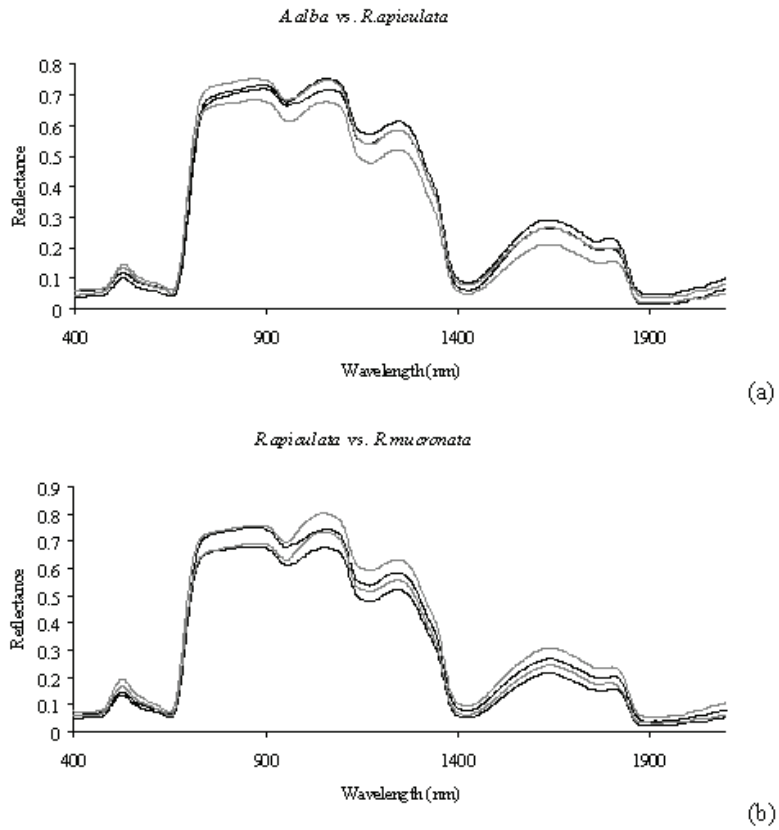


***Figure 3.6**: (a) A min-max plot between A. alba (black line) and R. apiculata (grey line) and (b) A min-max plot between R. apiculata (black line) and R. mucronata (grey line)*

Additionally, the results of hypothesis testing in Table 3.3 also confirmed that band selection carried out by the GA-based band selector was significantly better than by random selection. By majority, spectral separability between 16 mangrove species when the spectral bands were selected by chance was significantly lower than when the spectral bands were selected by the GA-based band selector (i.e., with a 95% level of confidence).

The success of the GA-based band selector can be explained by the chosen spectral locations (Figure 3.5), as each location directly relates to principal physico-chemical properties of plants that helps distinguish between the species. The details of the relationships between these spectral locations and plants can be found in the literature (Elvidge,

1987; Himmelsbach et al., 1988; Curran, 1989; Elvidge, 1990; Menon and Neelakantan, 1992; Tomlinson, 1994; Basak et al., 1996; Kumar et al., 2001; Williams and Norris, 2001; Das et al., 2002). In brief, the band selected from the visible area is needed for discriminating between mangroves that possessed different leaf pigments. The band on the red-edge slope is useful for separating mangrove species that contain different internal leaf structure and water. Similar to the red edge band, the near-infrared band helps sort different plants according to the dissimilarity of their leaf internal structure such as the size of intercellular volume. Finally, the spectral information of the mid-infrared region (i.e., the infrared slope, the mid-infrared absorption pitch, and the mid-infrared peak) is essential for dissolving the internal structure variables and foliar biochemical contents other than the leaf pigments.

The reader may note that the form of GA and its parameters used in this study are not the only options available. To tackle the problem at hand, it is also possible to alter, for example, the encoding scheme, fitness criterion, population size, crossover rate, and mutation rate. Even though the alteration may affect the evolution, it is expected that the robustness of the evolutionary search could still produce a similar outcome (see "freedom of choice" in Goldberg, (1989), page 80). In other words, GA is likely to find meaningful spectral bands that possess high spectral separability despite the alteration. It is, however, beyond the scope of this study to compare different designs of GA and the use of different search parameters.

The optimism gains from the results of this laboratory-level study (i.e., using laboratory spectra) encourages further investigation into the potential of the GA-based band selector for vegetation discrimination when hyperspectral images taken by airborne or satellite sensors are used. Furthermore, it is also anticipated that the use of the GA-based band selector is not limited to the application for vegetation discrimination. The GA-based band selector is now being tested by the author to detect spectral bands that show strong vegetation responses to different physico-chemical treatments (e.g., nitrogen, illumination) in both laboratory and field scenarios. It is hoped that the GA-based band selector could be used as an alternative to traditional methods such as statistical and derivative analyses that are normally used for detecting vegetation responses to external influences (Tsai and Philpot, 1998; Mutanga et al., 2003).

In conclusion, this study strengthens the confidence when using GA as band selection tools. The results confirm that the GA-based band selector is able to cope with spectral similarity at the species level. It selects spectral bands that related to principal physico-chemical properties of plants, and, simultaneously, maintains the separability between species classes at a high level. Additionally, the application of the GA-based band selector other than vegetation discrimination such as the investigation into vegetation spectra in response to different physico-chemical treatments is also anticipated.

Chapter 4*

Spectral Smoothing

* Vaiphasa, C. (in press). Consideration of Smoothing Techniques for Hyperspectral Remote Sensing. ISPRS Journal of Photogrammetry and Remote Sensing.

Abstract

Spectral smoothing filters are popularly used in a large number of modern hyperspectral remote sensing studies for removing noise from the data. However, most of these studies use subjective *ad hoc* measures for selecting filter types and their parameters. We argue that this subjectively-minded approach is not appropriate for choosing smoothing methods for hyperspectral applications. In our case study, it is proved that smoothing filters can cause undesirable changes to statistical characteristics of the spectral data, thereby affecting the results of the analyses that are based on statistical class models. If preserving statistical properties of the original hyperspectral data is desired, smoothing filters should then be used, if necessary, after careful consideration of which smoothing techniques will minimize disturbances to the statistical properties of the original data. A comparative t-test is proposed as choosing a smoothing filter for hyperspectral data at hand.

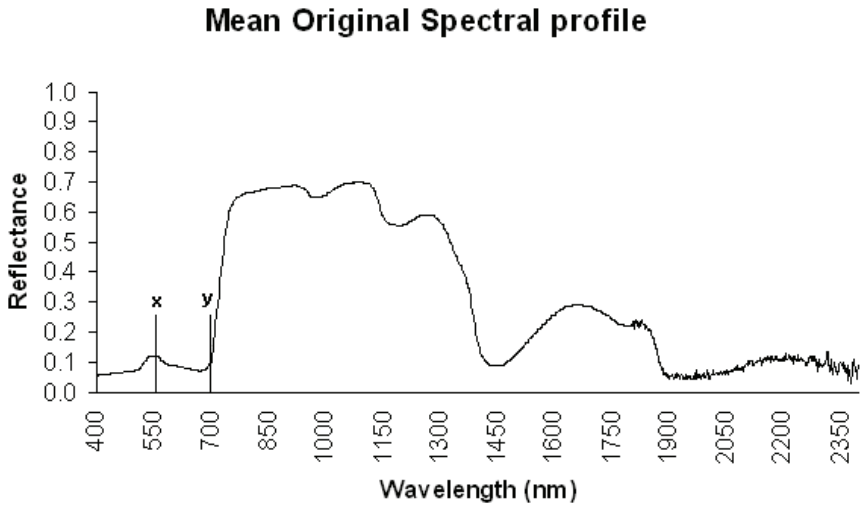
Keywords: aggregation; convolution; hyperspectral; Savitzky-Golay; spectral smoothing

4.1 Introduction

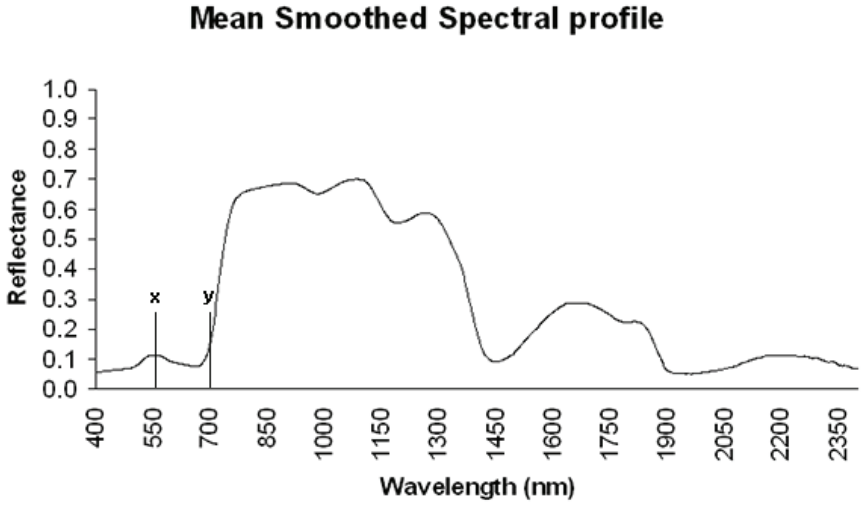
One of the most important problems when using hyperspectral sensors is noise levels. Each narrow-band spectral recorder of the hyperspectral sensor captures small amount of energy that is sometimes undermined by the self-generated noise inside the sensors. This problem is particularly obvious at long wavelength regions (e.g., short-wave infrared). Moreover, physical disturbances such as the fluctuation of light illumination and atmospheric states could further reduce the quality of the spectral signal recorded by the sensors (Oppenheim and Schafer, 1975; Landgrebe, 1997; Lyon, 2004). It is a common practice that *ad hoc* spectral smoothing techniques are used for solving such problem of high noise levels (Chen et al., 2001; Ben-Dor et al., 2002; Andréfouët et al., 2003; Vaughan et al., 2003; Rees et al., 2004; Smith et al., 2004; Thenkabail et al., 2004; Whiting et al., 2004; Zhang et al., 2004; Wu et al., 2005).

Nevertheless, smoothing methods cause changes to the original spectral data that could lead to incorrect results in subsequent analyses (Savitzky and Golay, 1964; Kawata and Minami, 1984; Tsai and Philpot, 1998; Gong et al., 2001; Schmidt and Skidmore, 2004). For example, image processing of remotely sensed imagery for accurate plant biophysical studies (Bruce and Li, 2001; Zarco-Tejada et al., 2001; Imanishi et al., 2004; le Maire et al., 2004; Meroni et al., 2004) and vegetation discrimination and classification (Tsai and Philpot, 2002; Hochberg and Atkinson, 2003; Foody et al., 2004; Schmidt et al., 2004) are dependent upon statistical estimates of spectral data that is often dampened by smoothing filters (Figure 4.1). Instead of *ad hoc* methods, a more objectively-minded approach should be used for selecting the right smoothing method for a particular hyperspectral application so as to minimize the disturbance to the original spectral data.

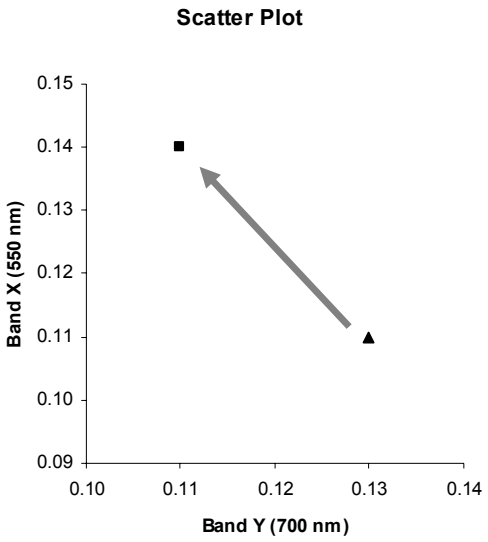
As a result, there are two major goals in this study. First, this study intends to demonstrate the effects of smoothing techniques on the statistical properties of the spectral response. Second, as a replacement for *ad hoc* measures, the study suggests the use of one statistical test (a pair t-test) as a tool for realising the trade-off between the choice of smoothing methods and their effects on statistical properties of the original data. A real example of plant reflectance responses is used for supporting the argument of this study.


Mean Original Spectral profile
Reflectance
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
x
y
400
550
700
850
1000
1150
1300
1450
1600
1750
1900
2050
2200
2350
Wavelength (nm)

(a)


Mean Smoothed Spectral profile
Reflectance
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
x
y
400
550
700
850
1000
1150
1300
1450
1600
1750
1900
2050
2200
2350
Wavelength (nm)

(b)



(c)

Figure 4.1*: (a) An average spectral profile of plants before smoothing; (b) An average spectral profile after smoothing; and (c) A scatter plot of the two principal wavelengths (550 nm and 700 nm) before (triangle) and after (square) smoothing*

4.2 Methods

4.2.1 Smoothing techniques

In the field of digital signal processing, the definition of a spectrum $s_o(\lambda)$ observed by a spectrometer is given by the sum of the true signal of the spectrum $s_t(\lambda)$ and the noise $n(\lambda)$:

$$s_o(\lambda) = s_t(\lambda) + n(\lambda) \qquad \text{[Eq.1]}$$

Please note that λ indicates a wavelength. Thus, the definition of spectral smoothing is the estimation of $s_t(\lambda)$ from the observed spectrum $s_o(\lambda)$. An estimate $\hat{s}_t(\lambda)$ can be calculated by the convolution of the observed spectrum $s_o(\lambda)$ with a weighting function (i.e., smoothing filter) $g(\lambda)$ chosen by the practitioner:

$$\hat{s}_t(\lambda) = s_o(\lambda) * g(\lambda) \qquad \text{[Eq.2]}$$

The operator * denotes convolution integral (Oppenheim and Schafer, 1975; Lyon, 2004). There are many types of smoothing filters $g(\lambda)$

adopted by remote sensing practitioners for hyperspectral applications including linear and non-linear methods (Savitzky and Golay, 1964; Kawata and Minami, 1984; Tsai and Philpot, 1998; Foody et al., 2004; Schmidt and Skidmore, 2004). The most popularly used ones are moving average and Savitzky-Golay filters. As a result, both of them are selected for this study.

4.2.1.1 Moving average

The moving average method for smoothing is well-known and has been called by different names, for example, running mean, mean average filter, etc. The concept of the moving average filter is simple as it takes the mean spectral value of all points within the specified window (i.e., filter size) as the new value of the middle point of the window (Tsai and Philpot, 1998). The method is solely based on linear calculations and has one key parameter, the filter size.

4.2.1.2 Savitzky-Golay

The concept of this method is based on simple polynomial least-square calculations. However, instead of fitting a least-square curve to the total length of spectrum at once, the method fits the spectral data piece-by-piece with the size equal to a user-defined value (i.e., filter size) by using a special form of matrix calculations (Savitzky and Golay, 1964; Steinier et al., 1972; Madden, 1978). This method requires two key parameters: the filter size and the degree of polynomial orders.

4.2.2 Hyperspectral data collection

The data used for testing the key hypothesis of this study are laboratory spectra of mangrove leaves. The data were recorded from the leaves of 16 tropical mangrove species listed in Table 4.1. The leaves were collected using a line-transect method from mangrove trees (higher than 2.5 m) in the natural mangrove forest of Ao Sawi (Sawi Bay), Chumporn, the south of Thailand (10º 15'N, 99º 7'E) on February, 6, 2001. There were ten transects randomly placed throughout the area so as to collect tree samples from pioneer, intermediate, and upper zones. Species identification was carried out by the staffs of Royal Thai Forestry Department; taxonomy follows Tomlinson (1994) and Teeratanatorn (2000).

Table 4.1: *Thirty reflectance curves were recorded per each mangrove species listed below.*

Mangrove species
1. *Avicennia alba*
2. *Acrostichum aureum*
3. *Bruguiera cylindrica*
4. *Bruguiera gymnorrhiza*
5. *Bruguiera parviflora*
6. *Ceriops tagal*
7. *Excoecaria agallocha*
8. *Heritiera littoralis*
9. *Lumnitzera littorea*
10. *Lumnitzera racemosa*
11. *Nypa fruticans*
12. *Pluchea indica*
13. *Rhizophora apiculata*
14. *Rhizophora mucronata*
15. *Sonneratia ovata*
16. *Xylocarpus granatum*

The leaves were randomly broken up into 30 piles per mangrove species. Each pile is about the same size (20-30 leaves). Each pile of leaves (top side up) was then placed on top of a black metal plate until the background could not be seen. We recorded spectral responses of each plate for 20 times. Each plate was rotated 45º horizontally after every fifth record so as to compensate for the bi-directional reflectance distribution function (BRDF). Next, the 20 records were averaged to create a radiance curve. Then, the radiance was converted to a reflectance curve by using a spectralon reference panel, as well as the correction of the spectrometer internal current (dark current). The abovementioned steps were repeated for all the leaf plates. Finally, we have 30 reflectance curves per each mangrove species (Table 4.1).

The whole operation was conducted using the FieldSpec Pro FR spectroradiometer (Analytical Spectral Device, Inc.) under a laboratory condition (i.e., dark room, 25ºC) in order to avoid ambient light sources unrelated to the true spectral signal of the leaves. The instrument was equipped with three spectrometers (i.e., VNIR, SWIR1, and SWIR2), covering 350 nm to 2500 nm (2151 bands in total), with sampling intervals of 1.4 nm between 350 nm and 1000 nm, and 2 nm between

1000 nm and 2500 nm. The spectral resolution of the spectrometers was 3 nm for the wavelength interval 350 nm to 1000 nm, and 10 nm for the wavelength interval 1000 nm to 2500 nm. The sensor, equipped with a field of view of 25°, was mounted on a tripod and positioned 0.5 m above the leaf plate at the nadir position. A halogen lamp fixed at the same position was used to illuminate the sample plate.

4.2.3 Experimental use of smoothing filters

4.2.3.1 Statistical comparisons

We smoothed all the leaf spectra of the mangrove species listed in Table 4.1 (N=480 spectra in total) by using two different types of smoothing filters: (i) the moving average, and (ii) 2nd order Savitzky-Golay. Filter sizes in use were exhaustively varied between 7 and 51 with the increment of 2 (i.e., i = 7, 9, 11,…, 51) for each filter type. Then, we statistically compared the smoothed data against the original spectral data by using pair t-tests for all the filters. The test was thoroughly conducted at every spectral location between 400 nm and 2400 nm. In other words, for each filter, H_o: $\mu_s(\lambda)-\mu_o(\lambda) = 0$ and H_a: $\mu_s(\lambda)-\mu_o(\lambda) \neq 0$ where $\mu_s(\lambda)$ is the mean of smoothed spectra at λ nm wavelength (λ = 400,…,2400 nm) and $\mu_o(\lambda)$ is the mean of original spectra at the same wavelength.

4.2.3.2 Spectral separability analysis

The square of Jeffries-Matusita (J-M) distance analysis was used to quantify the effect of smoothing. The calculation of the J-M distance in this study was based on the following equation (Eq.3) (Richards, 1993).

$$JM_{ij} = \sqrt{2(1-e^{-a})} \qquad \text{[Eq.3]}$$

$$\text{where } a = \frac{1}{8}(\mu_i - \mu_j)^T \left(\frac{C_i + C_j}{2}\right)^{-1} (\mu_i - \mu_j) + \frac{1}{2}\ln\left(\frac{\frac{1}{2}|C_i + C_j|}{\sqrt{|C_i| \times |C_j|}}\right)$$

Note: i and j are the spectral responses of two classes being compared. C is the covariance matrix of the spectral response. μ is the mean vector of the spectral response. ln is the natural logarithm function. T is the transposition function. $|C|$ is the determinant of C.

The J-M distance index was used for quantifying spectral separability between two very similar mangrove classes, *R. apiculata* and *R. mucronata*, to see whether smoothing reduced the separability between the two classes. Due to "the curse of dimensionality" (Bellman, 1961; Hughes, 1968; Fukunaga, 1990), a subset of the whole range of all spectral bands had to be chosen prior to the separability analysis. Thus, spectral bands used for calculating the J-M distance were: 513 nm, 717 nm, 1263 nm, 1385 nm, 1489 nm, and 1669 nm. These principal spectral locations were selected by a feature selection algorithm (please see further details of the algorithm in Siedlecki and Sklansky, 1989; Kavzoglu and Mather, 2002; Yu et al., 2002; Vaiphasa, 2003).

