# Advanced Remote Sensing Techniques for Monitoring Complex Ecosystems: Spectral Indices, Unmixing, and Classification of Wetlands

## 複雑な生態系を監視するための高度リモートセンシング技術：湿原のスペクトル指数、ミクセル分解、および分類

Yoshiki Yamagata
山形与志樹

# Advanced Remote Sensing Techniques for Monitoring Complex Ecosystems: Spectral Indices, Unmixing, and Classification of Wetlands

Yoshiki Yamagata

A dissertation submitted for the degree of
Doctor of Philosophy (Macro Science)
in The University of Tokyo
1997

# Abstract

New satellite sensors are now providing a huge number of time series of remotely sensed images of the earth's surface. There is an increasing need for improved techniques to extract information about the earth ecosystems from these remotely sensed data, because remote sensing is the only comprehensive approach to monitor the global environment. However, until recently, there have been few studies on the techniques for monitoring complex natural ecosystems such as wetlands.

This thesis investigates new techniques for environmental monitoring using remotely sensed images, with a special emphasis on wetland monitoring. There are 3 major original contributions in this thesis: First is the development of new spectral indices, such as PVI (Perpendicular Vegetation Index), WTI (Water Turbidity Index), and VSWI (Vegetation-Soil-Water Index), that can be used to monitor the vegetation, water, and soil conditions in wetland areas. These indices allow assessment of the states of wetland environments and changes in them. Second is the development of a new unmixing method called the subspace method. This new unmixing method is effective for delineating continuous vegetation distribution using spectral image data. Third is the development of a new classification method using gaussian process modeling. This new method is efficient in classifying sensor fusion data. All these new methods are theoretically formulated and then subjected to experiment using various time series, spectral, and radar data.

The thesis is composed of 10 chapters:

In Chapter 1, the fundamental relevance of wetland monitoring and remote sensing in general is discussed, and brief descriptions are provided of spectral indices, unmixing, classification, and the Bayesian approach.

In Chapter 2, several multivariate analyses are conducted using multitemporal Landsat TM image data to analyze the relationship between inundation by flooding and the response by vegetation.

In Chapter 3, new spectral indices - Perpendicular Vegetation Index (PVI) and Water Turbidity Index (WTI) - are developed for monitoring the degree of inundation and vegetation change. The WTI along the turbid water line (TWL) and the PVI for paddy rice, are defined. The relationships between flooding, water turbidity, and the vegetation change of paddy rice are analyzed using these indices.

In Chapter 4, the PVI is further developed to a Vegetation-Soil-Water Index (VSWI). An algorithm that can automatically determine the end-member spectral points of vegetation, soil, and water is developed. The VSWI is applied to wetland monitoring using multitemporal Landsat TM data.

In Chapter 5, a new approach of unmixing by the subspace method is developed, and applied to wetland vegetation unmixing using hyperspectral image data. Unmixing by the subspace method is superior to conventional methods in numerical stability and computational speed for hyperspectral imagery.

In Chapter 6, by applying feature-selection methods, effective band combinations for classifying wetland vegetation types are investigated using airborne MSS data. Feature selection is performed using a measure of separability, which is then cross-validated. An effective band combination is selected and used to produce a wetland vegetation map.

In Chapter 7, the method of wetland vegetation classification using multitemporal remotely sensed data is established using multitemporal Landsat TM data. Also, by conducting biomass sampling and from spectral reflectance measurements, I establish that wetland vegetation types have distinct temporal spectral patterns.

In Chapter 8, wetland vegetation classification is attempted using high-resolution airborne spectral image data (CASI), and the relationship between vegetation and elevation is analyzed by overlaying a digital elevation model on the CASI image.

In Chapter 9, a new classification method for classifying wetland vegetation types using Landsat TM, JERS1 SAR, and ERS1 SAR data based on the Gaussian process is formulated and tested experimentally. The accuracy is assessed by comparison with maximum likelihood and Bayesian neural network methods. The Gaussian process method outperforms other methods, especially for sensor fusion data.

In Chapter 10, the contributions of this thesis are summarized, and the possible applications of newly developed advanced techniques for monitoring wetlands and other complex ecosystems are discussed, and finally the remaining problems and directions for future research are described.

# Acknowledgments

I would like to express my gratitude to all of my committee members and, in particular, to Professor R. Hirasawa of the University of Tokyo, my committee chairman, for all of his guidance and support. I would also like to express my gratitude to Drs. Yasuoka and Tamura at NIES, Professor T. Iwakuma of the University of Hokkaido, and Dr. Brian Lees of the Australian National University for their advice, encouragement, and all sorts of support. In addition, I am grateful to Dr. Fujita, Mr. Oguma, Dr. Sugita, and all of the members of the remote sensing research group of my laboratory for their cooperation provided in carrying out much of the experimental work. Finally, I would also like to thank my wife Yumi and daughters Miki and Maki for their patience and understanding over the last several years.

# Table of Contents

# Chapter 1

# Introduction

> Our understanding of our planet is sadly deficient. True, we have grown familiar with spacecraft photos of that lonely-looking globe hanging in the void, covered with a life sustaining biosphere that makes our planet uniquely beautiful and leaves it uniquely fragile. Yet we know next to nothing about the workings of Earth's ecosystem. (Myers 1984, p.258.)

Environmental changes have been expanding from regional to global issues due to the rapid growth of the world's population and to technology developments. Environmental problems such as global warming are major issues for humankind in the next century (Mintzer 1992). We have to make a start on the road toward sensible stewardship of our planet. However, our understanding of the Earth's ecosystem is poor, as it is based on limited research and insufficient observational data. For example, scientists do not even agree, within an order of magnitude, on how much wetland there is on Earth (Mathews 1987, Aselmann and Crutzen 1989), crucially important as that is to global warming modeling.

The use of remotely sensed data in the study of environmental changes is substantial. Remotely sensed data like aerial photos are useful at the local scale, but at the global scale they are vital. Global environmental problems require information at the global scale. Earth observation from space offers unique opportunities to obtain that information. More importantly, remotely sensed data can be used to parameterize environmental models (Foody and Curran 1994).

Given remotely sensed data, there is an increasing need for improved techniques to extract the environmental information from the data. Moreover, new satellite sensors are now providing a huge amount of time series data for environmental monitoring. However, conventional research programs are not well developed to use these high dimensional data. Investigation has focused primarily on analyses of simple land cover

types using scenes from remotely sensed images. Monitoring of complex natural ecosystems such as wetlands using time series and spectral image data has been neglected.

This thesis investigates new techniques for analyzing high dimensional, remotely sensed data acquired by optical and radar sensors, specifically focusing on their application to wetland monitoring. The research includes most important aspects of remote sensing techniques, spectral indices, unmixing, and classification. Spectral indices measure the condition of the earth surface, unmixing techniques decompose the end-members from the mixed pixels (mixels) in the scene, while classification techniques discriminate land cover types from the pixels.

The new methods that have been specifically developed in this thesis are as follows: First, new spectral indices have been developed for wetland monitoring. They include a Perpendicular Vegetation Index (PVI), a Water Turbidity Index (WTI), and a Vegetation, Soil and Water Index (VSWI). Second, a new unmixing algorithm using the subspace method (SM) has been devised to unmix effectively the high dimensional, spectral imagery data. Third, multitemporal and hyperspectral remotely sensed image data have been tested for classifying wetland vegetation types. Fourth, in order to select effective band combinations, the feature selection method using cross-validation has been employed for selecting effective bands for wetland classification. And finally, the latest Bayesian approach, Bayesian neural networks (BNN), and the Gaussian process (GP) have been tested for classification using multitemporal sensor fusion of optical and microwave sensor data. In this thesis, all these new methodologies are theoretically formulated and tested experimentally for their effectiveness.

This introductory chapter discusses the fundamental relevance of wetland monitoring and remote sensing in general, and provides brief descriptions of conventional methods for spectral indices, unmixing, and classification. Finally, the latest theoretical Bayesian approach is briefly described. The basic knowledge described here is important for understanding the main chapters of this thesis.

The structure of the chapter is as follows: the first section describes the importance of global wetland monitoring, with the special emphasis on its connection with the global warming problem. The second section explains remote sensing methodology in general, as well as the principles and characteristics of the sensors used in the analyses. The

third section describes briefly several conventional spectral indices. The fourth section explains the basic idea of the unmixing approach. The fifth section describes supervised classification, with special emphasis on the maximum likelihood method. The sixth section explains the principles of the statistical feature selection method for classification. And the final section describes the latest Bayesian approach, which is employed as a new classification method in this thesis, with emphasis on the difference between the Bayesian and conventional statistical approaches. This chapter concludes with an overview of the structure of the thesis.

## 1.1 Global Wetland Monitoring

Wetlands are an important component of the earth geosphere-biosphere system. Wetlands have existed on land throughout the history of the Earth. During the Pleistocene, large climatic oscillations resulted in widely changing wetland patterns, which continued to a lesser extent during the Holocene (Masing et al. 1990). At present, wetlands are strongly influenced by man, and their distribution and extent have been largely changed by human activities over the last century. Many of the Earth's wetlands lie within the permafrost areas at high latitude. The amount of wetlands in humid climates and tropical zones is not known, nor has the size of the highly astatic wetlands within arid regions been determined.

Methane is a trace gas in the atmosphere whose concentration has been increasing at the rate of 1% per year in the last decade. Methane is one of the strongest green house gases and its contribution to global warming is almost comparable to that of carbon dioxide. Although major sources of methane have been identified, there exist large uncertainties in the estimates of the source strengths. Several investigations have shown that 10 - 40 % of total methane emission is from wetlands (Cicerone and Oremlund 1988). An important source of methane is anaerobic decomposition in wetlands and irrigated rice fields. It is also produced by enteric fermentation in ruminants, biomass burning, decomposition in landfills, and fossil fuel exploration, transport and combustion.

The wide range of estimation of methane emission in wetlands arises from large uncertainties about wetlands themselves as well as about their emission characteristics. The total size of wetland areas has been estimated by several authors. However, there is little information about their geographic distribution or ecological characteristics

(Mathews 1987). Estimates of global wetland areas range between 2 and $3.6 \times 10^6$ km$^2$. Some of these estimates consider only swamps and marshes, whereas others include peatland. However, the area of peatland alone ranges from 2.3 to $4.0 \times 10^6$ km$^2$. These numbers show that due to different classification, the distribution and the extent of wetlands are subject to rather high uncertainties. Data related to seasonality brought about by flooding and freezing are especially important for estimating the methane emission rate. Aselmann and Crutzen (1989) have included these data in estimating the geographical distribution of wetlands by compiling published maps. They arrived at a global wetland area of $5.7 \times 10^6$ km$^2$.

Accurate estimation of global wetland distribution requires seasonal information at the global scale, and remote sensing is the only methodology that can provide such information. In addition, remotely sensed data can be used to parameterize models of such wetland conditions as amount of vegetation biomass and degree of water inundation.

## 1.2 Principles of Remote Sensing

Remote sensing can be defined as the science of observation from a distance. Thus, it is contrasted with in situ sensing, in which measuring devices are either immersed in or at least touch the object of observation and measurement (Barret and Curtis 1992). Remote sensing of the Earth's surface began with the use of aerial photography in the early 1900s. Aerial-mapping cameras and photo interpretation were the tools used until the late 1960s, when the first multispectral scanner systems were flown on aircraft. In the 1960s, remote sensing became a distinct field of study. The launch of the first LANDSAT satellite in 1972 was actually the real starting point of remote sensing studies. Since the LANDSAT series of satellites, digital image processing and classification have become increasingly important aspects of remote sensing (Showengredt 1983).

Since 1983, NASA has been involved in a project on earth science, now called Earth System Science. This project studies the earth as an integrated, dynamic whole. Remote sensing has been a key technology in this project, and a new satellite platform for many earth observation sensors, the Earth Observing System (EOS), has been designed (Curran et al. 1990).

Remotely sensed imagery data acquired from airborne and satellite sensors are available in digital format. The imagery data are composed of discrete picture elements (pixels). The digital value of each pixel is radiometrically quantized into discrete brightness levels. The great advantage of digital data is that they can be processed by computer, either for further processing or for analysis.

A most important characteristic of remotely sensed image data is the wavelength band. Some sensors measure reflected solar radiation in the ultraviolet, visible, and near-to-middle infrared bands. In the microwave band, the sensors measure the relative return from the Earth's surface of the energy actually transmitted from the satellite itself. Sensors of this type are referred to as active, because the energy source is provided by the remote sensing platform. Sensors that depend on an external energy source, such as the sun, are called passive.

The properties of the sensors are characterized by the number and location of their spectral bands, the ground resolution (pixel size), and the radiometric resolution (dynamic range). Radiometric resolution means the range and number of discrete brightness values (or S/N ratio). Frequently, radiometric resolution is described by bits (8 bits has 256 levels of brightness). For example, the Landsat Thematic Mapper has 7 wavelength bands with 8-bit radiometric resolution, 6 of which have 30-m spatial resolution and a thermal band that has a spatial resolution of 120 m.

The purpose of remote sensing is to identify and quantify surface materials. Land cover type can be identified if the sensor has sufficient wavelength bands. For example, if the sensor has red and near-infrared bands, it should be easy to discriminate between vegetation, soil and water because vegetation is brighter in the near-infrared band and dark in the red band, while soil appears bright in both bands and water is black in both bands. If more than two bands are available more precise discrimination should be possible, even with land cover types that are spectrally similar to each other. Recently, the number of bands in sensors for remote sensing has been rapidly increasing.

A similar argument also applies to microwave image data. Several different wavelengths can be used to identify land cover types based on the different scattering cross-section with wavelength. However, further information can be obtained using microwave imaging, from the specific characteristic of the microwave band; that is the polarization of the transmitted and scattered radiation. The polarization of an

electromagnetic wave refers to the orientation of the propagated electric field. During scattering by surface materials, some polarization changes occur and energy can be received as horizontally and/or vertically polarized. The degree of polarization rotation that occurs can also be a useful indicator of surface material.

## 1.3 Spectral Indices

Spectral indices have been one of the most actively studied subjects in remote sensing during the last two decades. Vegetation indices in particular have been the focal point of environmental monitoring.

The spectral reflectance of a plant canopy is a combination of the reflectance spectra of plant and soil components, governed by the optical properties of these elements and photon exchanges within the canopy. As the vegetation grows, the soil contribution progressively decreases but may still remain significant, depending on plant density, row effects, canopy geometry, wind effects, and so on.

The cells in plant leaves are very effective at scattering light because of the high contrast in the index of refraction between the water-rich cell contents and the intercellular air spaces. Vegetation is very dark in visible bands (400-700 nm) because of the high absorption by pigments (mainly chlorophyll) in leaves. There is a slight increase in reflectivity around 550 nm (visible green) because the pigments are least absorptive there. In the spectral range of 700-1300 nm, plants are very bright. From 1300 nm to about 2500 nm, vegetation is relatively dark, primarily because of the absorption by leaf water. Cellulose, lignin, and other plant materials also absorb radiation in this spectral range.

Based on these spectral properties of vegetation reflectance, spectral indices have been developed for monitoring. Most of them are indices for monitoring vegetation parameters. The vegetation index is a number generated by some combination of remote sensing bands and has some empirically tested relationship with the amount of vegetation in a given image pixel. Most vegetation indices have been defined based on combinations of visible and near-infrared reflectances, such as the normalized difference vegetation index (NDVI) and simple ratio (SR), which have been used in remote sensing studies. These indices are closely related to the vegetation biomass, but are also sensitive to factors such as sensor look-angle, and soil and atmospheric

conditions. These factors complicate interpretation of the target reflectance. Several new indices, such as SAVI (soil-adjusted vegetation index) and ARVI (atmospherically resistant vegetation index) and combinations of both (SARVI), have also been developed in an attempt to minimize these influences (Rondeaux et al. 1996).

A basic assumption for these vegetation indices is that all bare soil spectra in a remotely sensed image form a line in spectral space. This is related to the concept of the soil line. Nearly all of the commonly used vegetation indices are only concerned with red and near-infrared (Red-NIR) space, so a red-near-infrared line for bare soil is assumed. This line is considered to be the line of zero vegetation. The soil line is a hypothetical line in spectral space that describes the variation in the spectrum of bare soil in the image. The line is found by locating two or more patches of bare soil pixels in the image and finding the line of best fit among them in spectral space.

Vegetation indices can be grouped into two types based on the iso-vegetation line in the Red-NIR spectral space:

1) Iso-vegetation lines converge at a simple point: The indices that use this assumption are the "ratio-based" indices, which measure the slope of the line between the point of convergence and the pixel spectrum. NDVI, SAVI, and RVI are of this type.

2) All iso-vegetation lines remain parallel to the soil line: These indices are typically called "perpendicular" indices and they measure the perpendicular distance from the soil line to the pixel spectrum. PVI, WDVI, and DVI are of this type.

NDVI is sensitive to vegetation change and has the highest dynamic range among vegetation indices. NDVI is moderately sensitive to the soil background and to atmospheric effects. PVI has a smaller dynamic range and is also sensitive to atmospheric effects. It is relatively easy to use but the soil line has to be determined first. PVI works better than NDVI when there is less vegetation cover. Many of the indices based on the soil line are inadequate if no atmospheric correction has been performed, and all vegetation indices change as the look-angle from the sensor changes.

## 1.4 Unmixing

Statistical unmixing has been intensively studied in recent decades and now it is a fairly well established method in remote sensing studies. However, the methods used are difficult to apply to unmixing using very high dimensional, remotely sensed data. In this section, the basic principles and limitations of statistical unmixing, especially the linear unmixing model, are briefly described.

In cases when ground resolution is almost the same as the ground cover units or much larger, or we need to study regions of natural or semi-natural vegetation where ground surface cover changes continuously, a suitable way of extracting information is to try to estimate how the area of each ground pixel is divided up among the different cover types. This approach is usually known as mixture modeling or unmixing. By unmixing, we can produce quantitative maps of the concentration of different cover types across the scene.