4.3 Results

Each small plot in Figure 4.2 demonstrates the p-values of the comparative t-tests between smoothed and original data for all spectral locations between 400 nm and 2400 nm. In brief, if a p-value at a particular spectral location was as low as 0.01, it demonstrated that the smoothed data were statistically different from the original spectral response with 99% confidence. This illustrated that the original spectra were significantly disturbed by the chosen smoothing method. For example, the 2nd order Savitzky-Golay filter at size 11 resulted in 560 statistically-disturbed bands ($N_{p<0.01}$=560). Additionally, original and smoothed spectral curves of *Avicennia alba* were also included in the figure so as to illustrate the smoothing effect of each filter.

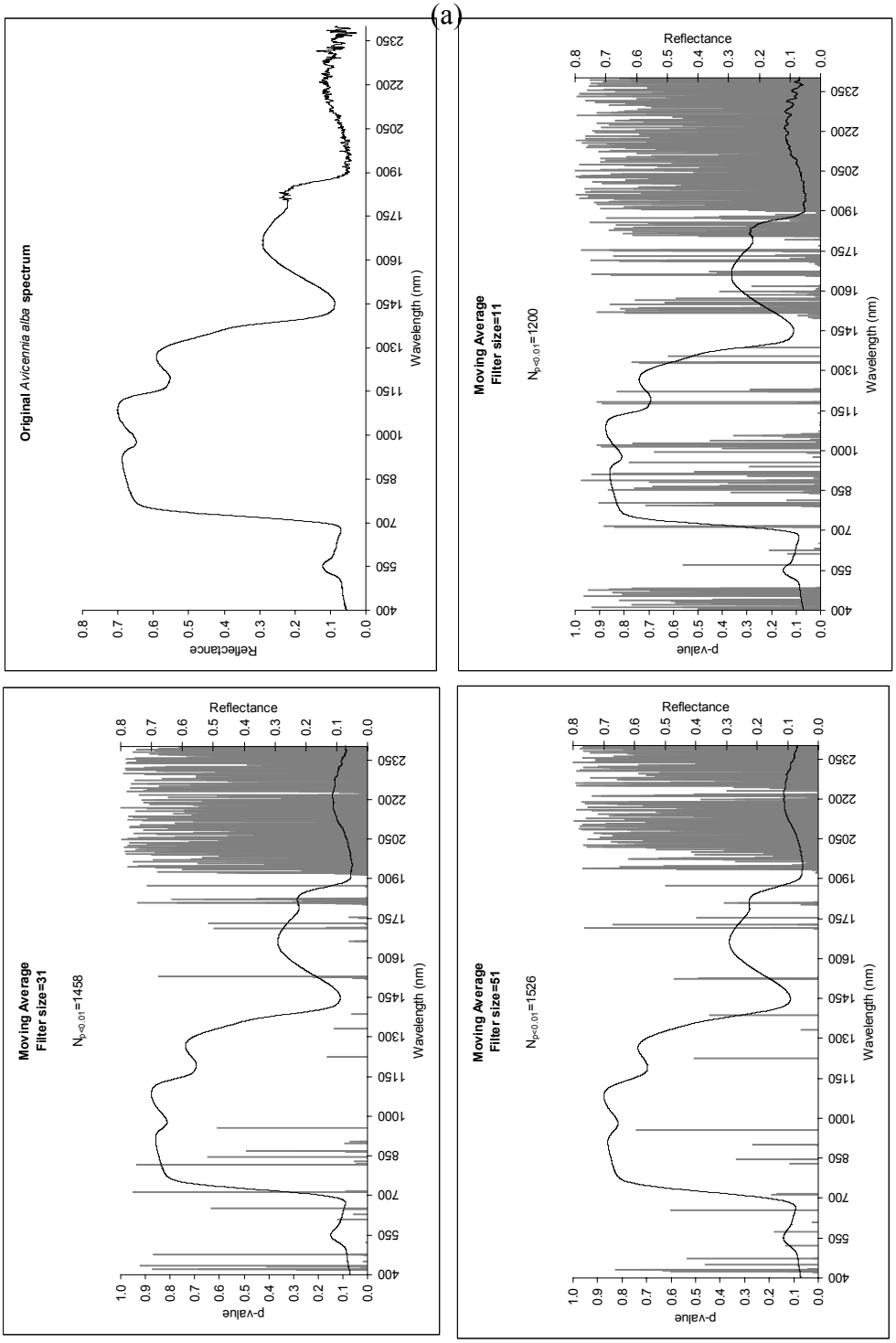
It is important to note that, even if we had exhaustively done the t-test for 2 filter types with sizes between 7 and 51, we only illustrated in Figure 4.2 the results of the 2 filters at 3 different sizes, including 11, 31, and 51 for the purpose of brevity. Instead, the results of the exhaustive experiment for all filter types and sizes were summarized in Figure 4.3. The plot demonstrated the number of statistically different spectral locations with 99% confidence ($N_{p<0.01}$) per filter size. Overall, it was found that the moving average filters statistically disturbed the original spectral response more than the Savitzky-Golay filters. In the case of the moving average filters, the number of statistically different locations steadily increased from 1000 to 1526. In contrast, the number of statistically different locations of the Savitzky-Golay filters increased from 606 to 1153. The general trends in Figure 4.3 of both filter types indicated that bigger filter sizes resulted in higher statistical disturbances

even if there were some exceptions in the trend of the Savitzky-Golay filters where local minima were noticeable.

The reader may note that, even if the number of spectral bands that had p-values less than 0.01 ($N_{p<0.01}$) were recorded for every filter, the p-value level at 0.01 were not used as a critical value α for accepting or rejecting the outcome of one individual filter. Instead, the $N_{p<0.01}$ was only used for the purpose of comparisons between the influences of smoothing filters on the spectral mean. In the case that the reader wish to make a decision to accept or reject the smoothing results of one individual filter on the basis of the reported p-values alone, the reader has to adjust α to the effect of "multiplicity" (i.e., increasing chance of having Type I error as the number of tests grows) (Rothman, 1990; Hsu, 1996; Perneger, 1998; Feise, 2002). Nevertheless, such consideration of one individual filter was not the intension of this study.

Finally, the results of the separability analysis between the two mangrove classes (*R. apiculata* and *R. mucronata*) were demonstrated in Table 4.2. It was found that none of the J-M distance indices of the smoothed data was higher than the J-M distance index of the original data. In other words, spectral smoothing reduced the separability between the two mangrove classes. Additionally, the results in Table 4.2 also reflected the trends in Figure 4.3. The separability index reduced while the filter size increased.

(a)


Original Avicennia alba spectrum
Reflectance
Wavelength (nm)
Moving Average
Filter size=11
N p<0.01=1200
Moving Average
Filter size=31
N p<0.01=1458
Moving Average
Filter size=51
N p<0.01=1526
p-value

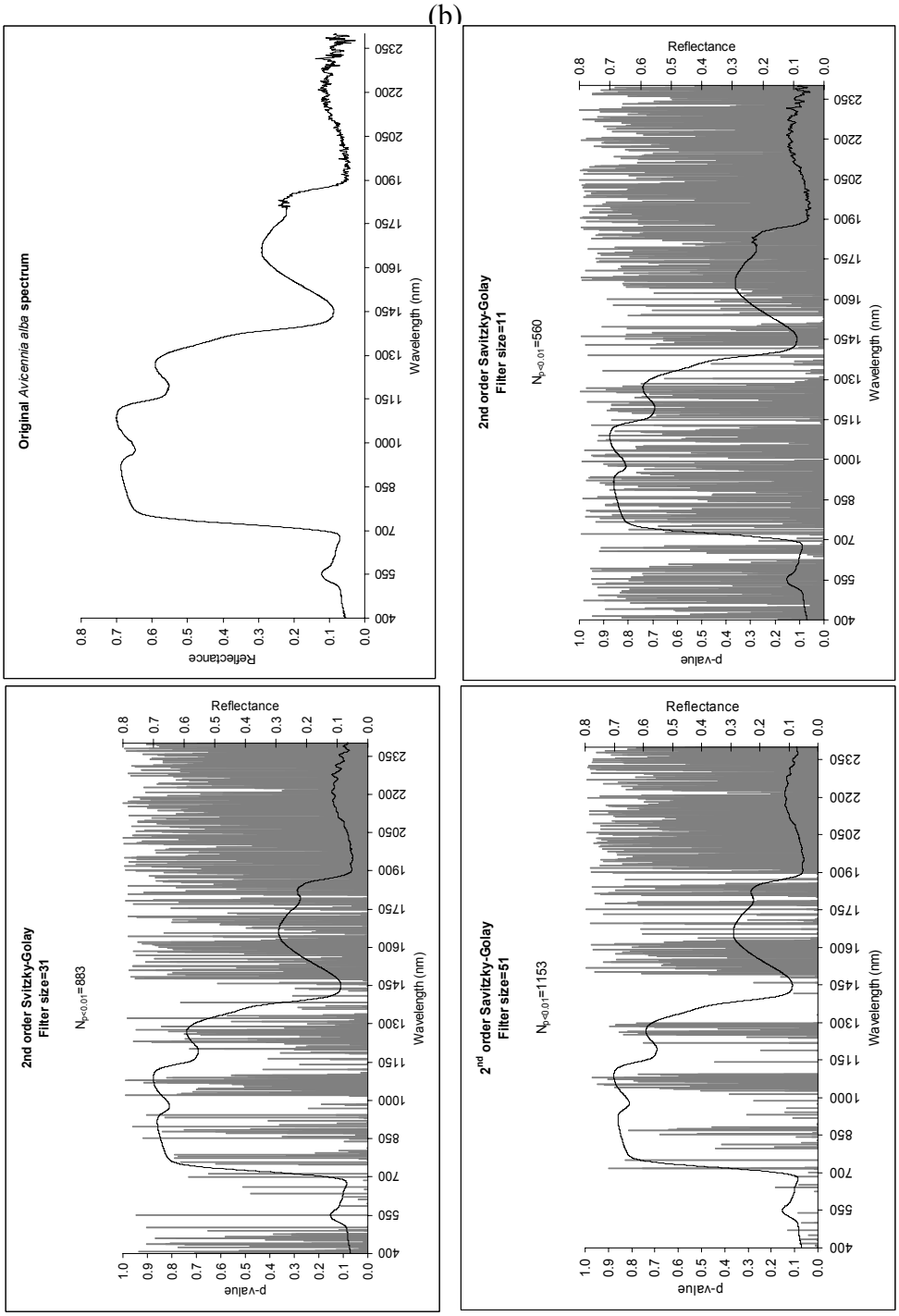


***Figure 4.2**: The effect of smoothing filters on the mean of the leaf-level spectral data of 16 mangrove species: (a) moving average filters; and (b) 2nd order Savitzky-Golay filters.*

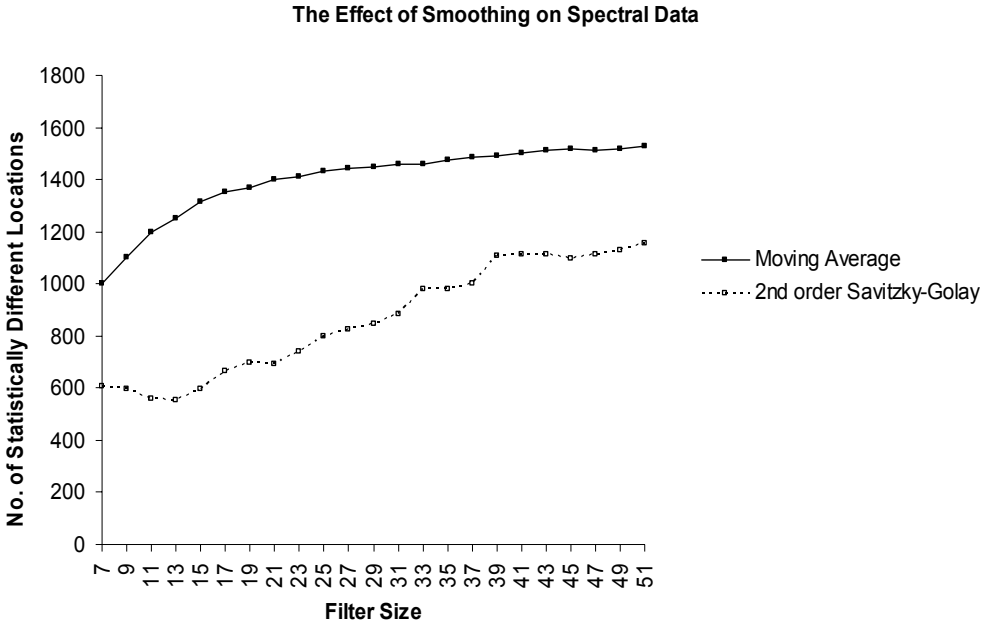


***Figure 4.3**: The number of statistically disturbed locations caused by the moving average and Savitzky-Golay filters*

***Table 4.2:** Comparisons between the J-M distances of two very similar mangrove classes, R. apiculata and R. mucronata, before and after smoothing*

Filter Type	Filter Size	J-M distance
Moving Average	11	1.9311
	31	1.9254
	51	1.9163
2^{nd} order Savitzky-Golay	11	1.9314
	31	1.9307
	51	1.9284
Original data	n/a	1.9315

4.4 Discussion & Conclusion

Our case study demonstrates that the effects of smoothing on the statistical estimate of spectral data are possible to affect negatively the outcome of the subsequent analyses that utilize statistical characteristics of the spectral data. In other words, spectral smoothing does not always enhance the data but make the classes of interest (e.g., *R. apiculata* and *R. mucronata*) more difficult to separate in the spectral feature space. This problem is evident in Table 4.2, as smoothing filters reduce

separability between the two mangrove classes, *R. apiculata* and *R. mucronata*. This argument may also be true for the following spectral discrimination and classification studies that used *ad hoc* measures for selecting smoothing filters without preserving statistical properties of the original data (Gong et al., 2001; Andréfouët et al., 2003; Hochberg and Atkinson, 2003; Schmidt and Skidmore, 2003; Vaughan et al., 2003; Yamano et al., 2003; Castro-Esau et al., 2004; Foody et al., 2004; Rees et al., 2004; Schmidt et al., 2004; Thenkabail et al., 2004).

An instructive example of how to use the t-test method for choosing smoothing filters is described as follows. In Figure 4.2, if an *ad hoc* criterion is used, a practitioner who is challenged to smooth the spectral profile of *A. alba* might face a dilemma to choose between the size-31 moving average filter and the size-51 Savitzky-Golay filter so as to remove the noise in the mid-infrared region of the spectrum (i.e., ≈1900–2400 nm) as well as try not to over-smooth the data. In contrast, when the t-test is applied, its results (Figure 4.2) persuade the practitioner to select the size-51 Savitzky-Golay filter. This is because the size-51 Savitzky-Golay filter does not disturb as many spectral bands as the size-31 moving average filter (i.e., the $N_{p<0.01}$ of the size-51 Savitzky-Golay filter is 1153 while the $N_{p<0.01}$ of the size-31 moving average filter is 1458) even if both filters have similar smoothing effects on the mid-infrared region of the spectrum.

Even though our case study is based on the use of two most-popular filters (moving average and Savitzky-Golay), the t-test method is a universal method that can be used for choosing smoothing filters other than the moving average and Savitzky-Golay. Moreover, this t-test method is not limited to the application for smoothing vegetation spectra. It could be used in other case studies, for example, in selecting appropriate smoothing filters for mineral spectra. However, the number of spectral samples required for t-statistics (e.g., N≈20 samples) is one obvious limitation of this method.

In the continuing study, the notion of using statistical measures (e.g., the t-test) for constraining the selection of spectral smoothing could be expanded to a greater extent. Instead of using only the t-test that constrains the disturbance on the statistical mean, the F-test could also be integrated to limit the disturbance on the statistical variance. In addition, designing smoothing filters that preserve other properties of the original

spectral data such as signal strength (i.e., maximising signal-to-noise ratios) could also be an interesting topic for future research. Only a few pioneer studies have been found in this particular area (Kawata and Minami, 1984, Bruce and Li, 2001; Schmidt and Skidmore, 2004).

In conclusion, hyperspectral smoothing should be used, if necessary, with objective justification of which smoothing technique causes minimal damages to the original data in terms of statistical differences. Therefore, we suggest a comparative t-test as a measure for choosing the right smoothing filter for the hyperspectral data at hand.

Chapter 5*

Ecological Data Integration

* Vaiphasa, C., Skidmore, A.K., de Boer, W.F. (accepted). A post-classifier for mangrove mapping using ecological data. ISPRS Journal of Photogrammetry and Remote Sensing.

Abstract

A global decline in the extent of tropical mangrove forests is one of the most serious problems of the world's coastal ecosystems. This problem results in an increasing demand for detailed mangrove maps at the species level for monitoring mangrove ecosystems and their diversity. Consequently, the goal of this research is to investigate, for the first time, the potential of exploiting mangrove-environment relationships for improving the quality of the final mangrove map at the species level. The relationships between mangroves and the surrounding environmental gradient were incorporated into the mapping process via a typical Bayesian probability model. The Bayesian model functioned as a post-classifier to improve the quality of an already-produced mangrove map. The environmental gradient in use was a GIS layer of soil pH data. Despite the remaining confusion between *R. mucronata* and *S. caseolaris*, the extra investment in collecting the soil pH data paid off. The addition of the soil pH data into the post-classification model helped increase the mapping accuracy from 76.0 % to 88.2%. It is therefore anticipated that the methodology presented in this study can be used as guidelines for producing mangrove maps at a finer level.

Keywords: classification; expert system; multispectral; remote sensing; vegetation

5.1 Introduction

Mangroves colonize tropical and subtropical regions of the world where physico-chemical conditions of the surrounding environment, namely temperature, rainfall, tidal frequency, sedimentation process, freshwater availability, wave and storm shelters, and soil chemistry are suitable for their growth. In addition, such biological conditions as herbivore disturbance, growth competition, and propagule dispersion are also critical factors that determine their survival (Lugo and Snedaker 1974; Hogarth, 1999; Manson et al., 2003).

The existence of mangrove communities benefits the surrounding ecosystem in many ways. For example, they help protect coastal environments from wave and storm surges, maintain coastal water quality, support estuarine food chains, and provide nursery habitats for invertebrates and juvenile fish (Linneweber and de Lacerda, 2002; Manson et al., 2003).

Mangroves forests at large are currently under pressure from anthropogenic activities (e.g., aquaculture, agriculture, urbanization) (Barbier and Sathiratai, 2004). Such activities lead to the alterations of sedimentation rates, freshwater runoff and tidal inundation patterns that subsequently affect the distribution of mangroves (Hogarth, 1999; Linneweber and de Lacerda, 2002). To understand the dynamics of mangrove ecosystems better, detailed assessment of mangrove distribution is required. To date, remote sensing is the only technique that can be used for assessing mangrove change over large areas and that can provide this spatial insight on a repetitive basis (Ramsey III and Jensen, 1996; Green et al., 2000; Ramsey III et al., 2005).

Even though multispectral remote sensing is popularly used at the operational level for mapping mangroves, the information provided by the multispectral sensors is not enough for studying mangrove ecosystems and their diversity owing to the lack of spatial and spectral details (Aschbacher et al., 1995; Ramsey III and Jensen, 1996; Gao, 1999; Green et al., 2000; Sulong et al., 2002; Verheyden et al., 2002; Demuro and Chisholm, 2003; Held et al., 2003). Advanced remote sensing applications are therefore required to serve such demands. The use of ancillary data for vegetation mapping is a prospective candidate (Skidmore et al., 1997a, 1997b; Lehmann and Lenz, 1998; Berberoglu et al., 2004; Comber et al., 2004; Schmidt et al., 2004). In general, ancillary

data are extracted from the relationships between plants and their environment, and then incorporated with remotely sensed data into the mapping model in order to improve the quality of the final map (Lillesand and Kiefer, 2000). Artificial neural networks (Skidmore et al., 1997a; Lehmann and Lenz, 1998; Berberoglu et al., 2004) and expert systems (Skidmore et al., 1997b; Comber et al., 2004; Schmidt et al., 2004) are the two most frequently-used mapping models.

Due to the strong relationships between mangroves and the surrounding environment (Macnae, 1968; Clough, 1982; Semeniuk, 1983; Tomlinson, 1994; Hogarth, 1999), it is normally found that mangrove species occur in a sequencing order. This particular characteristic of mangroves is termed "mangrove zonation." (Tomlinson, 1994; Hogarth, 1999; Vilarrubia, 2000; Satyanarayana et al., 2002). Thus, we hypothesise that these quantifiable spatial relationships between mangroves and their environment can be exploited for mangrove mapping at a finer level (e.g., species level).

Consequently, this study investigates yet another unexplored potential of remote sensing for mangrove mapping. The aim of this study is to test the possibility of integrating ecological data into the mapping model in order to map mangroves at the species level by using a well-established Bayesian expert system (Skidmore et al., 2001; Schmidt et al., 2004). Specifically, the expert system is used as a post-classifier for improving the quality of an already-produced mangrove map. The results before and after applying the post-classifier are statistically compared to see whether there is any improvement in the mapping accuracy.

5.2 Methods

5.2.1 Study site

The study site is located at Cape Talumpuk, Pak Phanang, Nakorn Sri Thammarat, Thailand (8º 31'N, 100º 9'E). A satellite photo of the cape is illustrated in Figure 5.1. A white narrow edge along the eastern end of the cape is an extended sand beach, while the majority of the area on the western side is extensively covered by a 56.8 km^2 dense mangrove forest. There are seven dominant mangrove species: *Avicennia alba*, *Avicennia marina*, *Avicennia officinalis*, *Bruguiera parviflora*, *Rhizophora apiculata*, *Rhizophora mucronata*, and *Sonneratia caseolaris*. The most prominent species is *R. apiculata*. This species covers approximately one

third of the cape. In most cases, each species possesses an almost homogeneous stand of its own community. The morphological characteristics of the area are dominated by a long process of sedimentation from the river at the south end of the cape, as well as by the tidal influence from the east and north. The area is under the influence of a tropical climate with two seasons in a year. The dry period is between February and April, and the rest of the year is dominated by monsoons (Teeratanatorn, 2000).

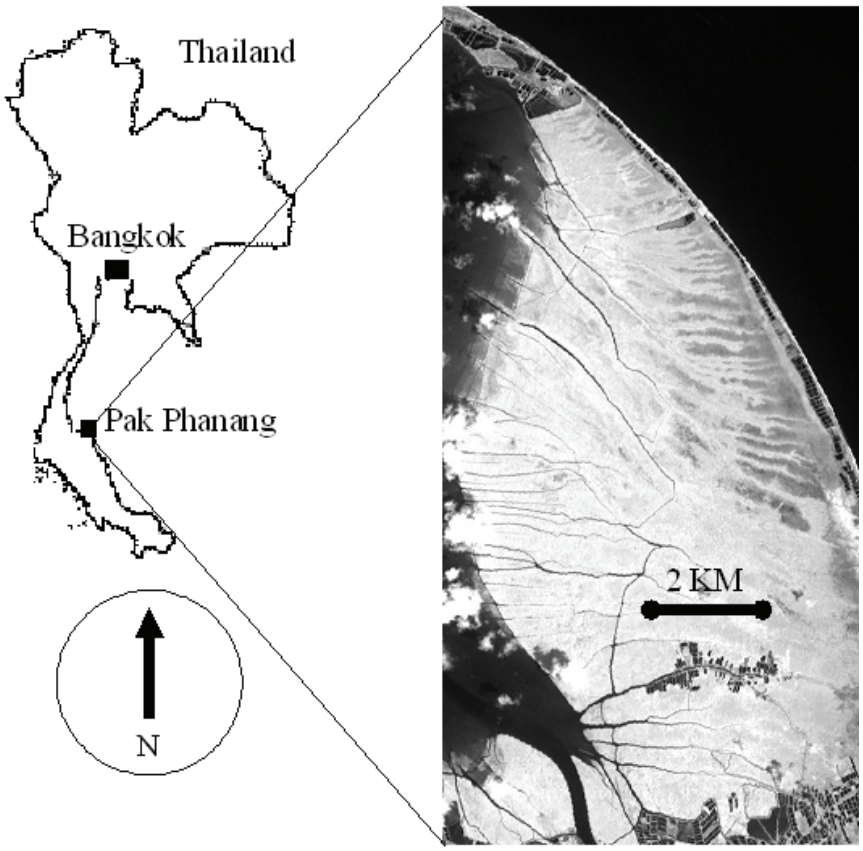


Figure 5.1*: The location of the study area, Cape Talumpuk, Pak Phanang, Nakorn Sri Thammarat, Thailand (8° 31'N, 100° 9'E)*

5.2.2 Field survey

5.2.2.1 Ecological data collection

In this study, soil pH was selected as input ancillary data for the post-classifier (§ 5.2.4.). The soil pH was chosen because it is an important environmental gradient that possesses strong relationships with the spatial distribution of tropical mangrove species (Macnae, 1968; Clough, 1982; Semeniuk, 1983; Hogarth, 1999; Tomlinson, 1994). In other words, a mangrove species is likely to be found in the area where the pH level is suitable for its growth. For example, *Avicennia officinalis* is less abundant in acidic conditions (i.e., pH < 7.0) than *Rhizophora apiculata*.

The field campaign for soil sampling was carried out in March, 2004 (dry season). The distribution of 200 soil sampling stations is shown in Figure 5.2. A stratified random sampling method was used for selecting these 200 locations with two conditions in mind: (i) the sampling locations have to be evenly-distributed throughout the cape; and (ii) the sampling locations must contain as various mangrove species as possible. Both an image generated by unsupervised classification and an existing vegetation map were used as stratification constraints. Good data distribution over the study area was the key requirement for point-data interpolation in the following section (§ 5.2.3.1.).

Figure 5.2*: Two hundred soil sampling stations: 150 plots for interpolation (•) and 50 plots for validation (V)*

A 500-cm^3 soil sample was collected at the centre of each sampling station at 15 cm depth beneath the ground level. Then, all the soil samples were sent to the soil laboratory at the faculty of agricultural technology, King Mongkut's Institute of Technology Ladkrabang (KMITL), Bangkok for further analyses of soil pH. All the samples were air-dried and sieved. Then, a glass-electrode pH meter was used for measuring the samples (McLean, 1982).

Even though using dried soil samples instead of fresh soil samples is not an ideal practice (Ahern et al., 2004), it was necessary because of the hostile conditions of the mangrove forest (i.e., too difficult to perform *in situ* analyses). Therefore, one should note the differences of soil pH values between dried and fresh samples that could affect the outcome of the study. Nevertheless, based on the assumption that there are strong correlations between the two sample types, the differences of soil pH values between dried and fresh samples are expected to have little effect on the outcome of the study. This is because the absolute values of soil pH have not been used in our mapping model (i.e., Eq.1), but they were transformed and used in a relative fashion instead (see the probability values in the expert table (Table 5.2)). Thus, either when dried or fresh samples are used, the relative probability values in the expert table are expected to be similar as both dried and fresh soil pH values are correlated.

5.2.2.2 Mangrove sampling

523 sampling stations were set up for mangrove sampling. The size of each sampling station was 15x15 m^2. Similar to soil sampling, a stratified random sampling method was used for selecting the locations of the sampling stations. Both an image generated by unsupervised classification and an existing vegetation map were used as stratification constraints. The stratified random sampling was chosen to make sure that the stations are evenly distributed through out the mangrove zones (i.e., pioneer, landward, and upper zones). For the sake of brevity, the locations of 523 sampling stations are not graphically shown in this thesis.

The field campaign for mangrove sampling was carried out during the wet season between October, 2003, and January, 2004. Mangrove species composition of all trees (i.e., $\geq$ 2.5 m high) was recorded from each sampling station. The floristic parameters recorded were the species name, tree height, diameter at breast height, crown cover area, and DGPS coordinates in the UTM system. It was found that each station comprised an almost pure vegetation stand (i.e., dominated by one mangrove species only). As a result, the floristic composition of each sampling station was classified into one of the seven available species (*A. alba*, *A. marina A. officinalis*, *B. parviflora*, *R. apiculata*, *R. mucronata*, or *S. caseolaris*). Species identification was under the supervision of the staffs of Royal Thai Forestry Department; taxonomy follows Tomlinson (1994) and

Teeratanatorn (2000). Then, the plots were randomly split into two groups for the purpose of image classification and validation (Table 5.1).

Table 5.1: *Mangrove samples and their coded names used for image classification*

Vegetation Types	Training samples	Testing samples
Avicennia alba (AA)	19	20
Avicennia marina (AM)	28	29
Avicennia officinalis (AO)	39	39
Bruguiera parviflora (BP)	27	27
Rhizophora apiculata (RA)	69	70
Rhizophora mucronata (RM)	45	45
Sonneratia caseolaris (SC)	33	33
Total	260	263

5.2.3 Input data for the post-classifier

5.2.3.1 Soil pH interpolation

The post-classifier in use required input ancillary data (i.e., soil pH) in a format of a GIS map. In this study, geo-statistical interpolation was used for producing the soil pH map. The reliability of geo-statistical interpolation is well-known among soil scientists (Burgess and Webster, 1980; Gajem et al., 1981; Yost et al., 1982; Samra et al., 1988; Yates et al. 1988; Ardahanlioglu et al., 2003).

Prior to the interpolation, the soil pH data were thoroughly checked for outliers, but none was found. Next, directional semi-variograms were used to check for any sign of directional anisotropy, but it turned out that the distribution of the soil pH data was isotopic. Then, all 200 data points were separated into two groups, the interpolation and validation sets. The distribution of both interpolation (150 points) and validation (50 points) points is shown in Figure 5.2. Then, the interpolation points were interpolated over the study area to create a GIS layer of soil pH. The interpolation method in use was an ordinary kriging method based on a spherical model. The point-data interpolation procedure of this study followed the guideline of Isaaks and Srivastava (1989) using commercial software (ILWIS v.3.0, ITC institute, the Netherlands). The interpolation result is demonstrated in Figure 5.3, and its variogram model can be found in Appendix III. Finally, a Root Mean Square (RMS) error of the

interpolation map was calculated using the validation points. It was found that the RMS error of the interpolation result was pH ±0.43.

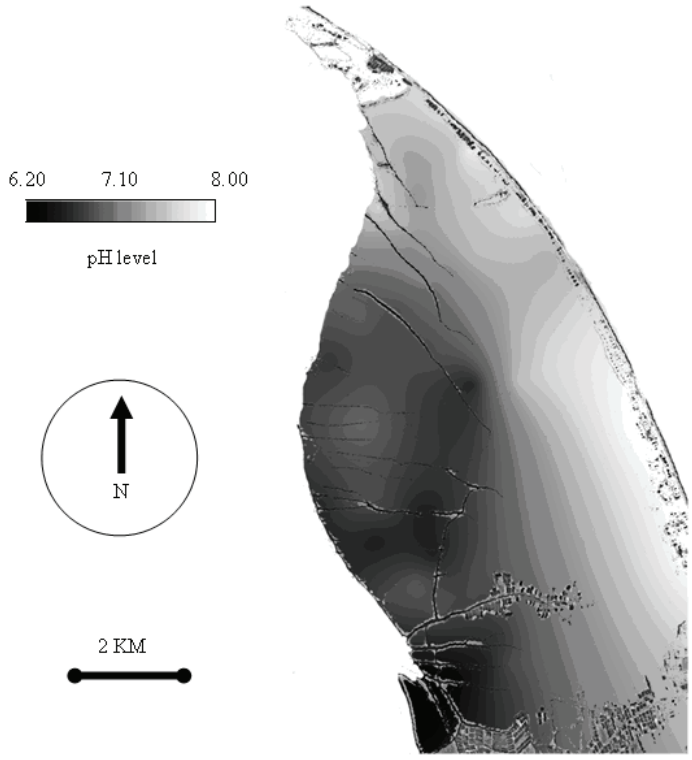


***Figure 5.3**: The interpolation map of soil pH using an ordinary kriging method*

5.2.3.2 Plant-environment relationships

Relationships between soil pH and mangrove species are illustrated in Table 5.2 in terms of conditional probability values. Each value in the table indicated the chance to find a particular environmental condition (e.g., a pH range) at an area where the mangrove species was known. Each column of the table was actually the normalised histogram of soil pH per each mangrove species (i.e., the sum of all elements in each column = 1.00). In this study, the mangrove-pH relationships derived from the 200 points of the soil pH data were used for the construction of Table 5.2. The construction of Table 5.2 was, however, not restricted to the use of empirical data collected from the field. Subjective knowledge of mangrove experts could be used to replace (or improve) the values in the table if appropriate. Subsequently, Table 5.2 was termed "the expert table".

Table 5.2: *Mangrove-pH relationships in terms of conditional probability values (the expert table)*

pH	AA	AM	AO	BP	RA	RM	SC
6.20-6.39	0.00	0.00	0.00	0.00	0.26	0.50	0.10
6.40-6.59	0.00	0.00	0.00	0.00	0.37	0.07	0.30
6.60-6.79	0.10	0.00	0.00	0.00	0.11	0.00	0.00
6.80-6.99	0.18	0.00	0.00	0.00	0.24	0.22	0.40
7.00-7.19	0.27	0.00	0.00	0.12	0.02	0.00	0.20
7.20-7.39	0.45	0.00	0.00	0.22	0.00	0.21	0.00
7.40-7.59	0.00	0.40	0.00	0.44	0.00	0.00	0.00
7.60-7.79	0.00	0.40	0.00	0.22	0.00	0.00	0.00
7.80-7.99	0.00	0.20	1.00	0.00	0.00	0.00	0.00
Total	1.00	1.00	1.00	1.00	1.00	1.00	1.00

5.2.3.3 The classified image

An already-classified satellite image was also required for the calculation of the post-classifier. In this study, the satellite image was originally taken by the ASTER instrument (Terra spacecraft) in the dry season on March, 6, 2002. After eliminating the off-nadir band, the 14-band satellite image was then corrected for radiometric and geographic distortion and resampled into a UTM coordinate system (zone 47 North) with a 15x15m^2 pixel size. Subsequently, the image was classified using a maximum likelihood classifier (MLC). The classifier was trained using the mangrove species data of 260 training plots (Table 5.1). As a result, each pixel was assign to one of the seven dominant mangrove classes. The result of the MLC classification is demonstrated in Figure 5.4a with its confusion matrix (Table 5.3). 263 independent testing samples were used for the calculation of the confusion matrix. Additionally, the MLC did not only produce a classified image, but also seven rule maps containing convertible information of the Chi-square-based likelihood of all mangrove species. The MLC used in this study is a function in commercial software (ENVI v.3.6, RSI, Inc).

The time difference between the acquisition of the image and the collection of soil samples should be noted even if the collection of both data was carried out in the same season (i.e., dry season) without any extreme weather condition reported in between. The time difference bars both data from being used in some applications that require a precise snapshot of what happens to the mangroves and soil pH conditions simultaneously. However, the drawback of the time difference is deemed to be of little significance to this study since this study exploited

mangrove-soil relationships in a relative sense. In other words, the soil pH data were not used for pinpointing an exact location of a mangrove species but for comparing relatively between the likelihood of seven different mangrove species.

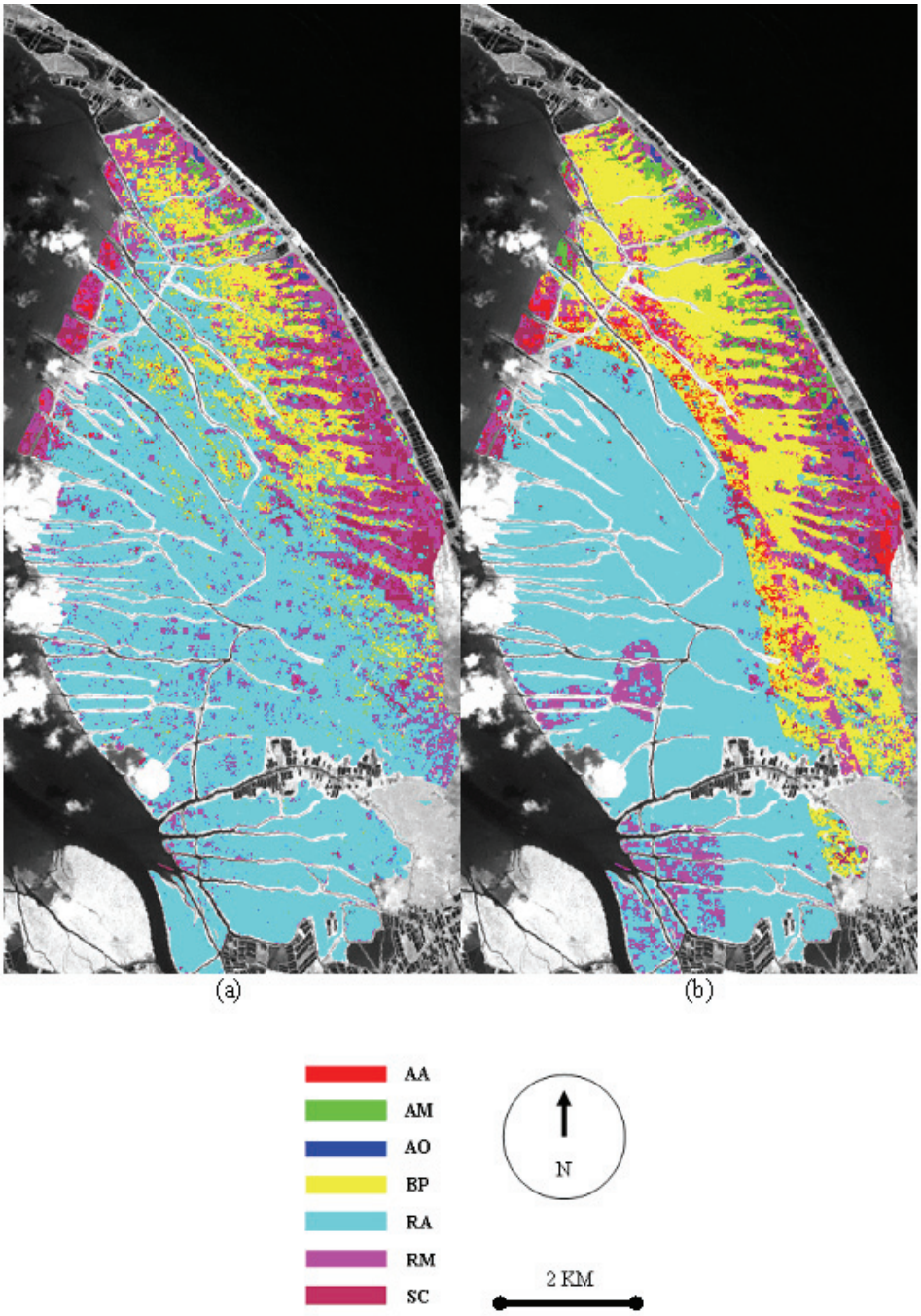


***Figure 5.4**: A classification result when using MLC only (a) and a classification result when applying the post-classifier (b). Non-mangrove areas are shown in grey.*

***Table 5.3**: Accuracy tables before applying the post-classifier (overall accuracy = 76.0%)*

Class	AA	AM	AO	BP	RA	RM	SC	Total
AA	10	0	0	0	0	0	4	14
AM	0	25	0	0	0	0	0	25
AO	0	0	33	0	0	0	0	33
BP	0	0	0	19	0	0	0	19
RA	2	0	6	8	60	10	0	86
RM	6	4	0	0	10	28	4	52
SC	2	0	0	0	0	7	25	34
Total	20	29	39	27	70	45	33	263

Class	Producer's accuracy	User's accuracy
AA	50.0	71.4
AM	86.2	100.0
AO	84.6	100.0
BP	70.4	100.0
RA	85.7	69.8
RM	62.2	53.8
SC	75.8	73.5

5.2.4 The post-classifier

The post-classifier in use is a computer programme in IDL language developed by the institute (ITC, Enschede, the Netherlands). It is being employed at the operational level for mapping salt marsh vegetation in the Netherlands (Skidmore et al., 2001; Schmidt et al., 2004). The post-classifier is designed as a pixel-based tool in which the post-classification process is carried out at one pixel at a time. The key mechanism of the post-classifier is Bayes' Rule (Bayesian). $B_1, B_2, \ldots, B_n$ are independent events that are members of a sample space (S) and $\bigcup_{i=1}^{n} B_i$ = S. Given a situation A and $P(A) \neq 0$, the Bayes' rule can then be illustrated in Eq. 1.