In unmixing, we need to estimate for each pixel in the image the proportions covered by each ground category, given the spectrum of the pixel. The estimation should be based on a model of the spectral mixture of the components. Ideally, we need to know the exact function of the spectral mixture. The linear mixing model is the simplest approach, although even here, the function cannot always be inverted unambiguously. Ambiguous inversion happens when we have more ground cover components than spectral bands. The mixture function depends on the reflectance spectra of the cover types of which the scene is composed. However, inadequate calibration of an image and atmospheric effects often prevent the calibration of the data to the reflectance values; hence, laboratory reflectances cannot be used for unmixing. A much more likely situation is that we have detailed field knowledge of proportional ground cover for a number of pixels; then, based on the number of pixels for which both the mixture function and spectrum of each component are known, we can unmix the model by statistical inference.

The basic physical assumption underlying the linear mixing model is that there is no significant multiple scattering between the different cover types; each photon that reaches the sensor has interacted with just one cover type. Under these conditions, the received energy can be modeled as the simple sum of energies received for cover types in that field and the coefficient of the model is proportional to the area covered. The ground cover proportions are well defined and add up to unity.

Pure component spectra of the land cover types are referred to as the end-member spectra, and the categories themselves as the end-members. The term end-member here means the spectral response without noise. Each end-member spectrum is a point in spectral space. Pixels that are mixtures of just two components will, in the absence of noise, lie in feature space along the line joining the points corresponding to the two end-member spectra. Similarly, mixtures of three end-member spectra with non-negative proportions should lie inside a triangle. In higher dimensional space, mixed pixels (mixels) lie in the hyperplane defined by the end-member spectra of the components.

Let $c$ be the number of the land cover types and $n$ be the number of spectrum dimensions. In the linear unmixing model, from the theory of linear equations, the solution of the unmixing will be unique if $c=n+1$, and if $c>n+1$ there will be an infinity of exact solutions. Finally, when $c<n+1$ there may well be no exact solution. The condition must be modified to n<c if we are relaxing the sum-to-unity constraint. The number of dimensions of the spectra here is not really the number of spectral bands, but rather the intrinsic dimensional number of the spectral data that can be known by principal component analysis. In the case of the 6 bands of Landsat Thematic Mapper data, if we find that the fifth and sixth principal components of the data contain nothing but noise, then the true dimensional number of the data is four. In this case we can unmix only up to 5 components using these data (Settle and Drake 1993).

Ideally the end-member spectra should be determined by laboratory measurements of the end-member components. However, there are sometimes substantial problems in correcting satellite data sufficiently well for atmospheric effects to allow direct comparison between laboratory data and remotely sensed image data. Thus, the end-member spectra are often determined from pure pixel values sampled from the image itself.

## 1.5 Classification

The most widely used method for extracting information on surface cover from remotely sensed data is image classification. With this technique, each image pixel is allocated exclusively to one of a small number of known categories, producing an image containing thematic information. The resulting thematic map can be used to

estimate the area of each category, if the numbers of boundary pixels or mixed pixels are small.

In this thesis, classification studies using multitemporal and hyperspectral remotely sensed data in Chapters 6 and 7, and also the feature selection study in Chapter 8, are all conducted using the supervised maximum likelihood classification method. New classification methods using the Bayesian approach developed in Chapter 8 are also compared with the maximum likelihood method. Because the latter forms the foundation of the statistical classification approach, it is important to be aware of the details of this method in order to understand the studies in this thesis. In this chapter, the basic principles of the supervised classification method and maximum likelihood method are described following Richards (1986).

### 1.5.1 Supervised Classification

Supervised classification is the most common technique for remote sensing image analysis. It classifies the pixels in an image to one of the cover types, or classes. Many algorithms have been developed for this. Some are based upon probability distribution models for the classes of interest; others partition the multispectral space into class regions using optimally located surfaces. Irrespective of the particular method chosen, the essential practical steps are:

1. Determine the set of ground cover types into which the image is to be segmented. These are land cover classes, such as water, urban regions, cropland, and range lands.

2. Choose representative or prototype pixels from each of the desired sets of classes. These pixels are called training data. Training sets for each class can be established using ground truth by site visits, maps, aerial photographs, or even photo interpretation of color-composite products formed from the image data. Often the training pixels for a given class will lie in a common region enclosed in a border. That region is then often called a training area.

3. Use the training data to estimate the parameters of the particular classifier algorithms to be used; these parameters will be the properties of the probability model used or will be equations that define partitions in the multispectral space.

The set of parameters for a given class is sometimes called the signature of that class.

4. Using the trained classifier, label or classify every pixel in the image into one of the desired ground cover types. Here the whole image is classified.

5. Produce tabular summaries or thematic (class) maps which summarize the results of the classification.

### 1.5.2 Maximum Likelihood Method

Maximum likelihood is the most common supervised classification method. When the distributions of the spectra are normal, it can be shown theoretically that the maximum likelihood method gives the best performance. In the following, the principles of the maximum likelihood method are described.

Let the land cover classes be represented by

$$\omega_i, i = 1,\ldots, M \tag{1.1}$$

Where M is the total number of classes. The method determines the class of a pixel at a location $\mathbf{x}$ by conditional probabilities

$$p(\omega_i|\mathbf{x}), i = 1,\ldots, M \tag{1.2}$$

Here, $\mathbf{x}$ is a vector of brightness values for the pixel. The probability $p(\omega_i|\mathbf{x})$ gives the likelihood that the correct class is $\omega_i$ for a pixel at position $\mathbf{x}$. Then, classification can be performed according to

$$\mathbf{x} \in \omega_i \text{ if } p(\omega_i|\mathbf{x}) > p(\omega_j|\mathbf{x}) \text{ for all } j \neq i \tag{1.3}$$

The pixel at $\mathbf{x}$ belongs to class $\omega_i$ if $p(\omega_i|\mathbf{x})$ is highest. This intuitive decision rule is a special case of a more general rule in which decisions can be biased according to different degrees of significance being attached to different incorrect classifications. The general approach is called Baye's classification.

However, the $p(\omega_i|\mathbf{x})$ values are generally unknown. If sufficient training data are available for each land cover type, then these data can be used to estimate a probability distribution $p(\mathbf{x}|\omega_i)$ for a cover type that describes the chance of finding a pixel from class $\omega_i$ at the position $\mathbf{x}$. There will be as many $p(\mathbf{x}|\omega_i)$ as there are ground cover classes. In other words, for a pixel at a position $\mathbf{x}$ in multispectral space, a set of probabilities can be computed that give the relative likelihood that the pixel belongs to each available class.

The desired $p(\omega_i|\mathbf{x})$ in (1.3) and the available $p(\mathbf{x}|\omega_i)$ - estimated from training data - are related by Baye's theorem:

$$p(\omega_i|\mathbf{x}) = p(\mathbf{x}|\omega_i)p(\omega_i) / p(\mathbf{x}) \tag{1.4}$$

where $p(\omega_i)$ is the probability that class $\omega_i$ occurs in the image, and $p(\mathbf{x})$ is the probability of finding a pixel from any class at location $\mathbf{x}$. Although $p(\mathbf{x})$ itself is not important for classification, it can be calculated as

$$p(\mathbf{x}) = \sum_{i=1}^{M} p(\mathbf{x}|\omega_i)p(\omega_i) \tag{1.5}$$

The $p(\omega_i)$ are called a priori (or prior to analysis) probabilities, since they are the probabilities with which class membership of a pixel could be guessed before classification. By comparison the $p(\omega_i|\mathbf{x})$ are called a posteriori (or after the fact) probabilities. Using (1.4) it can be seen that the classification rule of (1.3) is

$$\mathbf{x} \in \omega_i \text{ if } p(\mathbf{x}|\omega_i)p(\omega_i) > p(\mathbf{x}|\omega_j)p(\omega_j) \text{ for all } j \neq i \tag{1.6}$$

where $p(\mathbf{x})$ has been removed as a common factor. The rule of (1.6) is more acceptable than that of (1.3) since the $p(\omega_i|\mathbf{x})$ are known from training data, and it is conceivable that the $p(\omega_i)$ are also known or can be estimated from the analyst's knowledge of the image. For mathematical convenience, if the definition

$$g_i(\mathbf{x}) = \ln\{p(\mathbf{x}|\omega_i)p(\omega_i)\} = \ln p(\mathbf{x}|\omega_i) + \ln p(\omega_i) \tag{1.7}$$

is used, where ln is the natural logarithm, then (1.6) can be restated as

$$\text{if } g_i(\mathbf{x}) > g_j(\mathbf{x}) \text{ for all } j \neq i \tag{1.8}$$

where $g_i(\mathbf{x})$ are referred to as discriminant functions. This is, with one modification to follow, the decision rule used in maximum likelihood classification.

At this stage, it is assumed that the probability distributions for each class are of the form of normal multivariate models. This is an assumption, rather than a demonstrable property of natural spectral or information classes. However it leads to mathematical simplification in the following.