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^{n} P(A|B_i)P(B_i)} \quad ; i=1,2,\ldots,n \qquad \text{[Eq.1]}$$

Each term of Eq. 1 can be elaborated as follows.

- $P(B_j|A)$ is the output term of the equation. It describes the probability that a mangrove species (B_j) occurs when a soil pH condition (A) of the area is known.

- $P(A|B_j)$ describes the chance of finding a particular soil pH condition (A) at an area where the mangrove species (B_j) is known. Thus, $P(A|B_j)$ can be looked up from the expert table (Table 5.2).

- $P(B_j)$ is *a priori* probability of occurrence of a mangrove species (B_j) at a certain location. In this study, $P(B_j)$ is the Chi-square based probability value that is converted from the MLC rule map.

- The dividing term of Eq. 1 is the sum of all possible cases (i.e., $i=1, 2,..., n$) where n is the total number of mangrove species under study. In this study, the total number of species is 7 (i.e., n = 7).

At each pixel of the already-classified image, the post-classifier repeatedly calculates the value of $P(B_j|A)$ for every mangrove class (i.e., from $j=1$ to $j=7$). Then, all of the values are compared, and the pixel is finally re-labelled as the mangrove species that possesses the highest $P(B_j|A)$ value.

5.2.5 Statistical test

In this study, a z-test based on KHAT or K statistics in Eq. 2 was used for comparing between the results before and after applying the post-classifier. Given the null hypothesis H_o: K_1-K_2=0 and the alternative H_1: K_1-K_2≠0, H_o is rejected if $Z \geq Z_{\alpha/2}$. A 95% confidence limit was used as a critical value (i.e., $Z_{\alpha/2}$=1.96). Congalton and Green (1999) provide more mathematical details about the test.

$$Z = \frac{|K_1 - K_2|}{\sqrt{Var(K_1) + Var(K_2)}} \qquad \text{[Eq.2]}$$

5.3 Results

The classification accuracy before applying the post-classifier was reported in Table 5.3, and its corresponding classified map was illustrated in Figure 5.4a. The overall accuracy of the map (OA) and an estimate of kappa statistics (K) were OA_{MLC}=76.0% and K_{MLC}=0.71, respectively. For comparison, the post-classified result was illustrated in Figure 5.4b and Table 5.4. Overall, it was found that the post-classifier helped improve the mapping accuracy. The overall accuracy and an estimate of kappa statistics after applying the post-classifier were OA_{pH}= 88.2% and K_{pH}= 0.86, respectively.

Before applying the post-classifier (Table 5.3), the members of the Rhizophoraceae family (BP, RA, and RM) were spectrally confused with the other mangrove species as well as within the family. This spectral confusion was responsible for almost all misclassified pixels in every column of Table 5.3.

***Table 5.4**: Accuracy tables after applying the post-classifier (overall accuracy = 88.21%)*

Class	AA	AM	AO	BP	RA	RM	SC	Total
AA	15	0	0	1	0	1	0	17
AM	0	29	0	0	0	0	0	29
AO	0	0	39	0	0	0	0	39
BP	0	0	0	26	0	0	0	26
RA	1	0	0	0	66	2	0	69
RM	2	0	0	0	4	24	0	30
SC	2	0	0	0	0	18	33	53
Total	20	29	39	27	70	45	33	263

Class	Producer's accuracy	User's accuracy
AA	75.0	88.2
AM	100.0	100.0
AO	100.0	100.0
BP	96.3	100.0
RA	94.3	95.7
RM	53.3	80.0
SC	100.0	62.3

By comparing between the situations before and after applying the post-classifier (i.e., between Table 5.3 and Table 5.4), it was found that the

spectral confusion caused by the members of Rhizophoraceae in every column was improved except in the RM column. This resulted in the improvement in the producer's accuracy of every mangrove class but RM. In addition, the result of the z-test based K statistics confirmed that the accuracy improvement was statistically significant. The z value reported was 3.76 and it was higher than the critical value (i.e., $Z_{\alpha/2}$=1.96).

The significant improvements in the mapping accuracy were also noticeable graphically by comparing between Figure 5.4a and Figure 5.4b. There were two distinct features that were responsible for most of the improvements. First, the "salt-and-pepper" patterns of RM in Figure 5.4a were mostly eliminated by the post-classifier and resulted in more homogeneous patterns in Figure 5.4b. This effect of the post-classifier helped reduce most of the class confusion between RM and the other two members of the Rhizophoraceae family (RA and BP). Second, the post-classifier created a distinctive pH limit in the middle of Figure 5.4b. This pH limit (pH ≈ 7.20) helped separate RA, RM, and SC from the rest of the mangroves, as the three species were more susceptible to alkaline conditions.

5.4 Discussion & conclusion

Another unexplored potential of remote sensing for mangrove mapping has been unveiled in this study. Overall, the results confirm that integrating a soil-related ecological parameter such as soil pH into the mapping process significantly improves the quality of the final mangrove map. The influence of soil pH via the Bayesian model of the post-classifier satisfactorily increases the overall mapping accuracy by 12% (i.e., from 76.0% to 88.2%) that is acceptable by the USGS classification standard (Anderson et al., 1976). The superiority of the addition of ancillary data over the solo use of multi-spectral data in mapping tropical mangrove species of Cape Talumpuk helps realise the possibility that the methodology presented in this study can be used as guidelines for producing mangrove maps at the community or species level.

The difficulty in discriminating the members of the Rhizophoraceae family in Table 5.3 may not be neglected. It is interesting to find that the spectral confusion between the members of the Rhizophoraceae family and the other mangrove species reported in this study is reconciled with the results of the laboratory experiment conducted by Vaiphasa and

Ongsomwang (2004). These authors studied the spectral separability indices between various tropical mangrove species under laboratory conditions. It was found that the laboratory spectra of the members of the Rhizophoraceae family were not only similar to the other mangrove families but also among themselves. As a result, it may be concluded that using spectral information alone is not adequate for mangrove species classification if the study area is dominated by Rhizophoraceae. In contrast, additional information (e.g., the soil pH map) is needed for resolving such spectral confusion.

Lastly, despite the optimism of the overall result, it should be noted that the addition of the soil pH data into the mapping process was not able to resolve the confusion between *R. mucronata* and *S. caseolaris* (please compare between the *R. mucronata* (RM) columns in Table 5.3 and Table 5.4). In fact, the addition of soil pH led to more confusion between the two species. As a result, it is recommended that extra treatments, in addition to the use of the soil pH map, should be done to improve specifically the separability between the two classes. Since both species possess totally different leaf textures, one possible solution is to exploit the leaf texture information extracted from aerial photographs.

In summary, this study explores one step beyond the existing studies. It is the first time that the possibility of exploiting mangrove-environment relationships for improving the quality of the final mangrove map is investigated. Despite the remaining confusion between *R. mucronata* and *S. caseolaris*, the integration of ecological information such as soil pH into the mangrove mapping process is worth the extra fieldwork effort. It significantly increases the mapping accuracy from 76.0 % to 88.2%. Therefore, we anticipate that the methodology presented in this study can be used as guidelines for producing a mangrove map at a finer level (e.g., community or species level).

Chapter 6

The Synthesis

6.1 Introduction

Mangrove forests are part of the coastal environment and stretch throughout the tropics and sub-tropics of the world (Tomlinson, 1994; Hogarth, 1999). They cover up to 75% of the world's tropical coastlines (Spalding et al., 1997). Their importance is recognisable in such aspects as forestry, fisheries, and environmental conservation (Barbier and Sathiratai, 2004). Similar to many other natural resources, mangroves are declining because of the influence of natural disturbance and human intervention. This has negative effects on economic development and ultimately on the environment as a whole (Barbier and Sathiratai, 2004). These repercussions have subsequently drawn considerable attention to the conservation and management of this unique estuarine ecosystem (Ramsar Convention, 1971; Linneweber and de Lacerda, 2002).

Precise and up-to-date spatial information on the current status of mangroves is a prerequisite for the sustainable conservation of mangrove ecosystems. It is almost impossible to gather this information by using traditional field surveys because mangrove swamps are extremely difficult to access. Fortunately, it has been discovered that remote sensing technology is a promising solution to this problem of accessibility (Green et al., 2000; Held et al., 2003).

To date, the use of remote sensing technology for gathering spatial information from mangrove forests (e.g., mapping and monitoring) at the community levels has been extensive (Aschbacher et al., 1995; Ramsey III and Jensen, 1996; Gao, 1999; Sulong et al., 2002), but the application at the species level, which is necessary for studying mangrove species diversity, is still inconclusive (Demuro and Chisholm, 2003; Held et al., 2003). Therefore, this thesis further explores the capability of remote sensing technology to map mangroves at the species level, using two important ingredients: (i) the use of narrow-band hyperspectral data and (ii) the integration of ecological knowledge of mangrove-environment relationships into the mapping process.

The key objectives of this study are:

(1) to demonstrate the potential of hyperspectral technology for discriminating mangroves at the species level

(2) to test whether a form of genetic algorithms can be used for selecting a meaningful subset of spectral bands that maintains spectral separability between mangrove species
(3) to investigate one of the most popular methods of reducing noise levels in hyperspectral data (i.e., spectral smoothing), as well as propose a technique for selecting an appropriate smoothing filter for the data at hand
(4) to test whether mangrove-environment relationships can be exploited in order to improve the mapping accuracy.

6.2 The main results

6.2.1 Hyperspectral data for mangrove discrimination

Although multispectral sensors are the most cost-effective remote sensing solutions for mangrove mapping (Aschbacher et al., 1995; Ramsey III and Jensen, 1996; Gao, 1999; Green et al., 2000; Sulong et al., 2002; Held et al., 2003), they are still limited to applications for mapping at the regional scale. One of their major constraints is the lack of spectral detail.

Unlike multispectral sensors, hyperspectral sensors that possess 100 or more narrow spectral bands between the visible and shortwave infrared regions have already proved to have the potential for discriminating terrestrial plants at the species level (Cochrane, 2000; Schmidt and Skidmore, 2003). Nevertheless, the hyperspectral research on mangroves published to date (Green et al., 2000; Demuro and Chisholm, 2003; Held et al., 2003; Hirano et al., 2003) remains inconclusive when it comes to using the technology for tropical mangrove species discrimination.

The prerequisite study described in Chapter 2 took the investigation into this issue one step further. It was a laboratory investigation to see whether hyperspectral data contained adequate spectral information for discriminating mangroves at the species level. The study helped us in deciding whether to invest in the expensive acquisition of airborne or satellite hyperspectral data. In brief, the spectral responses of 16 tropical mangrove species were recorded from the leaves, using a 2151-band spectrometer under laboratory conditions. Then, the mangrove spectra at every spectral location were statistically compared using one-way ANOVA to see whether they significantly differed. Finally, the spectral

separability between each pair of mangrove species was calculated using the J-M distance in order to confirm the results.

It turned out that the leaf spectra of different mangrove species were statistically different at most spectral locations, with a 95% confidence level (Figure 6.1). Specifically, the total number of spectral bands that had p-values < 0.05 was 1941, of which 477 bands had p-values < 0.01. Moreover, the J-M distance indices calculated for all pairs of the mangrove species also confirmed that the mangroves were spectrally separable (i.e., J-M distance ≥ 1.90), except the pairs that comprised members of Rhizophoraceae (Table 6.1).

Overall, the results encourage further investigation into the use of airborne and satellite hyperspectral sensors for discriminating mangrove species. However, one should bear in mind the difficulty in discriminating the members of the Rhizophoraceae family. Since the Rhizophoraceae family usually dominates tropical mangrove forests, difficulty in discriminating these mangroves is expected when implementing the on-board hyperspectral sensors.

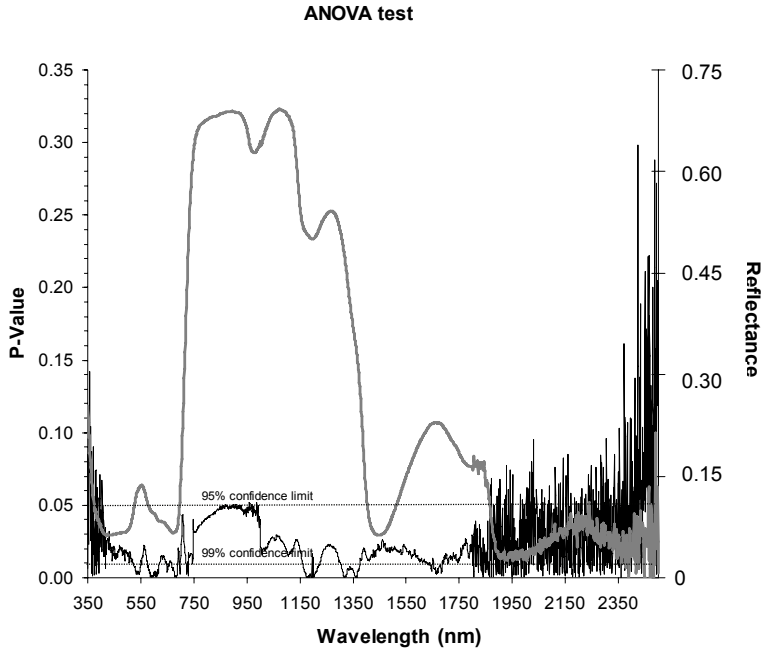


Figure 6.1*: The plot of p-values of the ANOVA test (black line) showing against a laboratory reflectance of Rhizophora apiculata (grey line)*

Table 6.1: *The J-M distances between all pairs of 16 mangrove species (120 pairs in total). The species names are coded in Chapter 2. The pairs that possess separability levels lower than 1.90 are highlighted in grey. Mangrove species are grouped by family name.*

		Avicenniaceae	Pteridaceae	Rhizophoraceae						Euphorbiaceae	Sterculiaceae	Combretaceae		Wurmb	Asteraceae	Sonneratiaceae	Meliaceae
		AVA	ACA	BC	BG	BP	CT	RA	RM	EA	HL	LL	LR	NF	PI	SO	XG
Avicenniaceae	AVA																
Pteridaceae	ACA	1.99															
Rhizophoraceae	BC	1.99	2.00														
	BG	2.00	1.56	1.99													
	BP	1.99	1.99	1.99	1.82												
	CT	2.00	1.99	1.97	1.99	1.90											
	RA	1.99	1.99	1.93	1.99	1.99	1.99										
	RM	2.00	1.99	1.72	1.99	1.73	1.85	1.86									
Euphorbiaceae	EA	1.94	1.99	2.00	1.99	1.99	2.00	1.99	2.00								
Sterculiaceae	HL	1.99	1.99	2.00	1.94	1.99	1.99	1.99	1.99	1.98							
Combretaceae	LL	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00						
	LR	2.00	2.00	2.00	2.00	2.00	2.00	1.99	2.00	2.00	2.00	1.99					
Wurmb	NF	1.99	1.99	2.00	1.99	1.99	2.00	1.99	1.99	1.99	1.99	2.00	2.00				
Asteraceae	PI	2.00	2.00	1.98	1.99	1.99	1.87	1.99	1.89	2.00	2.00	1.99	2.00	2.00			
Sonneratiaceae	SO	1.99	1.99	1.99	1.95	1.84	1.99	1.98	1.99	1.99	1.99	1.99	2.00	1.99	1.99		
Meliaceae	XG	1.97	1.99	1.99	1.82	1.96	2.00	1.99	1.99	1.98	1.99	2.00	2.00	1.99	2.00	1.99	

6.2.2 Hyper-dimensionality problems

The high-dimensional characteristics of hyperspectral data can trigger the phenomenon known as "the curse of dimensionality" (Bellman, 1961). This phenomenon causes imprecise class estimates in the spectral feature space, which result in low output classification accuracy (Bellman, 1961; Hughes, 1968). Consequently, this situation demands more training samples in order to construct better class estimates, thereby dramatically increasing the cost of the field survey.

Chapter 3 demonstrated an alternative to the existing account of feature selection tools to deal with the curse of dimensionality. This alternative feature selection tool was a form of genetic search algorithms (GA). Pioneering work that gained significant insight into this issue was carried out by Siedlecki and Sklansky (1989). The authors reported that the GA-based band selector performed better than many other popular band selection algorithms (e.g., branch and bound search, exhaustive search, and sequential forward selection). The authors rigorously tested their hypothesis, using a synthetic error model instead of real remotely sensed data in order to eliminate the variables (e.g., sample size, the number of spectral bands, and the number of classes of interest) that could have biased the outcome. Further evidence of the success of GA-based band selection tools can be found in recent hyperspectral remote sensing publications (Yu et al., 2002; Fang et al., 2003; Kooistra et al., 2003, Cogdill et al., 2004).

In contrast to the acid tests completed so far (Lofy and Sklansky, 2001; Kavzoglu and Mather, 2002; Yu et al., 2002; Ulfarsson et al., 2003), the work presented in Chapter 3 was the first time that the GA-based band selector had been tested on spectrometer records of very high dimensionality, comprising 2151 bands of leaf spectra of 16 tropical mangrove species. It turned out that the GA-based band selector was able to cope with spectral similarity at the species level. It selected spectral bands that related to the principal physico-chemical properties of plants (Curran, 1989; Elvidge, 1990; Kumar et al., 2001) and, simultaneously, maintained the separability between species classes at an 80% level of classification accuracy. The selection result is shown in Figure 6.2.

It is worth noting that only one of the six spectral locations illustrated in Figure 6.2 is in the visible region where electro-magnetic energy interacts with mangrove leaf pigments (e.g., chlorophylls, carotenoids) (Menon

and Neelakantan, 1992; Basak et al., 1996; Das et al., 2002). This outcome may be interpreted as an indication that the spectral responses of mangrove pigments contain less important spectral information for mangrove species discrimination than the information from the spectral responses of the other leaf components that interact with electro-magnetic energy at longer wavelengths. Unfortunately, the results of studies so far on the physico-chemical properties of leaves of different mangrove species are still inconclusive when it comes to pinpointing which components of mangrove leaves are spectrally separable (Menon and Neelakantan, 1992; Tomlinson, 1994; Basak et al., 1996; Das et al., 2002). A thorough comparative study is therefore recommended in order to confirm this part of the findings.

Lastly, the capability of the GA-based band selector to cope with a very complex band selection problem reported in Chapter 3 encourages the future use of the band selector for detecting spectral bands that show strong vegetation responses to different physico-chemical treatments (e.g., nitrogen, illumination) in both laboratory and field scenarios. It is anticipated that the GA-based band selector will be a viable alternative to the statistical and derivative analyses popularly used at the moment (Tsai and Philpot, 1998; Mutanga et al., 2003).

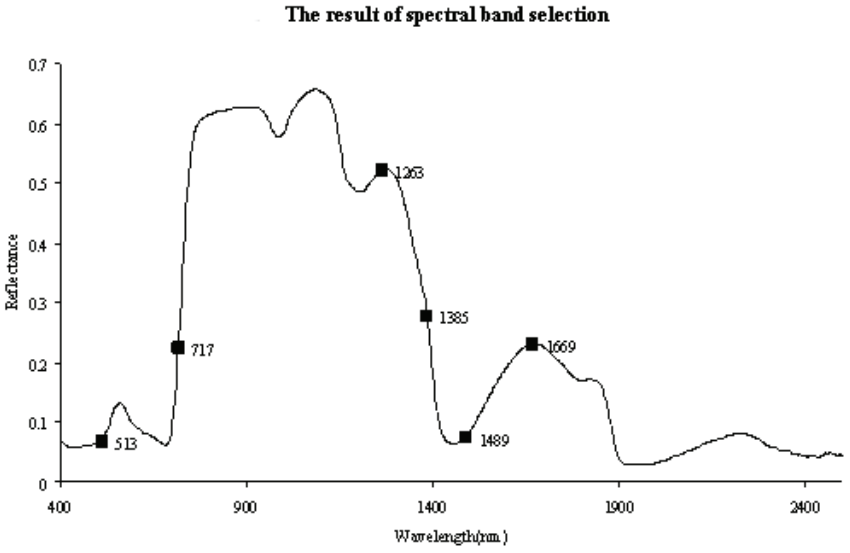


***Figure 6.2**: Six average spectral positions selected by the GA-based feature selection algorithm*

6.2.3 Noise levels

Another important problem when using hyperspectral data is low signal-to-noise ratios. This problem is normally solved by applying spectral smoothing filters to each spectral profile in order to create convolutions of spectral values, thereby reducing the noise level. According to the review in Chapter 4, however, it was found that at least 20 recently published reports used subjective *ad hoc* inspections as their measures for selecting filter types and the parameters. In other words, they did not employ any strict optimizing criterion to select suitable smoothing filters for their studies. It is believed that the *ad hoc* approach is not the most appropriate way. Furthermore, it is hypothesized that smoothing filters can cause significant changes to the statistical properties (e.g., mean) of spectral data (see Figure 6.3). This statistical disturbance could then affect the outcome of subsequent analyses (e.g., maximum likelihood classifier, Jeffries-Matusita distance) that are based on statistical estimates of the data.

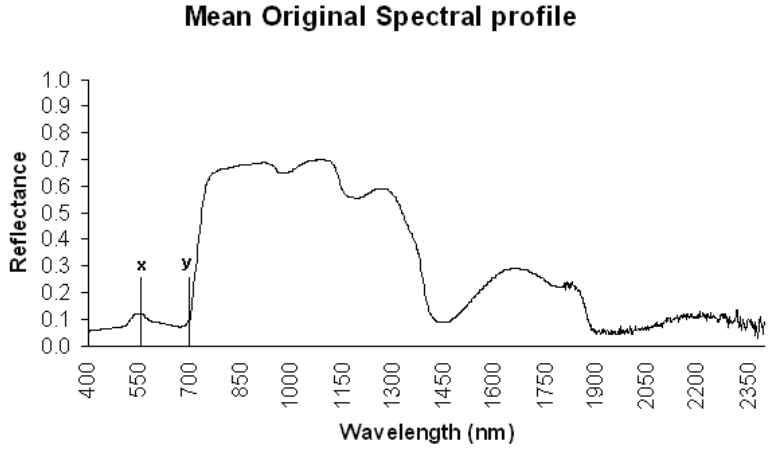
In Chapter 4, it was proved that the above hypothesis is true (i.e., the effect of the smoothing disturbances on the class statistics is evident in Table 4.2). Thus, if preserving statistical properties of the original hyperspectral data is desired, smoothing filters that cause the minimum disturbance to the statistical properties of the original data should be objectively applied. One possible solution is to use a simple comparative t-test as a post-smoothing measure for choosing an optimum smoothing filter for the hyperspectral data at hand.

The purpose of the post-smoothing method (the t-test) proposed in Chapter 4 was to control the effect on the statistical estimate of popular smoothing filters such as the moving average and Savitzky-Golay that have no built-in ability to preserve the original statistical properties of the spectral data (i.e., these popular filters are based on underlying non-parametric mathematics that does not preserve statistical estimates of class information) (Kay, 1993). However, the post-smoothing method could have been omitted if the filter used had had the ability to preserve the statistical properties of the data. This ability could be achieved by designing a specialized smoothing filter, using an estimation theory such as maximum likelihood estimation (Oppenheim and Schafer, 1975; Kay, 1993; Deng and Shen, 1997). In this way, class statistics of the original spectral data could be preserved by the filter after smoothing without the

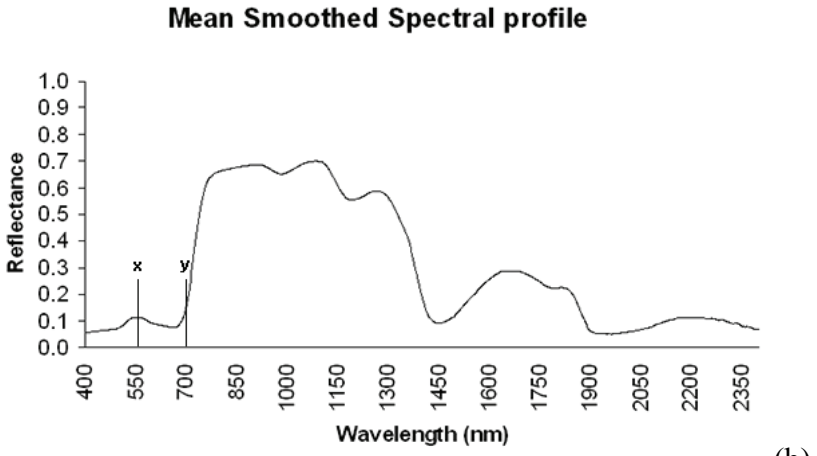
need for the post-smoothing statistical verification (i.e., no need to use the t-test method presented in Chapter 4).

Similarly, if preserving other properties of the spectral data, including signal phases or signal-to-noise ratios, is desired, specialized smoothing techniques could be used as a replacement for the generic techniques (e.g., moving average and Savitzky-Golay). With respect to the first case, preserving signal phases is particularly desirable for specific applications such as spectral derivative analyses. In this regard, specialized methods such as the Fourier transformation and wavelet decomposition would be the right choice because it has been proved that they can preserve the signal phase better than the generic smoothing methods (Curran et al., 1992; Schmidt and Skidmore, 2004). As for the second case, signal-to-noise ratios could be preserved by specific filters such as the Kawata-Minami filter, which is equipped with a least mean-square criterion that helps to maximize the signal-to-noise ratio (Kawata and Minami, 1984; Tsai and Philpot, 1998).

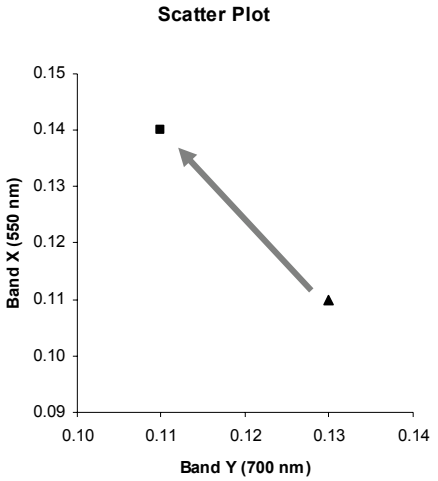
Nevertheless, according to the review in Chapter 4, application-specific smoothing methods such as the Fourier and wavelet transformation and the Kawata-Minami filter are less popular in the field of remote sensing than generic methods such as the moving average and Savitzky-Golay filters. Consequently, tailoring specialized smoothing filters to specific requirements as a replacement for generic methods could be an interesting topic for future research.



(a)



(b)



(c)

Figure 6.3*: An average spectral profile of plant leaves (a) before smoothing and (b) after smoothing; (c) a scatter plot of two principal wavelengths before (triangle) and after (square) smoothing*

6.2.4 Utilizing mangrove-environment relationships

Spatial relationships between mangroves and the environment are well known (Macnae, 1968; Clough, 1982; Semeniuk, 1983; Tomlinson, 1994; Hogarth, 1999). These relationships result in the mangrove zonations that are usually found in tropical mangrove forests (Tomlinson, 1994; Hogarth, 1999; Vilarrubia, 2000; Satyanarayana et al., 2002). As a result, it is hypothesized in this thesis that these quantifiable spatial

relationships between mangroves and their environment can be exploited for mangrove mapping.

In Chapter 5, the relationships between mangroves and the surrounding environmental gradient were utilized. The relationships were incorporated into the mapping process via a typical Bayesian probability model. The Bayesian model functioned as a post-classifier to improve the quality of a mangrove map already produced. The environmental gradient used was a GIS layer of soil pH data.

The integration of soil pH into the mapping process turned out to be worthwhile as it significantly increased the mapping accuracy: from 76.0% to 88.2%. However, the remaining confusion between *R. mucronata* and *S. caseolaris* points to the fact that soil pH data cannot help to resolve the similarity between the two species, and, as a result, more ancillary data such as leaf texture (i.e., captured by aerial photos) are recommended. Overall, it is anticipated that the methodology presented in this study will be used as a guideline for producing a mangrove map at the community or species level.

Lastly, follow-up research is already underway. First, the performances of other inference engines, such as artificial neural networks and the Dempster-Shafer theory, are now being compared with the outcome of Bayes' rule used in this thesis. Second, despite the problem relating to interoperability (i.e., data incompatibility) (Bishr, 1998), the research question of how to draw a consensus from expert knowledge from different spatial and non-spatial data sources (e.g., mangrove scientific publications, empirical data from other study areas, and opinions from local mangrove experts) is being resolved using recent advances in geo-information theories, including (i) suitability modelling (Bonham-Carter, 1994; Yamada et al., 2003), (ii) the personal construct theory (Kelly, 1955; Zhu, 1999), and (iii) the semantics look-up table method (Comber et al., 2004).

6.3 This thesis in a nutshell!

This thesis is all about exploring remote sensing methods that can be used for mapping mangroves at the species level. The significance of this thesis can be synthesised into three major points. First, I point out the reason why there is the need for the continuation of my research on the use of hyperspectral sensors for mapping mangroves at the species level.

Second, I tell the reader that there is no real reason to feel panic with the technical complications when working with hyperspectral data. Third, I explain why my attempt to incorporate the mangrove-environment relationships into the mapping process (i.e., an idea that seems to be too expensive to implement) could lead to an operational method in the near future.

6.3.1 Why is the follow-on research needed?

The evolution of remote sensing sensors from multispectral sensors to hyperspectral sensors gives birth to a practical tool for detailed mangrove studies. The hyperspectral sensor does not only share advantageous characteristics of its ancestor, multispectral sensor, particularly on the aspect of cost-effectiveness. It allows us to exploit the relationships between mangroves and their spectral characteristics in finer detail. This successful exploitation is evident in the following examples. Hirano et al. (2003) used the 224-band Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) sensor for producing accurately the map of the mangrove communities of the Everglades, Florida. Furthermore, Demuro and Chisholm (2003) successfully used the HYPERION hyperspectral sensor, accommodated on the satellite platform, for discriminating 8-class mangrove species communities in Australia. The situation looked even more optimistic when we discovered that pure mangrove spectra (laboratory spectra) contained enough information for discriminating most of mangrove species (Chapter 2). Nevertheless, the hyperspectral research on mangroves published to date (Green et al., 2000; Demuro and Chisholm, 2003; Held et al., 2003; Hirano et al., 2003) is still inconclusive. As a result, more research is still needed as to see whether the mangroves can be mapped at the species level when airborne or satellite hyperspectral sensors are used under the field conditions where there are numerous factors that could degrade the spectral signal received by the hyperspectral sensor, thereby making it harder to separate mangrove species.

6.3.2 Be at ease with hyperspectral data

With respect to spatial resolution of hyperspectral sensors, studying mangroves in detail does not require expensive high spatial resolution data as one might think. According to the spatial sensitivity analysis of tropical mangrove distribution reported by Manson et al. (2003) and our own experience in the tropical mangrove forests, it is clear that mangrove forests possess low spatial heterogeneity of mangrove species

distribution and, therefore, spatial resolution of commercial hyperspectral sensors installed on the satellite platform such as HYPERION (i.e., 30 m spatial resolution) should be adequate for the study of mangroves at the species level. This means that mapping mangrove species does not require expensive airborne hyperspectral sensors as in case studies of other terrestrial plants at the species level (Schmidt and Skidmore, 2003; Clark et al., 2005).

In addition, using hyperspectral sensors for mapping mangroves at the species level does not necessarily require more complex data treatments than the case of typical multispectral sensors. In other words, one can still use those existing methods (e.g., statistical-based classifiers etc.) that are normally used for the case of multispectral analyses for analysing hyperspectral data except for the requirements of special treatments for (i) high dimensionality, and (ii) high noise levels.

First, high dimensionality of hyperspectral technology is a two-sided sword. On the one hand, the inter-band correlations provide useful information about the shape of the spectral distribution in the feature space. This shape information has been proved that it helps increase the mapping accuracy (Landgrebe, 1997). This author falsified the old notion that the inter-band correlations are not good for classification (Ramsey III and Jensen, 1996). On the other hand, when there is limited number of field samples, using too many spectral bands (e.g., > 20 bands) at the same time could reduce the precision of the mathematical model of class information in the feature space (i.e., the curse of dimensionality (Bellman, 1961; Hughes, 1968)). To our relief, this problem of high-dimensionality can be solved straightforwardly by the use of feature extraction/selection algorithms (Lee and Landgrebe, 1993; Du and Chang, 2001; Kavzoglu and Mather, 2002). In the light of the existing tools, we have proposed an innovative form of genetic search algorithms for reducing the number of bands and, at the same time, maintaining mangrove species separability (see Chapter 3).

Second, it is well-known that narrow-band sensors of hyperspectral instruments can capture a very low amount of energy, thereby resulting in poor signal quality (i.e., noisy signals). This problem could get worse when there are additional external disturbances such as the fluctuation of the atmospheric states (Oppenheim and Schafer, 1975; Landgrebe, 1997; Lyon, 2004). Moreover, the connection points between spectral detectors

of the hyperspectral instrument could also play an important role in the quality of the spectral signal recorded (Schmidt and Skidmore, 2004). We have raised awareness of this issue in Chapter 4 and discuss about the spectral convolutions, which are popularly used for solving this signal-noise problem. In addition, we have proposed a method that can be used to visualise the trade-off between the noise levels reduced and the statistical estimate of the original data disturbed by the spectral convolution.

6.3.3 Is exploiting non-spectral information promising?

This thesis supports the idea of incorporating ancillary ecological data into the mapping process. This concept of integrating extra information into the mapping process has been borrowed from successful case studies of mapping other plant species (Skidmore et al., 1997a, 1997b; Lehmann and Lenz, 1998; Berberoglu et al., 2004; Comber et al., 2004; Schmidt et al., 2004). In short, similar to the extra spectral bands (or layers) provided by the hyperspectral sensor, the extra GIS layer produced by exploiting the relationships between mangroves and the environmental gradients can be thought as if it is an extra non-spectral dimension (i.e., ecological dimension). The outcome of this thesis in Chapter 5 points out that the integration of ecological data (soil pH) into the mapping process is worth the extra fieldwork effort. It significantly increases the mapping accuracy of the final mangrove map of Cape Talumpuk from 76% to 88%. More importantly, soil pH is a cost-effective parameter. It is easy to analyse (i.e., using a pH probe), and, in some countries such as Thailand[*], soil-related parameters such as soil pH are often available for the research as they are collected regularly from the mangrove forests to monitor their conditions. In addition to the success of adding soil pH data into the mapping process, I have a plan to test other ecological gradients that can be gathered cost-effectively (e.g., leaf textures captured by aerial photos, LIDAR-derived elevation maps, and inundation frequency maps produced by incorporating elevation maps with automatic tidal records) for improving the mapping accuracy further (e.g., > 90% of accuracy). If this plan is successful, it will strengthen the possibility of exploiting non-spectral information for mangrove species mapping at the operational level.

[*] Forest Research Office, Royal Forest Department: www.forest.go.th/Research/English/index.htm

6.4 Conclusion

As the potential of hyperspectral and ecological data for detailed mangrove mapping has already been unveiled, the main goal of this study has been achieved. The achievement of this thesis can be summarized as follows:

(1) The thesis reports that hyperspectral data contain adequate spectral details for discriminating most mangrove species. Further studies using airborne and satellite hyperspectral sensors are therefore encouraged.
(2) A form of genetic search algorithms has been successfully tested on a species-level problem of very high dimensionality. The results point to the capability of the genetic search algorithm to help in solving the problem of high dimensionality.
(3) The mistreatment of hyperspectral smoothing has been investigated, and an alternative method of optimizing the smoothed result is proposed.
(4) Mangrove-environment relationships have been successfully exploited for mangrove mapping, using a Bayesian expert system. It has been found that the relationships help to increase the accuracy of the final mangrove map at the species level.

References

Ahern, C.R., McElnea, A.E., Sullivan L.A., 2004. Acid Sulfate Soils Laboratory Methods Guidelines. Queensland Department of Natural Resources, Mines and Energy, Queensland.

Anderson, J.R., Hardy, E.E., Roach, J.T., Witmer, R.E., 1976. A land use and land cover classification system for use with remote sensor data. U.S. Geological Survey Professional Paper 964, 41.

Anderson, T.W., 1984. An Introduction to Multivariate Statistical Analysis. Wiley, New York.

Andréfouët, S., Hochberg, E.J., Payri, C., Atkinson, M.J., Muller-Karger, F.E., Ripley, H., 2003. Multi-scale remote sensing of microbial mats in an atoll environment. International Journal of Remote Sensing 24, 2661-2682.

Ardahanlioglu, O., Oztas, T., Evren, S., Yilmaz, H., Yildirim, Z.N., 2003. Spatial variability of exchangeable sodium, electrical conductivity, soil pH and boron content in salt- and sodium-affected areas of the Igdir plain (Turkey). Journal of Arid Environments 54, 495–503.

Aschbacher, J., Tiangco, P., Giri, C.P., Ofren, R.S., Paudyal, D.R., Ang, Y.K., 1995. Comparison of different sensors and analysis techniques for tropical mangrove forest mapping. Proceedings of the International Conference IGARSS: 2109-2111.

Asner, G.P., Wessman, C.A., Bateson, C.A., Privette, J.L., 2000. Impact of tissue, canopy, and landscape factors on the hyperspectral reflectance variability of arid ecosystems. Remote Sensing of Environment 74, 69-84.

Bandyopadhyay, S., 2005. Satellite image classification using genetically guided fuzzy clustering with spatial information. International Journal of Remote Sensing 26, 579-593.

Barbier, E.B., Sathiratai, S., 2004. Shrimp farming and mangrove loss in Thailand. Edward Elgar, Cheltenham.

Basak, U.C., Das, A.B., Das, P., 1996. Chlorophyll, carotenoids, proteins and secondary metabolites in leaves of 14 species of mangroves. Bulletin of Marine Science 58, 645–659.

Bellman, R.E., 1961. Adaptive Control Processes, Princeton University Press, Princeton.

Ben-Dor, E., Patkin, K., Banin, A., Karnieli, A., 2002. Mapping of several soil properties using DAIS-7915 hyperspectral scanner data - A case study over soils in Israel. International Journal of Remote Sensing 23, 1043-1062.

Berberoglu, S., Yilmaz, K.T., Özkan, C., 2004. Mapping and monitoring of coastal wetlands of Çukurova Delta in the Eastern Mediterranean region. Biodiversity and Conservation 13, 615-633.

Berlanga-Robles, C.A., Ruiz-Luna, A., 2002. Land use mapping and change detection in the coastal zone of northwest Mexico using remote sensing techniques. Journal of Coastal Research 18, 514-522.

Bishr, Y., 1998. Overcoming the semantics and other barriers to GIS interoperability. International Journal of Geographical Information Science 12, 299-314.

Blasco, F., Gauquelin, T., Rasolofoharinoro, M., Denis, J., Aizpuru, M., Caldairou, V., 1998. Recent advances in mangrove studies using remote sensing data. Marine and Freshwater Research 49, 287-296.

Bonham-Carter, G., 1994. Geographic Information Systems for Geoscientists, Pergamon, New York.

Bruzzone, L., Serpico, S.B., 2000. A technique for feature selection in multiclass cases. International Journal of Remote Sensing 21, 549-563.

Burgess, T.M., Webster, R., 1980. Optimal interpolation and isorithmic mapping of soil properties. I. The semivariogram and punctual kriging. Journal of Soil Science 31, 315–331.

Bruce, L.M., Li, J., 2001. Wavelets for computationally efficient hyperspectral derivative analysis. IEEE Transactions on Geoscience and Remote Sensing 39, 1540-1546.

Castro-Esau, K.L., Sánchez-Azofeifa, G.A., Caelli, T., 2004. Discrimination of lianas and trees with leaf-level hyperspectral data. Remote Sensing of Environment 90, 353-372.

Chalermwat, P., El-Ghazawi, T., LeMoigne, J., 2001. 2-phase GA-based image registration on parallel clusters. Future Generation Computer Systems 17, 467-476.

Chen, L., 2003. A study of applying genetic programming to reservoir trophic state evaluation using remote sensor data. International Journal of Remote Sensing 24, 2265-2275.

Chen, W., Kuze, H., Uchiyama, A., Suzuki, Y., Takeuchi, N., 2001. One-year observation of urban mixed layer characteristics at Tsukuba, Japan using a micro pulse lidar. Atmospheric Environment 35, 4273-4280.

Clark, M.L., Roberts, D.A., David, B.C., 2005. Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales. Remote Sensing of Environment 96, 375-398.

Clough, B.F., 1982. Mangrove Ecosystems in Australia: Structure, Function and Management. Australia Institute of Marine Science, ANU press, Sydney.

Cochrane, M.A., 2000. Using vegetation reflectance variability for species level classification of hyperspectral data. International Journal of Remote ***Sensing*** 21, 2075-2087.