In (1.7) therefore, it is now assumed for N bands that

$$p(\mathbf{x}|\omega_i) = (2\pi)^{-N/2}|\Sigma_i|^{-1/2} \exp\{-\frac{1}{2}(\mathbf{x}-\mathbf{m}_i)^t \Sigma_i^{-1}(\mathbf{x}-\mathbf{m}_i)\} \tag{1.9}$$

where $\mathbf{m}_i$ and $\Sigma_i$ are the mean vector and covariance matrix of the data in class $\omega_i$, respectively. The resulting term $-\frac{N}{2}\ln(2\pi)$ is common to all $g_i(\mathbf{x})$ and does not aid discrimination. Consequently, it is ignored and the final form of the discriminant function for maximum likelihood classification is

$$g_i(\mathbf{x}) = \ln p(\omega_i) - \frac{1}{2}\ln|\Sigma_i| - \frac{1}{2}(\mathbf{x}-\mathbf{m}_i)^t \Sigma_i^{-1}(\mathbf{x}-\mathbf{m}_i) \tag{1.10}$$

Often the analyst has no useful information about the $p(\omega_i)$, in which case equal prior probabilities are assumed; as a result $\ln p(\omega_i)$ can be removed from (1.10) since it is the same for all i. In that case the 1/2 common factor can also be removed leaving, as the discriminant function:

$$g_i(\mathbf{x}) = -\ln|\Sigma_i| - (\mathbf{x}-\mathbf{m}_i)^t \Sigma_i^{-1}(\mathbf{x}-\mathbf{m}_i) \tag{1.11}$$

Implementation of the maximum likelihood decision rule involves using either (1.10) or (1.11) in (1.8).

As a means for assessing the capabilities of the maximum likelihood decision rule, it is useful to determine the shapes of the surface that separate one class from another in multispectral space. These surfaces can be determined in the following manner.

Spectral classes are defined by those regions in multispectral space where their discriminant functions have the highest values.Clearly these regions are separated by surfaces where the discriminant functions for adjoining spectral classes are equal. The *i*th and *j*th spectral classes are separated therefore by the surface

$$g_i(\mathbf{x}) - g_j(\mathbf{x}) = 0 \tag{1.12}$$

This is referred to as a decision surface, since if all the surfaces separating spectral classes are known, decisions about class membership of an image pixel can be made on the basis of the position relative to the complete set of surfaces.

The construction $(\mathbf{x}-\mathbf{m}_i)^t \sum_i^{-1} (\mathbf{x}-\mathbf{m}_i)$ in (1.10) and (1.11) is a quadratic function of $\mathbf{x}$. Consequently, the decision surfaces implemented by maximum likelihood classification are quadratic and thus take the form of parabolas, circles, and ellipses.

## 1.6 Feature Selection

As the number of features (bands) of remotely sensed data increases, theoretically the accuracy of classification using these data should also increase. However, in real situations, this is not the case. A decrease in classification accuracy often occurs when the number of training samples is small. This is because when the number of features increases, the number of parameters also increases, and more rapidly, such that the number of training samples necessary for accurately estimating these parameters easily becomes larger than the actual training data available for classification. Feature selection or feature reduction is necessary to compensate for this problem and to realize higher classification accuracy using high dimensional remotely sensed data. In this section, the basic methods of feature selection employed in this thesis are described following Richards (1996).

### 1.6.1 Separability

As the number of features or bands increases, the classification cost also increases in any classification algorithm. In classifiers such as the minimum distance method, the cost increases linearly with the number of features. However in maximum likelihood classification, the cost increases quadratically with number of features. Therefore, it is

desirable to select only effective features to perform an efficient classification. Features that have low separability of spectral classes should be discarded not only to save computational costs but also to avoid failure in estimating the parameters of the classifier.

To perform feature selection, relative classification performance of a set of features has to be assessed quantitatively. A procedure commonly used is to determine the mathematical separability of classes. If the separability is not lowered by the removal of certain features, then those features are considered of little value in the classification process.

There are several ways to measure the separation or overlap between a pair of probability distributions of classification classes. Obviously the distance between means is insufficient since overlap will also be influenced by the variance of the distributions. Instead, a combination of both the distance between means and the covariance of the class distribution is necessary for the evaluation. Such measures are all referred to as separability, in which the separation of distributions is evaluated by statistical pattern classification.

### 1.6.2 Divergence

Divergence is a measure of the separability of a pair of probability distributions in their degree of overlap. It is defined in terms of the likelihood ratio

$$L_{ij}(\mathbf{x}) = p(\mathbf{x}|\omega_i) / p(\mathbf{x}|\omega_j) \tag{1.13}$$

This gives the divergence between a pair of spectral classes that are normally distributed. If there are more than two spectral classes, all pairwise divergences need to be checked to see whether a particular feature subset has sufficient separation. Then, the total separability is evaluated by the average divergence.

However, there is a problem with divergence. As spectral classes become separated from each other in multispectral space, the accuracy of classification approaches 1.0. The divergence measure will increase quadratically with separation between spectral classes. This rate of increase of divergence is quite misleading if divergence is used as an indication of how successfully the pattern in the corresponding spectral classes can

be classified. Moreover, if there are some outlying easily separable classes, they will increase the average divergence in a misleading way. This will result in the selection of sub-optimal feature subsets at best. This problem makes divergence an unsuitable indicator in some cases. However, the Jeffries-Matusita distance described in the next section does not suffer from this drawback.

### 1.6.3 Jeffries-Matusita (JM) Distance

The JM distance (Bhattacharrya distance) between a pair of probability distributions (spectral classes) is defined as

$$J_{ij} = \int \{\sqrt{p(\mathbf{x}|\omega_i)} - \sqrt{p(\mathbf{x}|\omega_j)}\}^2 d\mathbf{x} \tag{1.14}$$

which is seen to be a measure of the average distance between the two class density functions. For normally distributed classes this becomes

$$J_{ij} = 1000\sqrt{2(1-e^{-\alpha})} \tag{1.15}$$

in which

$$\alpha = \frac{1}{8}(m_i - m_j)^t \{\frac{\Sigma_i + \Sigma_j}{2}\}^{-1}(m_i - m_j) + \frac{1}{2}\ln\{\frac{(\Sigma_i + \Sigma_j)/2}{|\Sigma_i|^{1/2}|\Sigma_j|^{1/2}}\} \tag{1.16}$$

It is of interest to note that the first term in (1.16) is akin to the square of the normalized distance between the class means. The presence of the exponential factor gives an exponentially decreasing weight to increasing separations between spectral classes. The function is asymptotic to 1414 so that a JM distance of 1414 between spectral classes would imply 100% classification accuracy for those classes. This saturating behavior is highly desirable since it does not suffer from the difficulty that affects the divergence parameter.

## 1.7 Bayesian Learning

The latest Gaussian process is a Bayesian approach first developed in regression and function estimation studies based on Bayesian learning. Bayesian learning provides a completely different interpretation of probability and modeling from the conventional statistical approach (frequentist point of view). In this chapter, in order to make it easier to understand the later chapter using the Bayesian approach, the principles of Bayesian learning are briefly described, with special emphasis on the difference between Bayesian and frequentist approaches, following Neal (1996).

### 1.7.1 Bayesian and Frequentist Approaches

Bayesian learning uses the term probability to include all forms of uncertainty as to the degree of belief in various possibilities. Learning and inference are then performed by simple applications of the probability distribution over all unknown quantities In contrast, the conventional frequentist approach in statistics uses probabilities as long-run frequencies of repeatable events. In frequentist learning, estimates are made of unknown quantities to produce a good model.

To see the difference between Bayesian and frequentist learning, consider the case of tossing a coin. In Bayesian learning, there is an uncertainty about each toss: the coin has a certain probability of landing heads or tails. Because we do not know the probability of its landing heads, we will update this probability distribution using the rule of probability theory after each toss. In the frequentist approach, such a probability update is not necessary. Instead, we choose some estimator for the unknown probability of heads according to some criterion using the frequency of heads in past tosses.

### 1.7.2 Models and Likelihood

Consider a data series, $x^{(1)}, x^{(2)}, \ldots$, generated by an independent random process following a probability density $P(x^{(i)}|\theta)$ with unknown model parameter $\theta$. If this density is Gaussian, we can write this with $\theta = \{\mu, \sigma\}$ as:

$$P(x^{(i)}|\mu,\sigma) = \exp(-(x^{(i)} - \mu)^2 / 2\sigma^2) / \sqrt{2\pi}\sigma$$

Learning about $\theta$ is possible after some observation of $x^{(i)}$. For both Bayesian and frequentist learning, these observations can be assessed by the likelihood function

$L(\theta)$, which gives the probability of the observed data as a function of the unknown model parameters

$$L(\theta) = L(\theta | x^{(1)},\ldots,x^{(n)}) \propto P(x^{(1)},\ldots,x^{(n)}|\theta) = \prod_{i=1}^{n} P(x^{(i)}|\theta) \tag{1.17}$$

In the maximum likelihood method, the unknown parameters are estimated to maximize the likelihood $L(\theta)$. In the case of tossing a coin, the maximum likelihood estimate $\overline{\theta}$ is the frequency of heads among $x^{(1)},\ldots,x^{(n)}$. It is known that the maximum likelihood estimates converge on the true value as the amount of observational data increases.

In a real problem, we are interested in the prediction $P(x^{(n+1)}|\overline{\theta})$, not in the value of $\theta$ itself. In a frequentist approach, a prediction is based on an estimated $\theta$. However, by using the Bayesian approach we can make a prediction that takes account of the remaining uncertainty in $\theta$.

### 1.7.3 Bayesian Learning and Prediction

Bayesian learning gives a probability distribution over model parameters that expresses our belief regarding the likelihood of the different parameter values. To start the process of Bayesian learning, we must define an a priori distribution, $P(\theta)$, that expresses our initial belief about $\theta$. When we observe $x^{(1)},\ldots,x^{(n)}$ we can update this a priori distribution to an a posteriori distribution, using Baye's Rule:

$$P(\theta | x^{(1)},\ldots,x^{(n)}) = \frac{P(x^{(1)},\ldots,x^{(n)}|\theta)}{P(x^{(1)},\ldots,x^{(n)})} \propto L(\theta | x^{(1)},\ldots,x^{(n)})P(\theta) \tag{1.18}$$

The a posteriori distribution combines the likelihood function, which contains the information about $\theta$ derived from observation, with the a priori estimate, which contains the information about $\theta$ derived from our background knowledge.

In the coin tossing example, we might start with a uniform a priori estimate for $\theta$, the probability of heads. As we see the result of several tosses, the a posteriori distribution, obtained by combining the a priori estimate with the likelihood function, will approach the observed frequency of heads. To predict the value of an unknown quantity $x^{(n+1)}$,

Bayesian learning integrates the predictions of the model with the a posteriori distribution of the parameters as

$$P(x^{(n+1)}|x^{(1)},\ldots,x^{(n)}) = \int P(x^{(n+1)}|\theta)P(\theta|x^{(1)},\ldots,x^{(n)})d\theta \tag{1.19}$$

This predictive distribution for $x^{(n+1)}$ given $x^{(1)},\ldots,x^{(n)}$ is the complete Bayesian inference regarding $x^{(n+1)}$, which can be used for many purposes.

In the coin tossing example, if we use a uniform a priori estimate for the probability of heads, the Bayesian prediction for the result of toss n+1, given the results of the first n tosses, is

$$\begin{aligned} P(x^{(n+1)}|x^{(1)},\ldots,x^{(n)}) &= (h+1)/(n+2) \text{ if } x^{(n+1)} = \text{heads} \\ &\quad (t+1)/(n+2) \text{ if } x^{(n+1)} = \text{tails} \end{aligned} \tag{1.20}$$

where h and t are the numbers of head and tails in $x^{(1)},\ldots,x^{(n)}$.

However, even in this simple problem, we can see the effect of prediction by integration rather than maximization. If we have tossed the coin twice, and each time it landed heads, naïve application of maximum likelihood will lead us to conclude that the coin is certain to land heads on the next toss, since $\overline{\theta}=1$. The Bayesian prediction with a uniform a priori estimate gives a more reasonable probability of 3/4 for heads. The Bayesian procedure avoids jumping to conclusions by considering not just the value of $\theta$ that explains the data best, but also other values of $\theta$ that explain the data reasonably well, and hence also contribute to the integral of equation (1.19)

### 1.7.4 Learning Complicated Models

There is a principle for modeling studies called "Occam's Razor". It states that we should select the simplest model for inference. However, in some complicated pattern-recognition applications, we often cannot find any simple procedure for recognition. There is no reason to suppose that there is always a limit to the complexity of a model that will give good performance. However, the complexity of models is often limited, not just to save the computational time, but also to avoid overfitting of the training data. But there is no need for concern about overfitting in the Bayesian approach.

The overfitting problem in a frequentist approach occurs because of a trade-off between the bias and the variance of an estimator. The bias of an estimator measures any systematic error of the prediction; the variance measures the degree to which the estimate is sensitive to the randomness of the training data. One way of learning is to minimize the sum of the squares bias and the variance. Since reducing bias often increases variance, and vice versa, one way to compromise would be to minimize their sum. A complicated model that is flexible enough to represent the true process can have low bias, but may suffer from high variance, since its flexibility also lets it fit the random variation in the training data. A simple model will have high bias, unless the true process is really that simple, but will have lower variance. There are also other ways to trade off bias against variance, such as by use of a penalty function, but adjusting the model complexity is perhaps the most common method.

This complexity adjustment leads to a choice of model that varies with the amount of training data available - the more data, the more complex the model used. In this way, one can sometimes guarantee that the performance achieved will approach the optimum as the size of the training set approaches infinity because the bias will go down with increasing model complexity, while the variance will also go down due to the increasing amount of data. Some information criteria can be used to determine the optimal complexity of a model at a given training set size (Ripley 1996).

In a Bayesian approach, there is no necessity for restricting the complexity of the model based on the amount of available training data. In order to perform Bayesian learning, we need to select a model type, make an a priori guess, collect data, and then compute the a posteriori values to make a prediction. There is no need to change the model or the a priori estimates depending on how much data is available. If the model and estimates are correct for a thousand observations, they should be also correct for ten observations. Thus, the Bayesian approach can employ a suitable model that is as complex as we can afford computationally, regardless of the size of the training set. In other words, in a Bayesian approach we need not be concerned about Occam's Razor, because it will be applied automatically in the Bayesian learning process.

## 1.8 Outline of the Thesis

The main part of this thesis describes newly developed remote sensing techniques and their application to wetland monitoring.

In Chapter 2, as an introduction to the spectral index analysis of flooding presented in Chapter 3, multivariate analyses are conducted using multitemporal image data. In this study, Landsat TM images acquired immediately after a flood and one month after the flood are used. The flood - rice damage relationship is analyzed by multivariate analyses, including correlation analysis between the turbidity level of flooding water and rice plant damage, multiple regression related to the decrease in rice yield, and cluster analysis to discern the flood - rice damage relationship. Based on these analyses, relationships between flooding condition and flood damage are delineated.

In Chapter 3, new spectral indices for vegetation in the inundated area are developed and applied to the flooding analysis. A water turbidity index (WTI), analogous to soil brightness, and a perpendicular vegetation index (PVI) are devised and applied to the analysis of flooding and its relationship with damage, using multitemporal Landsat TM data. Rice yields are determined and related to the PVI calculated from the TM scenes, while the WTI is used for monitoring the floodwater turbidity. The relationship between floodwater turbidity and rice yield is investigated in two test sites.

In Chapter 4, a new spectral index, the Vegetation-Soil-Water Index (VSWI), is developed and applied to wetland monitoring. The VSWI is a natural extension of PVI for monitoring not only vegetation conditions but also soil and water conditions. I also develop an algorithm to fit a triangle to the spectral distributions to determine the end-member points for vegetation, soil, and water. VSWI is tested for wetland monitoring using 6 multitemporal Landsat TM scenes, which show seasonal vegetation changes over 5 years.

In Chapter 5, a new unmixing approach by the subspace method is developed and an experiment using hyperspectral imagery is conducted. In the subspace method, unmixing is calculated as the projection of each unknown pixel vector on the subspace of each class. The performance of this method was tested in an unmixing experiment using acquired, hyperspectral, Compact Airborne Spectral Imager (CASI) images.

Unmixing by the subspace method is tested against the wetland vegetation classes and the result is compared with conventional methods by least squares, quadratic programming, and orthogonal subspace projection. Finally, the results of the unmixing experiment are evaluated with regard to wetland vegetation monitoring.

In Chapter 6, several feature selection methods are used to determine the effective spectral bands for classifying the wetland vegetation types. Optimal band combinations are selected using airborne MSS data. For feature selection, both the separability measure (the Jeffries-Matusita (JM) distance) and the cross-validation method are used. Then, the dependency of the selection on the classification method is checked by maximum likelihood and minimum distance methods. The wetland vegetation is classified using selected bands to show that the wetland vegetation type in a bog area can be classified into several community-level classes using spectral information.

Chapter 7 shows that wetland vegetation types can be accurately classified using multitemporal Landsat TM images. The growth patterns of wetland vegetation change according to vegetation type, and we can use this feature in multitemporal images for classifying the vegetation types. To clarify temporal growth patterns of wetland vegetation types, biomass sampling experiments are conducted to measure vegetation growth during the summer. Spectral reflectance measurements are conducted to see the spectral differences between the vegetation types. In Chapter 8, wetland vegetation classification is investigated using spectral image data. Wetland vegetation types in a bog area are classified using the spectral information from CASI data. Detailed elevation data are overlaid on the CASI data to see the relationship between vegetation distribution and elevation differences in the wetland area. In the classification, the k-means clustering (unsupervised learning) method is used to classify the CASI images.

In Chapter 9, a new Bayesian classification method using a Gaussian process is developed and tested using sensor fusion data from optical and radar sensors. This Gaussian process has been developed from Bayesian neural networks with an infinite number of nodes in the hidden layer. It is also a Bayesian model-averaging approach, which integrates the model's predictions with the a posteriori probability of the parameters. The basic theory of the Gaussian process for classifying satellite remote sensing data is introduced and an experiment is made using multitemporal Landsat TM, JERS1, and ERS1 SAR data. The accuracy of the classifications is compared with the maximum likelihood and Bayesian neural network methods.