Cogdill, R.P., Hurburgh Jr., C.R., Rippke, G.R., 2004. Single-kernel maize analysis by near-infrared hyperspectral imaging. Transactions of the American Society of Agricultural Engineers 47, 311-320.

Comber, A.J., Law, A.N.R., Lishman, J.R., 2004. Application of knowledge for automated land cover change monitoring. International Journal of Remote Sensing 25, 3177-3192.

Congalton, R.G., Green K., 1999. Assessing the accuracy of remotely sensed data: principles and practices. Lewis, Boca Raton.

Curran, P.J., 1989. Remote sensing of foliar chemistry. Remote Sensing of Environment 30, 271-278.

Curran, P.J., Dungan, J.L., Macler, B.A., Plummer, S.E., Peterson, D.L., 1992. Reflectance spectroscopy of fresh whole leaves for the estimation of chemical concentration. Remote Sensing of Environment 39, 153-166.

Curran, P.J., Dungan, J.L., Peterson, D.L., 2001. Estimating the foliar biochemical concentration of leaves with reflectance spectrometry: Testing the Kokaly and Clark methodologies. Remote Sensing of Environment 76, 349-359.

Das, A.B., Parida, A., Basak, U.C., Das, P., 2002. Studies on pigments, proteins and photosynthetic rates in some mangroves and mangrove associates from Bhitarkanika, Orissa. Marine Biology 141, 415-422.

Demuro, M., Chisholm, L., 2003. Assessment of Hyperion for characterizing mangrove communities. Proceedings of the International Conference the AVIRIS 2003 workshop: 18-23.

Deng, L., Shen, X., 1997. Maximum likelihood in statistical estimation of dynamic systems: Decomposition algorithm and simulation results. Signal Processing 57, 65-79.

Du, Q., Chang, C., 2001. A linear constrained distance-based discriminant analysis for hyperspectral image classification. Pattern Recognition 34, 361-373.

Duda, R., Hart. P., 1973. Pattern Classification and Scene Analysis. Wiley, New York.

Elvidge, C.D., 1987. Reflectance characteristics of dry plant materials. Proceedings of the International Conference Remote Sensing of Environment: 721-733.

Elvidge, C.D., 1990. Visible and near infrared reflectance characteristics of dry plant materials. International Journal of Remote Sensing 11, 1775-1795.

Everitt, J.H., Judd, F.W., Escobar, D.E., Davis, M.R., 1996. Integration of remote sensing and spatial information technologies for mapping black mangrove on the Texas gulf coast. Journal of Coastal Research 12, 64-69.

Fang, H., Liang, S., Kuusk, A., 2003. Retrieving leaf area index using a genetic algorithm with a canopy radiative transfer model. Remote Sensing of Environment 85, 257-270.

Feise, R.J., 2002. Do multiple outcome measures require p-value adjustment? BMC Medical Research Methodology 2, 8.

Flink P., Lindell, L.T., Ostlund, C., 2001. Statistical analysis of hyperspectral data from two Swedish lakes. Science of the Total Environment 268, 155-169.

Foody, G.M., Sargent, I.M.J., Atkinson, P.M., Williams, J.W., 2004. Thematic labelling from hyperspectral remotely sensed imagery: Trade-offs in image properties. International Journal of Remote Sensing 25, 2337-2363.

Fukunaga, K., 1990. Introduction to Statistical Pattern Recognition, Academic Press, Orlando.

Gajem, Y.M., Warrick, A.W., Myers, D.E., 1981. Spatial structure of physical properties of a Typic Torrifluvent soil. Soil Science Society America Journal 45, 709–715.

Gao, J., 1999. A comparative study on spatial and spectral resolutions of satellite data in mapping mangrove forests. International Journal of Remote Sensing 20, 2823-2833.

Gates, D.M., Keegan, H.J., Schleter, J.C., Weidner, V.R., 1965. Spectral properties of plants. Applied Optics 9, 545-552.

Goel, P.K., Prasher, S.O., Landry, J.A., Patel, R.M., Viau, A.A., Miller, J.R., 2003. Estimation of crop biophysical parameters through airborne and field hyperspectral remote Sensing. Transactions of the American Society of Agricultural Engineers 46, 1235-1246.

Goldberg, D., 1989. Genetic Algorithms in Search, Optimization and Machine Learning, Addison-Wesley, Reading.

Gong P., Pu, R., Heald, R.C., 2002. Analysis of in situ hyperspectral data for nutrient estimation of giant sequoia. International Journal of Remote Sensing 23, 1827-1850.

Gong, P., Pu, R., Yu, B., 2001. Conifer species recognition: Effects of data transformation. International Journal of Remote Sensing 22, 3471-3481.

Green, E.P., Clark, C.D., Edwards, A.J., 2000. Image classification and habitat mapping. In: Edwards, A.J. (Ed.), Remote Sensing Handbook for Tropical Coastal Management. UNESCO, Paris, pp. 141-154.

Green, R.O., Eastwood, M.L., Sarture, C.M., Chrien, T.G., Aronsson, M., Chippendale, B.J., Faust, J.A., Pavri, B.E., Chovit, C.J., Solis, M., Olah, M.R., Williams, O., 1998. Imaging spectroscopy and the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS). Remote Sensing of Environment 65, 227-248.

Harsanyi, J. C., Chang, C.I., 1994. Hyperspectral image classification and dimensionality reduction: an orthogonal subspace projection approach. IEEE Transactions on Geoscience and Remote Sensing 32, 779-785.

Harvey, N.R., Theiler, J., Brumby, S.P., Perkins, S., Szymanski, J.J., Bloch, J.J., Porter, R.B., Galassi, M., Young, A.C., 2002. Comparison of GENIE and conventional supervised classifiers for multispectral image feature extraction. IEEE Transactions on Geoscience and Remote Sensing 40, 393-404.

Held, A., Ticehurst, C., Lymburner, L., Williams, N., 2003. High resolution mapping of tropical mangrove ecosystems using hyperspectral

and radar remote sensing. International Journal of Remote Sensing 24, 2739-2759.

Himmelsbach, D.S., Boer, J.D., Akin, D.E., Barton, E.E., 1988. Solid-state carbon-13 NMR, FTIR, and NIR spectroscopic studies of ruminant silage digestion. In: Creaser, C.S., Davies, A.M.C. (Eds.), Analytical Applications of Spectroscopy. Royal Society of Chemistry, London, pp. 410-413.

Hirano, A., Madden, M., Welch, R., 2003. Hyperspectral image data for mapping wetland vegetation. Wetlands 23, 436-448.

Hochberg, E.J., Atkinson, M.J., 2003. Capabilities of remote sensors to classify coral, algae, and sand as pure and mixed spectra. Remote Sensing of Environment 85, 174-189.

Hoffer, R.M., 1978. Biological and physical considerations in applying computer aided analysis techniques to the remote sensor data. In: Swain, P.H., Davis, S.M. (Eds.), Remote Sensing: the Quantitative Approach. McGraw-Hill, New York, pp. 227-289.

Hogarth, P.J., 1999. The Biology of Mangroves. Oxford University Press, Oxford.

Holland, J.H., 1975. Adaptation in Natural and Artificial Systems, University of Michigan, Ann Arbor.

Hsu, J.C., 1996. Multiple comparisons: theory and methods. Chapman and Hall, London.

Hughes, G.F., 1968. On the mean accuracy of statistical pattern recognizers. IEEE Transactions on Information Theory 14, 55-63.

Imanishi, J., Sugimoto, K., Morimoto, Y., 2004. Detecting drought status and LAI of two Quercus species canopies using derivative spectra. Computers and Electronics in Agriculture 43, 109-129.

Isaaks, E.H., Srivastava, R.M., 1989. Applied Geostatistics. Oxford University Press, Oxford.

Jin, Y.Q., Wang, Y.A., 2001. Genetic algorithm to simultaneously retrieve land surface roughness and soil wetness. International Journal of Remote Sensing 22, 3093-3099.

John, G.H., Kohavi, R., Pfleger, K., 1994. Irrelevant features and the subset selection problem. In: Cohen, W.W., Hirsh, H. (Eds.), Machine Learning. Morgan Kaufmann, San Francisco, pp. 121-129.

Jones, G.A., Greenhill, D., Orwell, J., Forte, P., 2000. Coastline registration: Efficient optimization in large dimensions using genetic algorithms. Geographical and Environmental Modelling 4, 21-41.

Kailath, T., 1967. The divergence and the Bhattacharyya distance measures in signal selection. IEEE Transactions on Communication Technology 15, 52-60.

Kavzoglu, T., Mather, P.M., 2002. The role of feature selection in artificial neural network applications. International Journal of Remote Sensing 23, 2919-2937.

Kawata, S., Minami, S., 1984. Adaptive smoothing of spectroscopic data by linear mean-square estimation. Applied Spectroscopy 38, 49-58.

Kay, S.M., 1993. Fundamentals of statistical signal processing : estimation theory, Prentice-Hall, New Jersey.

Kelly, G.A., 1955. The Psychology of Personal Constructs, Norton, New York.

Kendall, M.G., 1961. A Course in the Geometry of n Dimensions, Dover Publications, Inc., New York.

Keshava, N., 2004. Distance matrices and band selection in hyperspectral processing with applications to material identification and spectral libraries. IEEE Transaction on Geoscience and Remote Sensing 42, 1552-1565.

Kohavi, R., John, G.H., 1997. Wrappers for feature subset selection. Artificial Intelligence 97, 273-324.

Kokaly, R.F., 2001. Investigating a physical basis for spectroscopic estimates of leaf nitrogen concentration. Remote Sensing of Environment 75, 267-287.

Kooistra, L., Wanders, J., Epema, G.F., Leuven, R.S.E.W., Wehrens, R., Buydens, L.M.C., 2003. The potential of field spectroscopy for the assessment of sediment properties in river floodplains. Analytica Chimica Acta 484, 189-200.

Kovacs, J.M., Wang, J., Flores-Verdugo, F., 2005. Mapping mangrove leaf area index at the species level using IKONOS and LAI-2000 sensors for the Agua Brava Lagoon, Mexican Pacific. Estuarine, Coastal and Shelf Science 62, 377-384.

Kruse, F.A., Boardman, J.W., Lefkoff, A.B., Young, J.M., Kierein-Young, K.S., Cocks, T.D., Jenssen, R., Cocks, P.A., 2000. HyMap: an Australian hyperspectral sensor solving global problems – results from USA HyMap data acquisitions. Proceedings of the International Conference Australasian Remote Sensing and Photogrammetry: 18-23.

Kruse, F.A., Lefkoff, A.B., Boardman, J.W., Heidebrecht, K.B., Shapiro, A.T., Barloon, P.J., Goetz, A.F.H., 1993. The spectral image processing system (SIPS)--interactive visualization and analysis of imaging spectrometer data. Remote Sensing of Environment 44, 145-163.

Kumar, L., Schmidt, K.S., Dury, S., Skidmore, A.K. 2001. Review of hyperspectral remote sensing and vegetation science. In: van der Meer, F.D., de Jong, S.M. (Eds.), Imaging Spectrometry: Basic Principles and Prospective Applications. Kluwer Academic Press, Dordrecht, pp. 111-155.

Landgrebe, D., 1997. On information extraction principles for hyperspectral data.
http://dynamo.ecn.purdue.edu/~landgreb/whitepaper.pdf.

Le Maire, G., François, C., Dufrêne, E., 2004. Towards universal broad leaf chlorophyll indices using PROSPECT simulated database and hyperspectral reflectance measurements. Remote Sensing of Environment 89, 1-28.

Lee, C., Landgrebe, D.A., 1993. Feature extraction based on decision boundaries. IEEE Transactions on Geoscience and Remote Sensing 15, 388-400.

Lehmann, D., Lenz, R., 1998. Synthetical vegetation mapping within the perimeter of the Berchtesgaden National park, called Vorfeld. A methodical study using Geographical Information Systems and Neural Networks. Verhandlungen der Gesellschaft fur Okologie 29, 89-96.

Lillesand, T.M., Kiefer, R.W., 2000. Remote sensing and image interpretation. Wiley, Chichester.

Linneweber, V., de Lacerda, L.D., 2002. Mangrove Ecosystems: Function and Management. Springer, Berlin.

Liu, Z., Liu, A., Wang, C., Niu, Z., 2004. Evolving neural network using real coded genetic algorithm (GA) for multispectral image classification, Future Generation Computer Systems 20, 1119-1129.

Lofy, B., Sklansky, J., 2001. Segmenting multisensor aerial images in class-scale space. Pattern Recognition 34, 1825-1839.

Lu, F., Eriksson, L.O., 2000. Formation of harvest units with genetic algorithms. Forest Ecology and Management 130, 57-67.

Lugo, A.E., Snedaker, S.C., 1974. The ecology of mangroves. Annual Review of Ecological Systems 5, 39–64.

Luo, J.C., Zheng, J., Leung, Y., Zhou, C.H., 2003. A knowledge-integrated stepwise optimization model for feature mining in remotely sensed images. International Journal of Remote Sensing 24, 4661-4680.

Lyon, R.G., 2004. Understanding Digital Signal Processing, Prentice-Hall, New Jersey.

Macnae, W., 1968. A general account of the fauna and flora of mangrove swamps and forests in the Indo-West-Pacific region. Advances in Marine Biology 6, 73-270.

Madden, H.M., 1978. Comments on the Savitzky-Golay convolution method for least-Squares fit smoothing and differentiation of digital data. Analytical Chemistry 50, 1383.

Manson, F.J., Loneragan, N.R., Phinn, S.R., 2003. Spatial and temporal variation in distribution of mangroves in Moreton Bay, subtropical Australia: a comparison of pattern metrics and change detection analyses based on aerial photographs. Estuarine, Coastal and Shelf Science 57, 653-666.

Mausel, P.W., Kramber, W. J., Lee, J. K., 1990. Optimum band selection for supervised classification of multispectral data. Photogrammetric Engineering and Remote Sensing 56, 55–60.

McDonald, M.S., 2003. Photobiology of higher plants. Wiley, Chichester.

McLean, E.O., 1982. Soil pH and lime requirement. In: editors (eds.), Methods of Soil Analysis. Part II. Chemical and Microbiological Properties. ASA-SSSA, Madison, WI, pp.199–224.

Menon, G.G., Neelakantan, B., 1992. Chlorophyll and light attenuation from the leaves of mangrove species of Kali estuary. Indian Journal of Marine Sciences 21, 13–16.

Meroni, M., Colombo, R., Panigada, C., 2004. Inversion of a radiative transfer model with hyperspectral observations for LAI mapping in poplar plantations. Remote Sensing of Environment 92, 195-206.

Mertens, K.C., Verbeke, L.P.C., Ducheyne, E.I., de Wulf, R.R., Using genetic algorithms in sub-pixel mapping. International Journal of Remote Sensing 24, 4241-4247.

Metternicht, G.I., Zinck, J.A., 2003. Remote sensing of soil salinity: potentials and constraints. Remote Sensing of Environment 85, 1-20.

Mutanga, O., Skidmore, A.K., 2004. Integrating imaging spectroscopy and neural networks to map grass quality in the Kruger National Park, South Africa. Remote Sensing of Environment 90, 104-115.

Mutanga O., Skidmore, A.K., van Wieren, S., 2003. Discriminating tropical grass (*Cenchrus ciliaris*) canopies grown under different nitrogen treatments using spectroradiometry. ISPRS Journal of Photogrammetry and Remote Sensing 57, 263-272.

Narendra, P.M., Fukunaga, K., 1977. A branch and bound algorithm for feature subset selection. IEEE Transactions on Computers 26, 917–922.

Oppenheim, A.V., Schafer, R.W., 1975. Digital Signal Processing, Prentice-Hall, New Jersey.

Pal, S.K., Bandyopadhyay, S., Murthy, C.A., 2001. Genetic classifiers for remotely sensed images: Comparison with standard methods. International Journal of Remote Sensing 22, 2545-2569.

Pekkarinen, A., 2002. Image segment-based spectral features in the estimation of timber volume. Remote Sensing of Environment 82, 349-359.

Perneger, T.V., 1998. What's wrong with Bonferroni adjustments? British Medical Journal 316, 1236-1238.

Peterson, D.L., Hubbard, G.S., 1992. Scientific issues and potential remote sensing requirements for plant biochemical content. Journal of Image Science and Technology 36, 446-456.

Pudil, P., Novovicova, J., Kittler, J., 1994. Floating search methods in feature selection. Pattern Recognition Letters 15, 1119–1125.

Ramsar Convention, 1971. www.ramsar.org.

Ramsey III, E.W., Jensen, J.R., 1996. Remote sensing of mangrove wetlands: relating canopy spectra to site-specific data. Photogrammetric Engineering and Remote Sensing 62, 939-948.

Ramsey III, E.W., Rangoonwala, A., Nelson, G., Ehrlich, R., Martella, K., 2005. Generation and validation of characteristic spectra from EO1 Hyperion image data for detecting the occurrence of the invasive species, Chinese tallow. International Journal of Remote Sensing 26, 1611-1636.

Rasolofoharinoro, M., Blasco, F., Bellan, M.F., Aizpuru, M., Gauquelin, T., Denis, J., 1998. A remote sensing based methodology of mangrove studies in Madagascar. International Journal of Remote Sensing 19, 1873-1886.

Rees, W.G., Tutubalina, O.V., Golubeva, E.I., 2004. Reflectance spectra of subarctic lichens between 400 and 2400 nm. Remote Sensing of Environment 90, 281-292.

Richards, J.A., 1986. Remote Sensing Digital Image Analysis: An Introduction. Springer-Verlag, Berlin.

Richards, J.A., 1993. Remote Sensing Digital Image Analysis: An Introduction. Springer-Verlag, Berlin.

Rothman, K.J., 1990. No adjustments are needed for multiple comparisons. Epidemiology 1, 43-46.

Samra, J.S., Sharma, K.N.S., Tyaki, N.K., 1988. Analysis of spatial variability in sodic soils. I. Structural analysis. Soil Science 145, 180–187.

Satyanarayana B., Raman, A.V., Dehairs, F., Kalavati, C., Chandramohan, P., 2002. Mangrove floristic and zonation patterns of Coringa, Kakinada Bay, East Coast of India. Wetlands Ecology and Management 10, 25-39.

Savitzky, A., Golay, M.J.E., 1964. Smoothing and differentiation of data by simplified least squares procedures. Analytical Chemistry 36, 1627-1639.

Schmidt, K.S., Skidmore, A.K., 2003. Spectral discrimination of vegetation types in a coastal wetland. Remote Sensing of Environment 85, 92-108.

Schmidt, K.S., Skidmore, A.K., 2004. Smoothing vegetation spectra with wavelets. International Journal of Remote Sensing 25, 1167-1184.

Schmidt, K.S., Skidmore, A.K., Kloosterman, E.H., van Oosten, H., Kumar, L., Janssen, J.A.M., 2004. Mapping coastal vegetation using an

expert system and hyperspectral imagery. Photogrammetric Engineering and Remote Sensing 70, 703-715.

Schuerger, A.C., Capelle, G.A, Di Benedetto, J.A., Mao, C., Thai, C.N., Evans, M.D., Richards, J.T., Blank, T.A., Stryjewski, E.C., 2003. Comparison of two hyperspectral imaging and two laser-induced fluorescence instruments for the detection of zinc stress and chlorophyll concentration in bahia grass (*Paspalum notatum Flugge*.). Remote Sensing of Environment 84, 572-588.

Semeniuk, V., 1983. Mangrove distribution in northwestern Australia in relationship to regional and local freshwater seepage. Vegetation 53, 11-31.

Shahshahani, B.M., Landgrebe, D.A., 1994. The effect of unlabeled samples in reducing the small sample size problem and mitigating the Hughes phenomenon. IEEE Transactions on Geoscience and Remote Sensing 32, 1087-1095.

Siedlecki, W., Sklansky, J., 1989. A note on genetic algorithms for large-scale feature selection. Pattern Recognition Letters 10, 335-347.

Skidmore, A.K., Schmidt, K., Kloosterman, H., Kumar, L., van Oosten, H., 2001. Hyperspectral imagery for coastal wetland vegetation mapping. Beleids Commissie Remote Sensing (BCRS) NRSP-2 report.

Skidmore, A.K., Turner, B.J., Brinkhof, W., Knowles, E., 1997a. Performance of a neural network: mapping forests using GIS and remotely sensed data. Photogrammetric Engineering and Remote Sensing 63, 501-514.