In Chapter 10, I summarize the contributions of this thesis and discuss possible applications of the advanced techniques developed herein to several complex ecosystem studies. I describe specifically actual uses of the Perpendicular Vegetation Index, Water Turbidity Index, Vegetation-Soil-Water Index, Unmixing by Subspace Method, and Bayesian classification with a Gaussian process model for monitoring wetlands, forests, and inland waters. Finally, I indicate possible directions for future research.

# Chapter 2
# Flooding Analysis using Multitemporal Image Data

Rice fields, as artificial wetlands, have many common features with natural wetlands. The most important such features are inundation and vegetation growth. Remote sensing techniques that have been developed for monitoring these features in rice fields are also applicable to wetlands, and vice versa. In this chapter, as an introduction to the spectral index analysis in Chapter 3, multivariate analyses are conducted using multitemporal image data. Landsat TM images acquired immediately after a flood and one month after the flood are used. The flood - damage relationships are analyzed by multivariate analyses, including correlation analysis between the turbidity of floodwater and rice damage, a multiple regression model on the decrease of the rice yield, and cluster analysis to see the flood - damage relationship. Based on these analyses, relationships between degree of flooding and damage are delineated. This chapter is based on Yamagata et al. (1988a).

## 2.1 Introduction

Once an agricultural disaster occurs, agricultural agencies need to assess the crop damage before farmers harvest the crop. At present, the extent of such damage in a region is estimated from sampling surveys. However, a more accurate estimation of the damage cannot be obtained by ground-based methods only. Landsat and other earth observation satellite data are expected to compensate for such shortcomings.

To analyze the extent of inundation and crop damage caused by a flood, three approaches can be adopted using spectral reflectance data. First, the inundated area can be delineated using images acquired immediately after the flood (Green et al. 1983, Ramamoorthi and Rao 1985, Imhoff et al. 1987); information on the condition of flood water such as water turbidity can also be extracted (Khorram 1981, Lathrop and Lillesand 1986). Second, the extent of crop damage can be estimated using images showing crop conditions before harvest; in other words, the crop yield can be estimated

using the spectral response of the crop (Tucker et al. 1980, Pinter et al 1981, Miller et al. 1983, Aase et al. 1984, Patel et al. 1985, Shibayama and Munakata 1985). Third, the relationship between the extent of inundation and the crop damage caused can be analyzed.

Although we have few clear days in the crop growing season in Japan, fortunately we were able to acquire Landsat 5 TM (Thematic Mapper) images immediately after and one month after a flood, which showed the condition of the flood inundation and the rice crop damage before harvest, respectively. Using these multitemporal TM images, we attempted to estimate the rice crop damage and analyze the flood - damage relationships. The outline of the analyses follows.

1) We determined whether the crop damage could be estimated using image data that were acquired before the harvest and immediately after the flood. In the physical interpretation of the TM band reflectance, two notable relationships were already known from several previous investigations. First, water turbidity has a positive correlation with TM band 3 (red), because the mud in water has high reflectance in this band (Khorram 1981, Lathrop and Lillesand 1986). Second, the rice crop yield has a positive correlation with TM band 4 (near infrared); the near infrared reflectance is correlated with biomass, and biomass is linearly correlated with yield (Miller et al. 1983, Patel et al. 1985). Further, it seemed that the floodwater was more turbid where the flood stream flowed fast, and this fast stream caused the heavy rice damage. Based on these relationships, multiple regression models were used to estimate the amount of regional damage.

2) In order to classify the relationship between degree of flooding and rice damage, cluster analysis was employed (Townshend and Justice 1980). The cluster classes were delineated from the transformed image, which showed the enhanced flood-damage patterns. This transformation, prior to the clustering, was performed by principal component analysis using the composite multitemporal scene (Singh and Harrison 1985, Ingebritsen and Lyon 1985), which was already shown to be effective in enhancing land cover discrimination and detecting land-use change (Richards 1984, Townshend 1984). The objective of these studies was to delineate the qualitative pattern of land cover change, while that of this analysis was to classify the quantitative cause and effect relationship between the flood inundation and the damage caused to the rice crop.

## 2.2 Study area and image data

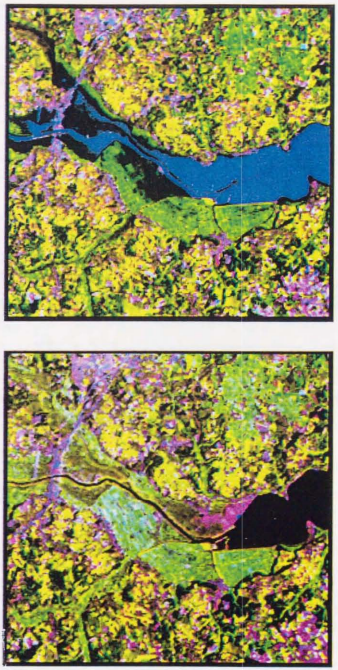
A typhoon which hit on August 4, 1986, brought about floods in several areas of the Kanto district. Many rice fields in northern Kanto district were inundated and heavily damaged. The rice fields along the Hinuma River, near the area where the river reaches Hinuma Lake (approximately 100 km north of Tokyo), were selected as the test site for our flood damage analyses.

Figure 2.1(a) is the color composite TM image using bands 3 (red), 4 (near infrared) and 5 (mid infrared) acquired on 6 August. It shows the spread of floodwater over the rice fields immediately after the flood (TM bands 3, 4, and 5 are shown in blue, green, and red, respectively). In this image, the area inundated with the turbid flood water appears blue, and the area inundated with clear water accumulated above the rice crop vegetation appears black. When this image was acquired, the rice crop was at the booting stage (approximately 10 days before the average heading day).

Figure 2.1(b) shows the same area one month after the flood (7 September) with the same color assignment. In this image, the heavily damaged rice fields near the lake appear red. This is because the damaged rice crop was mixed with the muddy floodwaters, and the reflectance in the near-infrared decreased while the reflectance in the mid-infrared band increased. The dark green area along the river and the light green area correspond to the somewhat damaged rice fields and to the undamaged rice fields, respectively. This scene was obtained almost 2 weeks before the rice crop was harvested (ripening stage). In this area, as is shown later, there were various levels of crop damage ranging from 100% to 0% yield decrease.

## 2.3 Method

The configuration of our analyses is depicted in Figure 2.2. Prior to the flood - damage analyses, several preprocessing steps were performed. First, the second scene was registered onto the first scene by Affine transformation. TM bands 2 (0.52-0.60 μm), 3 (0.63-0.69 μm), 4 (0.76-0.90 μm), and 5 (1.55-1.75 μm) of two registered scenes (250 x 240 pixels) were added to form multitemporal, 8-dimensional image data. Then, in order to extract the rice field pixels, a supervised classification was carried out using these multitemporal image data. Because the rice fields showed several different



(a)

(b)

**Fig. 2.1.** (a) TM image showing the inundated area on 6 August (immediately after the flood). bands 3, 4 and 5 shown blue, green and red respectively. The area inundated with turbid black. (b) TM image showing the flood damaged paddy fields on 7 September (1 month after the flood). Bands 3, 4 and 5 shown blue, green and red respectively. Heavily damaged paddy fields near the lake appear red and the dark green area along the river corresponds to the slightly damaged paddy fields.

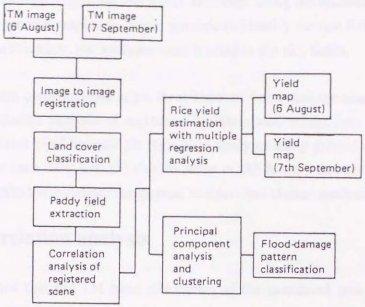



**Fig. 2.2.** Image analysis flow.

**Table 2.1.** Correlation coefficients between CCT counts of TM scenes obtained on 6 August (immediately after the flood) and 7 September (1 month after the flood).

| Scene | 6 August | | | |
|---|---|---|---|---|
| | Band | | | |
| | 2 | 3 | 4 | 5 |
| 7 September | | | | |
| Band 2 | −0·34 | −0·37 | 0·31 | 0·27 |
| 3 | 0·20 | 0·22 | −0·06 | −0·11 |
| 4 | −0·72 | −0·78 | 0·64 | 0·63 |
| 5 | 0·18 | 0·17 | 0·04 | 0·03 |

spectral features in each image, such as inundated or damaged, the rice field area could not be delineated directly from one scene. However, using the multitemporal features of the rice field spectral responses, it was possible to identify the rice field pixels correctly. After this classification, our analyses were limited to the rice fields.

Following these preprocessing steps, three kinds of flood-damage analyses were carried out: (1) correlation analysis of registered multitemporal image data in rice fields; (2) estimation of rice yield by multiple regression analyses using ground reference data and image data at each date; and (3) classification of the flood - damage relationship by a combined method of principal component analysis and cluster analysis.

## 2.4 Correlation analysis

It was expected that the TM band reflectance of the inundated area in the first scene was closely related to that of the damaged rice crop in the second scene. Correlation analyses were carried out to determine to what extent the TM band digital values of each scene were correlated statistically. Table 2.1 shows the correlation coefficients between the TM band CCT counts obtained on 6 August (immediately after the flood) and those of 7 September (before harvest). The highest negative correlation was observed between band 3 (red) on 6 August and band 4 (near infrared) on 7 September. As these bands are known to be related to the turbidity level of floodwater and the rice crop yield, respectively (Patel et al. 1985, Lathrop and Lillesand 1986), this negative correlation indicates the positive correlation between the floodwater turbidity and rice yield decrease. However it should be noted that since this correlation coefficient was calculated using CCT counts of all pixels in the rice fields, some fraction of the correlation was not directly due to this turbidity - damage relationship, but due to the spectral features of the non-inundated rice fields. The high positive correlation observed between band 4 (near infrared) in the two scenes was mainly due to the high near-infrared reflectance in the non-inundated rice fields. The detailed multitemporal spectral features of the inundated and non-inundated areas are discussed in the flood - damage pattern analysis further below.

## 2.5 Rice yield estimation

### 2.5.1 Data used

Rice yields were surveyed at seven test sites. The yields of 3 to 5 sampling plots (3.5 $m^2$) in each test site were measured on 25 September. The average yield for each test site was used as the ground reference yield, while the CCT counts for each test site were determined as the average of the corresponding 3x3 pixels.

### 2.5.2 Multiple regression analysis

Multiple regression models to estimate the decrease of rice yield were made using these ground reference yields as objective variables and CCT counts as explanatory variables.

Table 2.2 shows the results of a stepwise multiple regression analysis (Fin=Fout=2.0). Band 2 (green) and band 3 (red) were selected as significant variables in the stepwise regression analysis based on the scene obtained immediately after the flood, whereas only band 4 (near infrared) was significant in the model based on the scene obtained one month after the flood.

Considering the correlation of band 3 (red) with turbidity (Lathrop and Lillesand 1986), and of band 4 (near infrared) with biomass and yield (Patel et al.1985), this regression analysis shows the relationship between the floodwater turbidity and the rice crop damage, as mentioned in the previous chapter.

Although the number of ground reference yields was small, we judged these models reliable, for the range of measured yield data was sufficiently large and the models were used essentially as an interpolation method in each estimation. The multiple correlation for both models was greater than 0.95 and the standard estimation errors were less than 0.5 t/ha.

Figure 2.3 shows the relationship between the measured yield in each test site and the yield estimation from TM data at both dates.

**Table 2.2.** Regression analysis between yields (t/ha) and CCT counts of TM scenes obtained on 6 August (immediately after the flood) and 7 September (1 month after the flood).

| | 6 August | | 7 September | |
|---|---|---|---|---|
| Band | Coef. | *t* | Coef. | *t* |
| 2 | 0·598 | 2·7 | | |
| 3 | −0·414 | −3·5† | | |
| 4 | | | 0·093 | 8·3† |
| 5 | | | | |
| Const | −3·595 | | −5·165 | |
| *R* | 0·972 | | 0·965 | |
| S.E. | 0·432 | | 0·428 | |
| *N* = 7 | *F* = 34·3 | | *F* = 68·7 | |

† Significant at 5 per cent level.

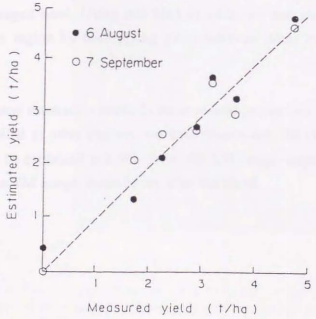



**Fig. 2.3.** Relationship between measured and estimated rice yield. Multiple regression models (table 3) were used for the estimations on 6 August (immediately after the flood) and 7 September (1 month after the flood).

### 2.5.3 Rice yield estimation map

Using these regression models, the rice yield was estimated pixel by pixel. Figures 2.4(a) and 2.4(b) show the distribution maps of the estimated rice yield using the TM data acquired on 6 August (immediately after the flood) and 7 September (one month after the flood), respectively. The estimated yields were divided into 9 levels at 0.5 t/ha yield intervals. When the results in Figure 2.4(a) and 2.4(b) were compared, some discrepancies were observed in the distribution of red pixels (I) which represent the rice fields with a yield less than 1.0 t/ha, and also in the distribution of blue pixels (A) which represent the fields with a yield of more than 4.5 t/ha. However the distribution of the medium-yield levels was similar in each map.

It seems that insufficient numbers of test sites with high and low yield levels, and saturation of band 3 (red) reflectance for water turbidity were the causes of the discrepancy observed. In addition, lodging of the rice crop which occurred after the flood in the non-inundated areas, might also be a cause of the discrepancy in the high yield distribution.

Table 2.3 is a comparison of the estimated rice field area for each yield level at 1.0 t/ha intervals between 6 August and 7 September. Estimated areas agreed well except for the heavily damaged level. Using this kind of table, we can assess the extent of the total damage in a region by multiplying yield decrease (t/ha) by corresponding estimated area (ha).

Although these estimation methods have several limitations in accuracy and cannot be directly applied to other regions, we have shown that the crop damage caused by the flood could be estimated not only from the TM image acquired before the harvest but also from the TM image immediately after the flood.

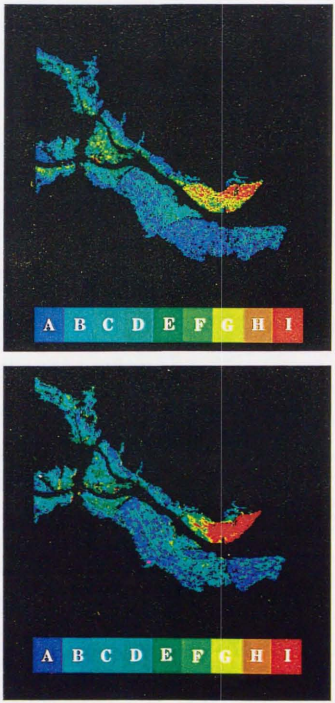



**Fig. 2.4.** (a) Distribution map of the rice yield, estimated by the multiple regression model for the scene obtained on 6 August (immediately after the flood). Letters in coloured bars represent yield levels. A $>$ 4.5; B = 4.0-4.5; C = 3.5-4.0; D = 3.0-3.5; ...; H = 1.0-1.5; I $<$ 1.0 t/ha. (b) Distribution map of rice yield, estimated by the multiple regression model for the scene obtained on 7 September (1 month after the flood). Letters in coloured bars represent yield levels. A $>$ 4.5; B = 4.0-4.5; C = 3.5-4.0; D = 3.0-3.5; ...; H = 1.0-1.5; I $<$ 1.0 t/ha.

**Table 2.3.** Comparison of estimated paddy field areas for each yield level between 6 August (immediately after the flood) and 7 September (1 month after the flood). Regression models were used to calculate the yield.

| | Area (ha) | | | | | |
|---|---|---|---|---|---|---|
| Yield level (t/ha) | 0~ | 1~ | 2~ | 3~ | 4~ | 5~ |
| 6 August | 25 | 56 | 102 | 137 | 192 | 109 |
| 7 September | 51 | 25 | 111 | 158 | 179 | 99 |

**Table 2.4.** Standardized principal component analysis using the eight-dimensional multitemporal TM images of paddy fields.

| | Eigenvector (6 August) | | | | Eigenvector (7 September) | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Principal | Band | | | | Band | | | | |
| Component | 2 | 3 | 4 | 5 | 2 | 3 | 4 | 5 | Eigenvalue |
| 1 | 0·36 | 0·34 | 0·33 | 0·33 | 0·37 | 0·37 | 0·37 | 0·37 | 7·2338 |
| 2 | −0·28 | −0·48 | 0·56 | 0·58 | −0·09 | −0·13 | 0·04 | −0·12 | 0·6508 |
| 3 | 0·25 | 0·60 | 0·27 | 0·25 | −0·32 | −0·23 | −0·49 | −0·22 | 0·0741 |
| 4 | 0·19 | 0·12 | 0·14 | −0·19 | −0·06 | −0·58 | 0·68 | −0·29 | 0·0181 |
| 5 | 0·04 | 0·05 | −0·69 | 0·66 | −0·04 | −0·24 | 0·13 | 0·09 | 0·0154 |
| 6 | −0·14 | 0·02 | 0·12 | −0·13 | −0·37 | −0·35 | 0·00 | 0·83 | 0·0057 |
| 7 | 0·75 | −0·49 | 0·00 | 0·05 | 0·19 | −0·26 | −0·29 | 0·11 | 0·0014 |
| 8 | −0·33 | 0·18 | 0·05 | −0·02 | 0·76 | −0·46 | −0·25 | 0·08 | 0·0006 |

## 2.6 Flood - damage pattern classification

### 2.6.1 Principal component analysis

According to several studies, principal component transformation of multitemporal, multispectral satellite image data is effective in enhancing regions with localized change (Richards 1984, Townshend 1984). To classify the flood - damage relationship, we also applied this method to the multitemporal TM image data on the flood and the crop damage.