Skidmore, A.K., Varekamp, C., Wilson, L., Knowles, E., Delaney, J., 1997b. Remote sensing of soils in a eucalypt forest environment. International Journal of Remote Sensing 18, 39-56.

Smith, K.L., Steven, M.D., Colls, J.J., 2004. Use of hyperspectral derivative ratios in the red-edge region to identify plant stress responses to gas leaks. Remote Sensing of Environment 92, 207-217.

Soukupová, J., Rock, B.N., Albrechtová, J., 2002. Spectral characteristics of lignin and soluble phenolics in the near infrared - A comparative study. International Journal of Remote Sensing 23, 3039-3055.

Spalding, M.D., Blasco, F., Field, C.D., 1997. World Mangrove Atlas. International Society for Mangrove Ecosystems, Okinawa.

Steinier, J., Termonia, Y., Deltour, J., 1972. Comments on smoothing and differentiation of data by simplified least square procedure. Analytical Chemistry 44, 1906.

Strachan, I.B., Pattey, E., Boisvert, J.B., 2002. Impact of nitrogen and environmental conditions on corn as detected by hyperspectral reflectance. Remote Sensing of Environment 80, 213-224.

Sulong, I., Mohd-Lokman, H., Mohd-Tarmizi, K., Ismail, A., 2002. Mangrove mapping using Landsat imagery and aerial photographs: Kemaman District, Terengganu. Malaysia Environment, Development and Sustainability 4, 135-152.

Swain, P.H., Davis, S.M., 1978. Remote sensing—The quantitative approach. McGraw-Hill, New York.

Swain, P.H., King, R.C., 1973. Two effective feature selection criteria for multispectral remote sensing. Proceedings of the First International Conference on Pattern Recognition: 536–540.

Teeratanatorn, W., 2000. Mangroves of Pak Phanang Bay (in Thai). Royal Forest Department, Bangkok.

Thenkabail, P.S., Enclona, E.A., Ashton, M.S., van Der Meer, B., 2004. Accuracy assessments of hyperspectral waveband performance for vegetation analysis applications. Remote Sensing of Environment 91, 354-376.

Thomas, V., Treitz, P., Jelinski, D., Miller, J., Lafleur, P., McCaughey, J.H., 2003. Image classification of a northern peatland complex using spectral and plant community data. Remote Sensing of Environment 84, 83-99.

Tomlinson, P.B., 1994. The botany of mangroves. Cambridge University Press, Cambridge.

Tsai, F., Philpot, W., 1998. Derivative analysis of hyperspectral data. Remote Sensing of Environment 66, 41-51.

Tsai, F., Philpot, W., 2002. A derivative-aided hyperspectral image analysis system for land-cover classification. IEEE transactions on Geoscience and Remote Sensing 40, 416-425.

Tseng, D.C., Lai, C.C., 1999. A genetic algorithm for MRF-based segmentation of multi-spectral textured images. Pattern Recognition Letters 20, 1499-1510.

Ulfarsson, M.O., Benediktsson, J.A., Sveinsson, J.R., 2003. Data fusion and feature extraction in the wavelet domain. International Journal of Remote Sensing 24, 3933-3945.

Vaiphasa, C., 2003. Innovative genetic algorithm for hyperspectral image classification. Proceedings of the International Conference MAP ASIA: 20.

Vaiphasa, C., Ongsomwang, S., 2004. Hyperspectral Data for Tropical Mangrove Species Discrimination. Proceedings of the 25th ACRS Conference: 22-28.

van Niel, T.G., McVicar, T.R., Fang, H., Liang, S., 2003. Calculating environmental moisture for per-field discrimination of rice crops. International Journal of Remote Sensing 24, 885-890.

Vaughan, R.G., Calvin, W.M., Taranik, J.V., 2003. SEBASS hyperspectral thermal infrared data: Surface emissivity measurement and mineral mapping. Remote Sensing of Environment 85, 48-63.

Verheyden, A., Dahdouh-Guebas, F., Thomaes, K., de Genst, W., Hettiarachchi, S., Koedam, N., 2002. High-resolution vegetation data for mangrove research as obtained from aerial photography. Environment, Development and Sustainability 4, 113-133.

Vilarrubia, T.V., 2000. Zonation pattern of an isolated mangrove community at Playa Medina, Venezuela. Wetlands Ecology and Management 8, 9-17.

Whiting, M.L., Li, L., Ustin, S.L., 2004. Predicting water content using Gaussian model on soil spectra. Remote Sensing of Environment 89, 535-552.

Williams, P., Norris, K., 2001. Near-infrared technology in the agricultural and food industries. American Association of Cereal Chemists Press, Minnesota.

Wu, D., Linders J., 2000. Comparison of three different methods to select feature for discriminating forest cover types using SAR imagery. International Journal of Remote Sensing 21, 2089-2099.

Wu, Y., Chen, J., Wu, X., Tian, Q., Ji, J., Qin, Z., 2005. Possibilities of reflectance spectroscopy for the assessment of contaminant elements in suburban soils. Applied Geochemistry 20, 1051-1059.

Yamada, K., Elith, J., McCarthy, M., Zerger, A., 2003. Eliciting and integrating expert knowledge for wildlife habitat modelling. Ecological Modelling 165, 251-264.

Yamano, H., Chen, J., Tamura, M., 2003. Hyperspectral identification of grassland vegetation in Xilinhot, Inner Mongolia, China. International Journal of Remote Sensing 24, 3171-3178.

Yates, S.R., Warrick, A.W., Matthias, A.D., Musil, S., 1988. Spatial variability of remotely sensed surface temperature at field scale. Soil Science Society America Journal 52, 40–45.

Yost, R.S., Uehara, G., Fox, R.L., 1982. Geostatistical analysis of soil chemical properties of large land areas: I. Semivariograms. Soil Science Society of America Journal 46, 1028–1032.

Yu, S., de Backer, S., Scheunders, P., 2002. Genetic feature selection combined with composite fuzzy nearest neighbor classifiers for hyperspectral satellite imagery. Pattern Recognition Letters 23, 183-190.

Zarco-Tejada P.J., Miller, J.R., Morales, A., Berjón, A., Agüera, J., 2004. Hyperspectral indices and model simulation for chlorophyll estimation in open-canopy tree crops. Remote Sensing of Environment 90, 463-476.

Zarco-Tejada, P.J., Miller, J.R, Noland, T.L., Mohammed, G.H., Sampson, P.H., 2001. Scaling-up and model inversion methods with narrowband optical indices for chlorophyll content estimation in closed forest canopies with hyperspectral data. IEEE transactions of Geoscience and Remote Sensing 39, 1491-1507.

Zarco-Tejada, P.J., Pushnik, J.C., Dobrowski, S., Ustin, S.L., 2003. Steady-state chlorophyll a fluorescence detection from canopy derivative reflectance and double-peak red-edge effects. Remote Sensing of Environment 84, 283-294.

Zhang, J., Rivard, B., Sanchez-Azofeifa, A., 2004. Derivative spectral unmixing of hyperspectral data applied to mixtures of lichen and rock. IEEE transactions on Geoscience and Remote Sensing 42, 1934-1940.

Zhu, A. 1999. A personal construct-based knowledge acquisition process for natural resource mapping. International Journal of Geographical Information Science 13, 119-141.

Curriculum vitae

Chaichoke Vaiphasa was born on 29th November 1975 in Bangkok, Thailand. He finished his bachelor's degree in survey engineering at the top of his class from the department of Survey Engineering, Faculty of Engineer, Chulalongkorn University, Bangkok, Thailand in 1997. During this time, he received the Siam Cement award for his academic performance. Next, he spent about two more years working for the university as a junior lecturer. After that, he started his master's degree project titled "Applying an artificial neural network to map salt marsh vegetation" at International Institute for Geo-Information Science and Earth Observation (ITC), Enschede, The Netherlands, and graduated in 2001. Right after his graduation, he was appointed as a PhD candidate at the ITC institute and Wageningen University. The title of his PhD thesis is "Remote sensing techniques for mangrove mapping." Throughout his time in The Netherlands, he was financially supported by the scholarship from the government of Thailand.

Author's bibliography

Vaiphasa, C. (2003). Innovative genetic algorithm for hyperspectral image classification. Proceedings of the International Conference MAP ASIA: 20.

Vaiphasa, C., Ongsomwang, S. (2004). Hyperspectral Data for Tropical Mangrove Species Discrimination. Proceedings of the 25th ACRS Conference: 22-28.

Vaiphasa, C., Ongsomwang, S., Vaiphasa, T., Skidmore, A.K. (2005). Tropical mangrove species discrimination using hyperspectral data: a laboratory study. Estuarine, Coastal, and Shelf Science 65, 371-379.

Vaiphasa, C. (in press). Consideration of smoothing techniques for hyperspectral remote sensing. ISPRS Journal of Photogrammetry and Remote Sensing.

Vaiphasa, C., Skidmore, A.K., de Boer, W.F. (accepted). A post-classifier for mangrove mapping using ecological data. ISPRS Journal of Photogrammetry and Remote Sensing.

Vaiphasa, C., de Boer, W.F., Skidmore, A.K., Panitchart, S., Vaiphasa, T., Bamrongrugsa, N. (in review). Impacts of shrimp pond waste materials on mangrove growth and mortality: a case study from Pak Phanang, Thailand. Hydrobiologia.

Vaiphasa, C., van Oosten, H., Skidmore, A.K., de Boer, W.F. (in review). A genetic algorithm for hyperspectral feature selection. Photogrammetric Engineering and Remote Sensing.

Vaiphasa, C. (in review), A hyperspectral band selector for plant species discrimination. Photogrammetric Engineering and Remote Sensing.

Appendix I: Genetic algorithms step by step

The details of the genetic search algorithm (GA) used in this thesis can be mapped into several steps in a computational flowchart (Figure A.1.1), and each step of the flowchart is described below.

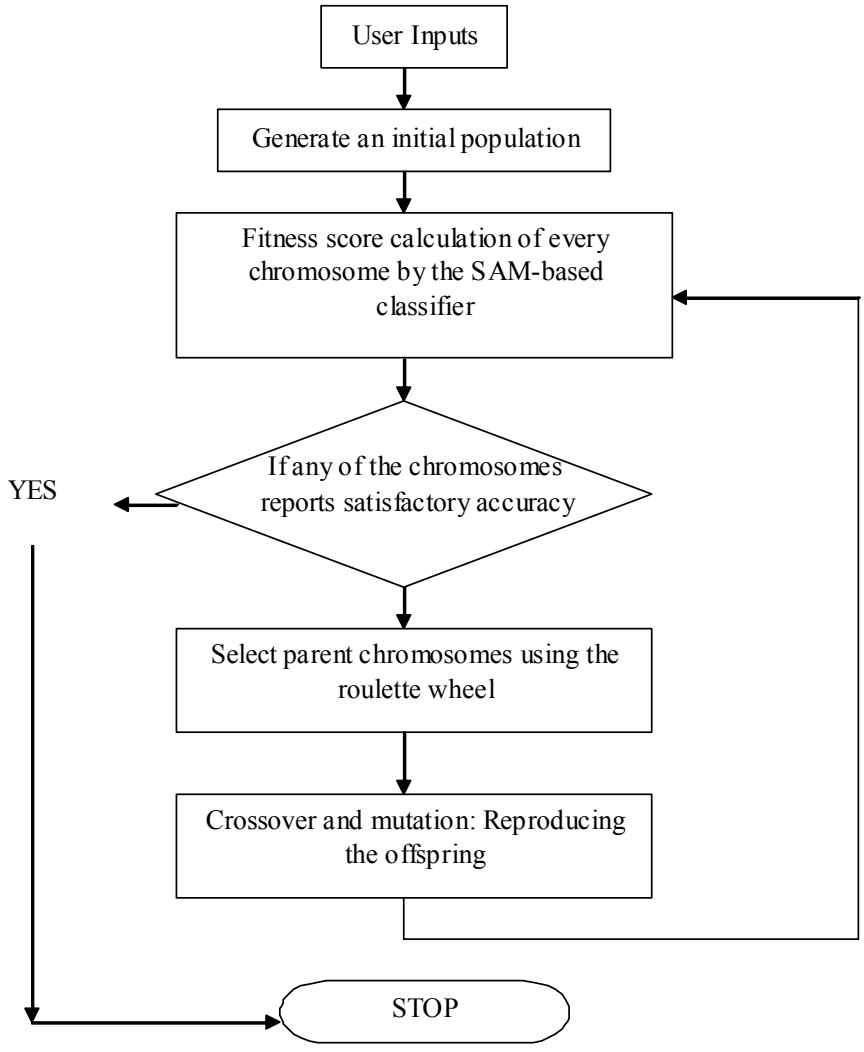


***Figure A.1.1**: A flowchart showing the details of the genetic algorithm used in this thesis, a combination between a genetic search algorithm and a spectral angle mapper based classifier*

STEP 1: First of all, the user has to provide a number of initial input requirements: (i) a population size (the number of chromosomes in a generation); (ii) a chromosome length (the number of image band labels per chromosome); (iii) a crossover rate; (iv) a mutation rate; (v) an input hyperspectral image with class samples; (vi) a fitness score threshold; and (vii) a maximum limit of evolutionary generations.

STEP 2: The population of the first generation is randomly generated according to the population size (N) specified by the user. Figure A.1.2 shows a population of randomly generated chromosomes of the first generation (e.g., N=100) when the user-defined chromosome length is 10. Each chromosome is randomly assigned with 10 image band labels without repetition. Given that the hyperspectral dataset in use has, for example, 2151 channels in total, the possible band labels to be assigned to a chromosome are between 1 and 2151.

B2	B7	B19	B55	B134	B337	B583	B799	B1200	B2120

1st chromosome with Fitness score=77%

B3	B5	B10	B45	B150	B368	B579	B791	B1210	B2121

2nd chromosome with Fitness score=78%

B3	B6	B11	B44	B131	B352	B569	B770	B1201	B2122

3rd chromosome with Fitness score=69%

……..

……..

……..

B1	B5	B13	B57	B132	B351	B567	B780	B1197	B2111

100th chromosome with Fitness score=55%

Figure A.1.2*: A randomly generated population of one hundred length-10 imaginary chromosomes is illustrated. Each chromosome is randomly assigned with 10 spectral bands without repetition.*

STEP 3: Before natural competition can start, each chromosome of the population has to have its own implicit strength relative to the others (i.e., a fitness score) (see Figure A.1.2). A fitness score of a chromosome indicates the chance for mating. A spectral angle mapper (SAM) based

nearest neighbour classifier is used as a tool for calculating the fitness scores for every chromosome so as to determine their relative strengths and weaknesses in terms of class separability. Following the definition of Goldberg (1989), the classifier works as a "fitness function".

STEP 4: Check the condition: if any chromosome in the current generation has its fitness score higher than the desired value indicated by the user in STEP 1 (e.g., 80% estimated accuracy), then stop and report the spectral band content of the winning chromosome. Otherwise, go on to the next step of the flowchart.

STEP 5: Use "the biased roulette wheel" (Goldberg, 1989) to select N chromosomes as parents of the next generation offspring. The biased roulette wheel is a metaphoric tool for making a selective competition. A chromosome with a larger fitness value obtains a larger area on the wheel. As a result, it has a higher chance to be selected. The roulette wheel has to be applied N times in order to select N parent chromosomes for reproduction.

STEP 6: After the selection process is done, the selected parents are paired and mated to produce the next generation. As an example, if the 1st and 2nd chromosomes in Figure A.1.2 are selected as parents, the image bands inside the two chromosomes will be exchanged and inherited to their offspring. This process is shown in Figure A.1.3 where the offspring share the first half from the 1st parent and the other half from the 2nd parent. Please note that the location of exchanging the bands is not always at the middle of the chromosome length like this example, but, in fact, it is randomly selected. Moreover, if the crossover rate selected by the user in STEP 1 is less than 100%, some of the selected parents may not have to mate to produce young, but, instead, they will be copied to the next generation. This reproduction process is repeated for all the parent chromosomes until the whole new generation (N offspring) is created. As a result, the new generation of N chromosomes tends to gain higher class separability because they inherit the image bands from the strong parent chromosomes.

B2	B7	B19	B55	B134	B337	B583	B799	B1200	B2120	1st parent chromosome
B3	B5	B10	B45	B150	B368	B579	B791	B1210	B2121	2nd parent chromosome
B2	B7	B19	B55	B134	B368	B579	B791	B1210	B2121	1st offspring chromosome
B3	B5	B10	B45	B150	B337	B583	B799	B1200	B2120	2nd offspring chromosome

Figure A.1.3*: The reproduction of the offspring of the first and second parents (from Figure A.1.2) by crossing over the image bands*

The effect of random mutation is also added at this phase (not shown). So, some of the chromosomes are mutated corresponding to a pre-defined chance that the user indicated (mutation rate). At the end, a new epoch of the calculation is to start over again if the maximum limit of generations has not been reached.

Appendix II: Spectral signatures of the mangroves

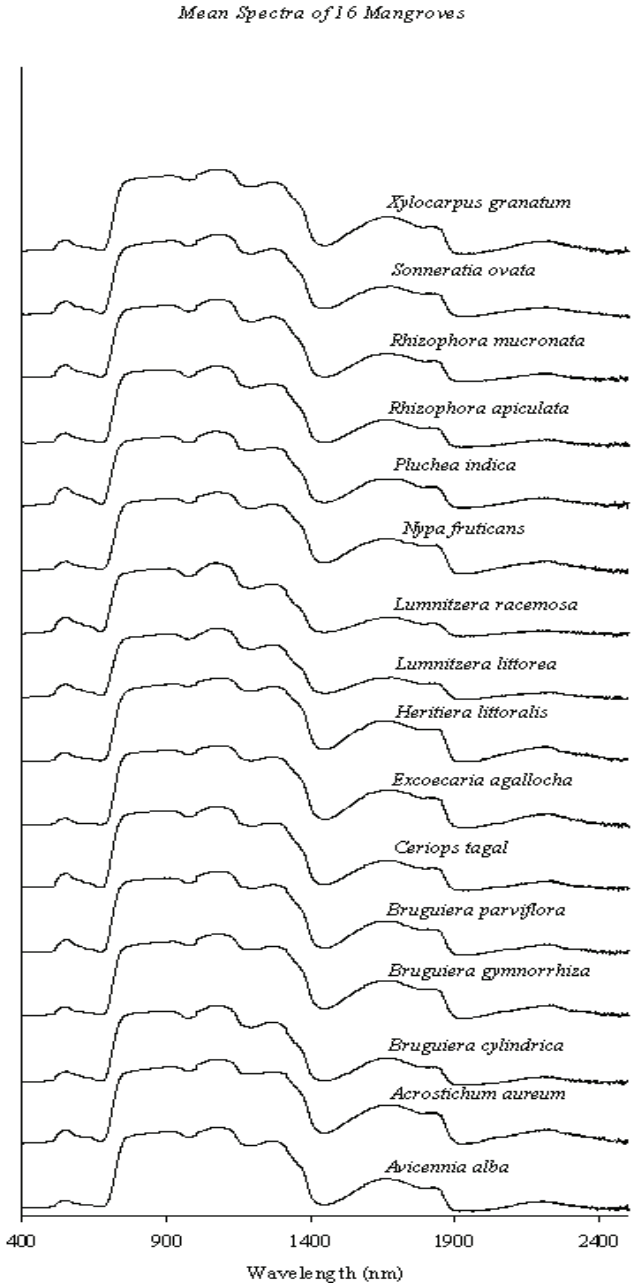


Figure A.2.1*: Mean leaf spectra of 16 mangrove species stacked on top of one another*

Appendix III: The variogram of soil pH interpolation

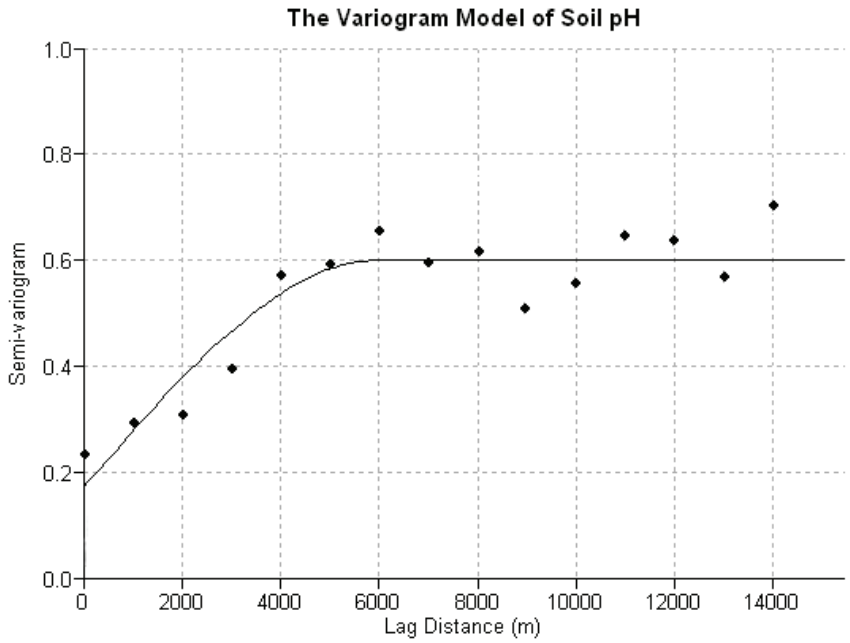


Figure A.3.1*: The omnidirectional variogram of soil pH (Spherical model, Range = 6,000, Sill = 0.6, Nugget = 0.17) used for the interpolation in Chapter 5.*

ITC Dissertation List

1. **Akinyede** (1990), Highway cost modelling and route selection using a geotechnical information system
2. **Pan He Ping** (1990), 90-9003-757-8, Spatial structure theory in machine vision and applications to structural and textural analysis of remotely sensed images
3. **Bocco Verdinelli, G.** (1990), Gully erosion analysis using remote sensing and geographic information systems: a case study in Central Mexico
4. **Sharif, M**. (1991), Composite sampling optimization for DTM in the context of GIS
5. **Drummond, J.** (1991), Determining and processing quality parameters in geographic information systems
6. **Groten, S.** (1991), Satellite monitoring of agro-ecosystems in the Sahel
7. **Sharifi, A.** (1991), 90-6164-074-1, Development of an appropriate resource information system to support agricultural management at farm enterprise level
8. **Zee, D. van der** (1991), 90-6164-075-X, Recreation studied from above: Air photo interpretation as input into land evaluation for recreation
9. **Mannaerts, C.** (1991), 90-6164-085-7, Assessment of the transferability of laboratory rainfall-runoff and rainfall - soil loss relationships to field and catchment scales: a study in the Cape Verde Islands
10. **Ze Shen Wang** (1991), 90-393-0333-9, An expert system for cartographic symbol design
11. **Zhou Yunxian** (1991), 90-6164-081-4, Application of Radon transforms to the processing of airborne geophysical data
12. **Zuviria, M. de** (1992), 90-6164-077-6, Mapping agro-topoclimates by integrating topographic, meteorological and land ecological data in a geographic information system: a case study of the Lom Sak area, North Central Thailand
13. **Westen, C. van** (1993), 90-6164-078-4, Application of Geographic Information Systems to landslide hazard zonation
14. **Shi Wenzhong** (1994), 90-6164-099-7, Modelling positional and thematic uncertainties in integration of remote sensing and geographic information systems
15. **Javelosa, R.** (1994), 90-6164-086-5, Active Quaternary environments in the Philippine mobile belt
16. **Lo King-Chang** (1994), 90-9006526-1, High Quality Automatic DEM, Digital Elevation Model Generation from Multiple Imagery
17. **Wokabi, S.** (1994), 90-6164-102-0, Quantified land evaluation for maize yield gap analysis at three sites on the eastern slope of Mt. Kenya
18. **Rodriguez, O.** (1995), Land Use conflicts and planning strategies in urban fringes: a case study of Western Caracas, Venezuela
19. **Meer, F. van der** (1995), 90-5485-385-9, Imaging spectrometry & the Ronda peridotites
20. **Kufoniyi, O.** (1995), 90-6164-105-5, Spatial coincidence: automated database updating and data consistency in vector GIS
21. **Zambezi, P.** (1995), Geochemistry of the Nkombwa Hill carbonatite complex of Isoka District, north-east Zambia, with special emphasis on economic minerals
22. **Woldai, T.** (1995), The application of remote sensing to the study of the geology and structure of the Carboniferous in the Calañas area, pyrite belt, SW Spain

23. **Verweij, P.** (1995), 90-6164-109-8, Spatial and temporal modelling of vegetation patterns: burning and grazing in the Paramo of Los Nevados National Park, Colombia
24. **Pohl, C.** (1996), 90-6164-121-7, Geometric Aspects of Multisensor Image Fusion for Topographic Map Updating in the Humid Tropics
25. **Jiang Bin** (1996), 90-6266-128-9, Fuzzy overlay analysis and visualization in GIS
26. **Metternicht, G.** (1996), 90-6164-118-7, Detecting and monitoring land degradation features and processes in the Cochabamba Valleys, Bolivia. A synergistic approach
27. **Hoanh Chu Thai** (1996), 90-6164-120-9, Development of a Computerized Aid to Integrated Land Use Planning (CAILUP) at regional level in irrigated areas: a case study for the Quan Lo Phung Hiep region in the Mekong Delta, Vietnam
28. **Roshannejad, A.** (1996), 90-9009284-6, The management of spatio-temporal data in a national geographic information system
29. **Terlien, M.** (1996), 90-6164-115-2, Modelling Spatial and Temporal Variations in Rainfall-Triggered Landslides: the integration of hydrologic models, slope stability models and GIS for the hazard zonation of rainfall-triggered landslides with examples from Manizales, Colombia
30. **Mahavir, J.** (1996), 90-6164-117-9, Modelling settlement patterns for metropolitan regions: inputs from remote sensing
31. **Al-Amir, S.** (1996), 90-6164-116-0, Modern spatial planning practice as supported by the multi-applicable tools of remote sensing and GIS: the Syrian case
32. **Pilouk, M.** (1996), 90-6164-122-5, Integrated modelling for 3D GIS
33. **Duan Zengshan** (1996), 90-6164-123-3, Optimization modelling of a river-aquifer system with technical interventions: a case study for the Huangshui river and the coastal aquifer, Shandong, China
34. **Man, W.H. de** (1996), 90-9009-775-9, Surveys: informatie als norm: een verkenning van de institutionalisering van dorp - surveys in Thailand en op de Filippijnen
35. **Vekerdy, Z.** (1996), 90-6164-119-5, GIS-based hydrological modelling of alluvial regions: using the example of the Kisaföld, Hungary
36. **Pereira, Luisa** (1996), 90-407-1385-5, A Robust and Adaptive Matching Procedure for Automatic Modelling of Terrain Relief
37. **Fandino Lozano, M.** (1996), 90-6164-129-2, A Framework of Ecological Evaluation oriented at the Establishment and Management of Protected Areas: a case study of the Santuario de Iguaque, Colombia
38. **Toxopeus, B.** (1996), 90-6164-126-8, ISM: an Interactive Spatial and temporal Modelling system as a tool in ecosystem management: with two case studies: Cibodas biosphere reserve, West Java Indonesia: Amboseli biosphere reserve, Kajiado district, Central Southern Kenya
39. **Wang Yiman** (1997), 90-6164-131-4, Satellite SAR imagery for topographic mapping of tidal flat areas in the Dutch Wadden Sea
40. **Saldana-Lopez, Asunción** (1997), 90-6164-133-0, Complexity of soils and Soilscape patterns on the southern slopes of the Ayllon Range, central Spain: a GIS assisted modelling approach

41. **Ceccarelli, T.** (1997), 90-6164-135-7, Towards a planning support system for communal areas in the Zambezi valley, Zimbabwe; a multi-criteria evaluation linking farm household analysis, land evaluation and geographic information systems
42. **Peng Wanning** (1997), 90-6164-134-9, Automated generalization in GIS
43. **Lawas, C.** (1997), 90-6164-137-3, The Resource Users' Knowledge, the neglected input in Land resource management: the case of the Kankanaey farmers in Benguet, Philippines
44. **Bijker, W.** (1997), 90-6164-139-X, Radar for rain forest: A monitoring system for land cover Change in the Colombian Amazon
45. **Farshad, A.** (1997), 90-6164-142-X, Analysis of integrated land and water management practices within different agricultural systems under semi-arid conditions of Iran and evaluation of their sustainability
46. **Orlic, B.** (1997), 90-6164-140-3, Predicting subsurface conditions for geotechnical modelling
47. **Bishr, Y.** (1997), 90-6164-141-1, Semantic Aspects of Interoperable GIS
48. **Zhang Xiangmin** (1998), 90-6164-144-6, Coal fires in Northwest China: detection, monitoring and prediction using remote sensing data
49. **Gens, R.** (1998), 90-6164-155-1, Quality assessment of SAR interferometric data
50. **Turkstra, J.** (1998), 90-6164-147-0, Urban development and geographical information: spatial and temporal patterns of urban development and land values using integrated geo-data, Villaviciencia, Colombia
51. **Cassells, C.** (1998), 90-6164-234-5, Thermal modelling of underground coal fires in northern China
52. **Naseri, M.** (1998), 90-6164-195-0, Characterization of Salt-affected Soils for Modelling Sustainable Land Management in Semi-arid Environment: a case study in the Gorgan Region, Northeast, Iran
53. **Gorte B.G.H.** (1998), 90-6164-157-8, Probabilistic Segmentation of Remotely Sensed Images
54. **Tegaye, Tenalem Ayenew** (1998), 90-6164-158-6, The hydrological system of the lake district basin, central main Ethiopian rift
55. **Wang Donggen** (1998), 90-6864-551-7, Conjoint approaches to developing activity-based models
56. **Bastidas de Calderon, M.** (1998), 90-6164-193-4, Environmental fragility and vulnerability of Amazonian landscapes and ecosystems in the middle Orinoco river basin, Venezuela
57. **Moameni, A.** (1999), Soil quality changes under long-term wheat cultivation in the Marvdasht plain, South-Central Iran
58. **Groenigen, J.W. van** (1999), 90-6164-156-X, Constrained optimisation of spatial sampling: a geostatistical approach
59. **Cheng Tao** (1999), 90-6164-164-0, A process-oriented data model for fuzzy spatial objects
60. **Wolski, Piotr** (1999), 90-6164-165-9, Application of reservoir modelling to hydrotopes identified by remote sensing
61. **Acharya, B.** (1999), 90-6164-168-3, Forest biodiversity assessment: A spatial analysis of tree species diversity in Nepal

62. **Akbar Abkar, Ali** (1999), 90-6164-169-1, Likelihood-based segmentation and classification of remotely sensed images
63. **Yanuariadi, T.** (1999), 90-5808-082-X, Sustainable Land Allocation: GIS-based decision support for industrial forest plantation development in Indonesia
64. **Abu Bakr**, **Mohamed** (1999), 90-6164-170-5, An Integrated Agro-Economic and Agro-Ecological Framework for Land Use Planning and Policy Analysis
65. **Eleveld, M.** (1999), 90-6461-166-7, Exploring coastal morphodynamics of Ameland (The Netherlands) with remote sensing monitoring techniques and dynamic modelling in GIS
66. **Yang Hong** (1999), 90-6164-172-1, Imaging Spectrometry for Hydrocarbon Microseepage
67. **Mainam, Félix** (1999), 90-6164-179-9, Modelling soil erodibility in the semiarid zone of Cameroon
68. **Bakr, Mahmoud** (2000), 90-6164-176-4, A Stochastic Inverse-Management Approach to Groundwater Quality
69. **Zlatanova, Z.** (2000), 90-6164-178-0, 3D GIS for Urban Development
70. **Ottichilo, Wilber K.** (2000), 90-5808-197-4, Wildlife Dynamics: An Analysis of Change in the Masai Mara Ecosystem
71. **Kaymakci, Nuri** (2000), 90-6164-181-0, Tectono-stratigraphical Evolution of the Cankori Basin (Central Anatolia, Turkey)
72. **Gonzalez, Rhodora** (2000), 90-5808-246-6, Platforms and Terraces: Bridging participation and GIS in joint-learning for watershed management with the Ifugaos of the Philippines
73. **Schetselaar, Ernst** (2000), 90-6164-180-2, Integrated analyses of granite-gneiss terrain from field and multisource remotely sensed data. A case study from the Canadian Shield
74. **Mesgari, Saadi** (2000), 90-3651-511-4, Topological Cell-Tuple Structure for Three-Dimensional Spatial Data
75. **Bie, Cees A.J.M. de** (2000), 90-5808-253-9, Comparative Performance Analysis of Agro-Ecosystems
76. **Khaemba, Wilson M**. (2000), 90-5808-280-6, Spatial Statistics for Natural Resource Management
77. **Shrestha, Dhruba** (2000), 90-6164-189-6, Aspects of erosion and sedimentation in the Nepalese Himalaya: highland-lowland relations
78. **Asadi Haroni, Hooshang** (2000), 90-6164-185-3, The Zarshuran Gold Deposit Model Applied in a Mineral Exploration GIS in Iran
79. **Raza, Ale** (2001), 90-3651-540-8, Object-oriented Temporal GIS for Urban Applications
80. **Farah, Hussein** (2001), 90-5808-331-4, Estimation of regional evaporation under different weather conditions from satellite and meteorological data. A case study in the Naivasha Basin, Kenya
81. **Zheng**, **Ding** (2001), 90-6164-190-X, A Neural - Fuzzy Approach to Linguistic Knowledge Acquisition and Assessment in Spatial Decision Making
82. **Sahu, B.K.** (2001), Aeromagnetics of continental areas flanking the Indian Ocean; with implications for geological correlation and reassembly of Central Gondwana

83. **Alfestawi, Y.** (2001), 90-6164-198-5, The structural, paleogeographical and hydrocarbon systems analysis of the Ghadamis and Murzuq Basins, West Libya, with emphasis on their relation to the intervening Al Qarqaf Arch
84. **Liu, Xuehua** (2001), 90-5808-496-5, Mapping and Modelling the Habitat of Giant Pandas in Foping Nature Reserve, China
85. **Oindo, Boniface Oluoch** (2001), 90-5808-495-7, Spatial Patterns of Species Diversity in Kenya
86. **Carranza, Emmanuel John** (2002), 90-6164-203-5, Geologically-constrained Mineral Potential Mapping
87. **Rugege, Denis** (2002), 90-5808-584-8, Regional Analysis of Maize-Based Land Use Systems for Early Warning Applications
88. **Liu, Yaolin** (2002), 90-5808-648-8, Categorical Database Generalization in GIS
89. **Ogao, Patrick** (2002), 90-6164-206-X, Exploratory Visualization of Temporal Geospatial Data using Animation
90. **Abadi, Abdulbaset M.** (2002), 90-6164-205-1, Tectonics of the Sirt Basin – Inferences from tectonic subsidence analysis, stress inversion and gravity modelling
91. **Geneletti, Davide** (2002), 90-5383-831-7, Ecological Evaluation for Environmental Impact Assessment
92. **Sedogo, Laurent G.** (2002), 90-5808-751-4, Integration of Participatory Local and Regional Planning for Resources Management using Remote Sensing and GIS
93. **Montoya, Lorena** (2002), 90-6164-208-6, Urban Disaster Management: a case study of earthquake risk assessment in Carthago, Costa Rica
94. **Ahmad, Mobin-ud-Din** (2002), 90-5808-761-1, Estimation of Net Groundwater Use in Irrigated River Basins using Geo-information Techniques: A case study in Rechna Doab, Pakistan
95. **Said, Mohammed Yahya** (2003), 90-5808-794-8, Multiscale perspectives of species richness in East Africa
96. **Schmidt, Karin** (2003), 90-5808-830-8, Hyperspectral Remote Sensing of Vegetation Species Distribution in a Saltmarsh
97. **Lopez Binnquist, Citlalli** (2003), 90-3651-900-4, The Endurance of Mexican Amate Paper: Exploring Additional Dimensions to the Sustainable Development Concept
98. **Huang, Zhengdong** (2003), 90-6164-211-6, Data Integration for Urban Transport Planning
99. **Cheng, Jianquan** (2003), 90-6164-212-4, Modelling Spatial and Temporal Urban Growth
100. **Campos dos Santos, Jose Laurindo** (2003), 90-6164-214-0, A Biodiversity Information System in an Open Data/Metadatabase Architecture
101. **Hengl, Tomislav** (2003), 90-5808-896-0, PEDOMETRIC MAPPING, Bridging the gaps between conventional and pedometric approaches
102. **Barrera Bassols, Narciso (**2003), 90-6164-217-5, Symbolism, Knowledge and management of Soil and Land Resources in Indigenous Communities: Ethnopedology at Global, Regional and Local Scales
103. **Zhan, Qingming** (2003), 90-5808-917-7, A Hierarchical Object-Based Approach for Urban Land-Use Classification from Remote Sensing Data

104. **Daag, Arturo S.** (2003), 90-6164-218-3, Modelling the Erosion of Pyroclastic Flow Deposits and the Occurrences of Lahars at Mt. Pinatubo, Philippines
105. **Bacic, Ivan** (2003), 90-5808-902-9, Demand-driven Land Evaluation with case studies in Santa Catarina, Brazil
106. **Murwira, Amon** (2003), 90-5808-951-7, Scale matters! A new approach to quantify spatial heterogeneity for predicting the distribution of wildlife
107. **Mazvimavi, Dominic** (2003), 90-5808-950-9, Estimation of Flow Characteristics of Ungauged Catchments. A case study in Zimbabwe
108. **Tang, Xinming** (2004), 90-6164-220-5, Spatial Object Modelling in Fuzzy Topological Spaces with Applications to Land Cover Change
109. **Kariuki, Patrick** (2004), 90-6164-221-3, Spectroscopy and Swelling Soils; an integrated approach
110. **Morales, Javier** (2004), 90-6164-222-1, Model Driven Methodology for the Design of Geo-information Services
111. **Mutanga, Onisimo** (2004), 90-5808-981-9, Hyperspectral Remote Sensing of Tropical Grass Quality and Quantity
112. **Šliužas, Ričardas V.** (2004), 90-6164-223-X, Managing Informal Settlements: a study using geo-information in Dar es Salaam, Tanzania
113. **Lucieer, Arko** (2004), 90-6164-225-6, Uncertainties in Segmentation and their Visualisation
114. **Corsi, Fabio** (2004), 90-8504-090-6, Applications of existing biodiversity information: Capacity to support decision-making
115. **Tuladhar, Arbind** (2004), 90-6164-224-8, Parcel-based Geo-information System: Concepts and Guidelines
116. **Elzakker, Corné van** (2004), 90-6809-365-7, The use of maps in the exploration of geographic data
117. **Nidumolu, Uday Bhaskar** (2004), 90-8504-138-4, Integrating Geo-information models with participatory approaches: applications in land use analysis
118. **Koua, Etien L.** (2005), 90-6164-229-9, Computational and Visual Support for Exploratory Geovisualization and Knowledge Construction
119. **Blok, Connie A.** (2005), Dynamic visualization variables in animation to support monitoring of spatial phenomena
120. **Meratnia, Nirvana** (2005), 90-365-2152-1, Towards Database Support for Moving Object Data
121. **Yemefack, Martin** (2005), 90-6164-233-7, Modelling and monitoring Soil and Land Use Dynamics within Shifting Agricultural Landscape Mosaic Systems
122. **Kheirkhah, Masoud** (2005), 90-8504-256-9, Decision support system for floodwater spreading site selection in Iran
123. **Nangendo, Grace** (2005), 90-8504-200-3, Changing forest-woodland-savanna mosaics in Uganda: with implications for conservation
124. **Mohamed, Yasir Abbas** (2005), 04-15-38483-4, The Nile Hydroclimatology: impact of the Sudd wetland (Distinction)
125. **Duker, Alfred, A.** (2005), 90-8504-243-7, Spatial analysis of factors implicated in *mycobacterium ulcerans* infection in Ghana
126. **Ferwerda, Jelle, G.,** (2005), 90-8504-209-7, Charting the Quality of Forage: Measuring and mapping the variation of chemical components in foliage with hyperspectral remote sensing

127. **Martinez, Javier** (2005), 90-6164-235-3, Monitoring intra-urban inequalities with GIS-based indicators. With a case study in Rosario, Argentina
128. **Saavedra, Carlos** (2005), 90-8504-289-5, Estimating spatial patterns of soil erosion and deposition in the Andean region using Geo-information techniques. A case study in Cochabamba, Bolivia