Table 2.4 shows the eigenvectors and the eigenvalues of the standardized principal component analysis using the 8-dimensional, multitemporal TM image of the rice fields (CCT counts for non-rice pixels were set to zero).

The first principal component was the so called brightness component, which had almost the same positive coefficient for each band. The second component represented a difference between the inundated and non-inundated areas. Both the third and the fourth components also represented flood - damage relationships. However, their eigenvalues were remarkably small compared with that of the second component.

Still higher components appeared to relate to subtle differences in the rice fields, such as the effects of the rice varieties planted and of mixed pixels.

### 2.6.2 Clustering

In order to categorize the flood - damage pattern, cluster analyses were applied to the image composed of principal components. After testing many kinds of initial cluster and band selection, nine stable cluster classes were extracted from the 2,3,4 principal components by the migrating means method.

### 2.6.3 Flood - damage relationships

The multitemporal spectral features of each cluster class are depicted in Figure 2.5. The CCT counts for each class were determined as the average for all the pixels belonging to that class, while the estimated rice yield for each cluster class was determined in the same way. The results are shown in Figure 2.6.

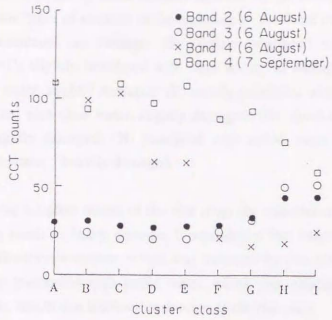



**Fig. 2.5.** The average CCT counts of TM bands 2, 3 and 4 obtained on 6 August (immediately after the flood) and TM band 4 obtained on 7 September (1 month after the flood) for each flood damage cluster class. Band 3 (6 August), band 4 (6 August) and band 4 (7 September) indicate the flood water turbidity, the rice vegetation above the flood water and the rice yield, respectively.

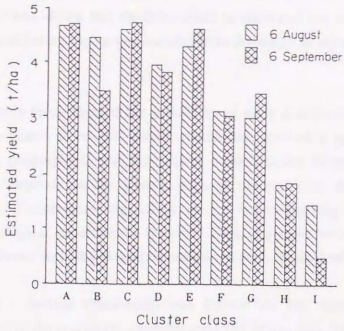



**Fig. 2.6.** Estimated average rice yield for each flood damage cluster class. Multiple regression models (table 4) were used for the estimations on 6 August (immediately after the flood) and 7 September (1 month after the flood).

To assist in interpreting these spectral features of the inundation and resulting crop damage, nine types of reaction of the rice crop to the flood inundation were identified: (A) not inundated, no damage; (B) slightly inundated with clear water, slightly damaged; (C) slightly inundated with clear water, no damage; (D) mostly inundated with clear water, slightly damaged; (E) mostly inundated with clear water, no damage; (F) inundated with clear water, slightly damaged; (G) inundated with somewhat turbid water, slightly damaged; (H) inundated with turbid water, damaged; (I) inundated with turbid water, heavily damaged.

Owing to the adaptive nature of the rice crop, the inundation with clear water did not necessarily result in heavy damage. It was shown that submergence of the rice crop under the floodwater surface, which was indicated by low band 4 (near infrared) CCT counts, and inundation with turbid water, which was indicated by high band 3 (red) CCT counts, resulted in irreversible damage to the rice crop.

## 2.7 Conclusion

TM image data were subjected to several multivariate analyses in order to assess the extent of flood damage and to analyze the relationship between the flood and the damage. Through the relationship between floodwater turbidity and the actual rice yield decrease, it was shown that the latter could be estimated not only from the TM image data acquired before the harvest but also from the image acquired immediately after the flood.

Although this result showed the possibility of early assessment of crop damage in a region, it remains to be determined whether this method is applicable to other areas where the geographical and agricultural conditions are different. The effect of the inundation depends on the growth stage of the rice crop and the duration of the inundation. Furthermore, the effect of the water turbidity is a function of the geomorphological conditions of the river. Thus, at present, estimation of crop damage from floodwater turbidity should be carried out using ground reference data.

The flood - damage relationship was defined by the clustering of the principal components of the registered, multitemporal TM image data. The CCT counts for each class showed the relationships between the features of the inundation in the rice fields and the resulting rice damage. Although this reaction pattern of the rice crop also

depends on the stage of growth of the plant when a flood occurs, the results reveal the adaptability of rice plants to flood inundation.

# Chapter 3
# Spectral Indices for Flooding Analysis

In this chapter, new spectral indices, the perpendicular vegetation index (PVI) and water turbidity index (WTI), are developed for flood inundation monitoring and vegetation change in wetland areas. These spectral indices are defined based on the analysis of rice field flooding and consequent crop damage. In the analysis, TM rice field images acquired during flooding and one month later were used to relate inundation damage to rice yield. This was accomplished by using the turbid water pixels to determine a turbid water line (TWL). The WTI along the TWL and PVI for paddy rice were defined using this line. The relationship is determined between floodwater turbidity at the rice booting stage, monitored using WTI, and the reduction in yield, measured by PVI. This chapter is based on Yamagata et al. (1988b).

## 3.1 Introduction

Landsat and other satellites have the capability to document conditions in individual rice fields and to provide a thematic overview. Several workers have estimated crop biomass and yield using satellite data (Barnett and Thompson 1982, Wiegand et al. 1979, Wiegand and Richardson 1984).Vegetation indices derived from field measurements of reflectance factors have been found to relate closely to leaf area index (LAI), phytomass, and yield of many kinds of crops (Jackson 1983, Jackson et al. 1983, Gallo et al. 1985, Wiegand and Richardson 1987). However, the indices derived from ground observed reflectance factors cannot be applied directly to satellite calibration (Price 1987).

As mentioned in Chapter 2, we were fortunate to acquire Landsat Thematic Mapper (TM) scenes that recorded inundated rice fields one day after a typhoon that struck the Kanto district on 4 and 5 August 1986. We also obtained images taken a month after the typhoon, which documented the damage (Yamagata and Akiyama 1988). We converted these data into reflectance factors using coefficients provided by Price (1987) and

defined the TWL, WTI, and the PVI (Richardson and Wiegand 1977) for inundation-damaged rice fields. TM bands 3 (630-690 nm: red) and 4 (760-900 nm: near-infrared or NIR) were used. The objective was to relate floodwater turbidity to resulting damage in terms of rice yield using WTI and PVI developed for two test sites, and to discuss how such relationships can be used in rice crop damage estimation.

## 3.2 Study Area and Image Data

We analyzed two test sites (Figure 3.1) in northern Kanto district in Japan, where rice crops were damaged by flooding from the typhoon. Both test sites were located along rivers near their confluence with lakes.

Figure 3.2 displays TM Band 3 images showing the inundated paddy fields in the Hinuma (a) and Ishioka (b) areas on 6 August 1986, immediately after the typhoon abated. In these images the more turbid the floodwater, the brighter the images appear. The image was acquired approximately 10 days prior to heading and the inundation continued for 1-5 days.

Figures 3.3(a) and (b) display the TM Band 4 (near-infrared) images, corresponding to those in Fig. 3.2, obtained on 7 September 1986, one month after the inundation. In these images, healthy rice appears bright and damaged rice appears dark. Heavily damaged rice was dead and abandoned (not harvested). When this scene was acquired, the rice crop was in the middle of the grain filling stage.

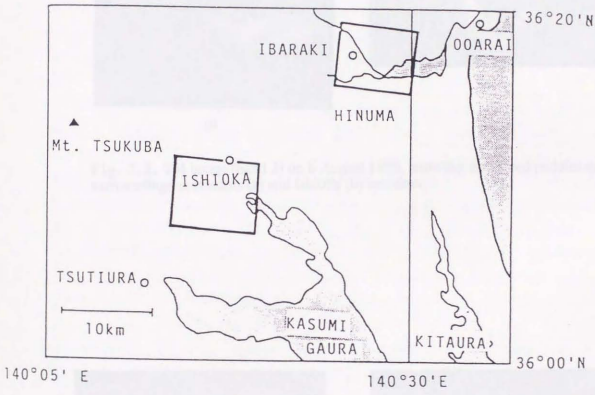
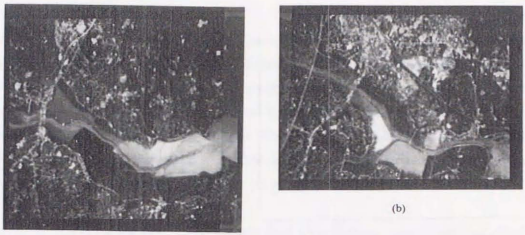



**Fig. 3.1.** Test site location map.



(a) (b)

**Fig. 3.2.** TM image (band 3) on 6 August 1986, showing inundated paddies and surroundings at Hinuma (a) and Ishioka (b) test sites.

(a) (b)

**Fig. 3.3.** TM image (band 4) on 7 September 1986, showing damaged paddies and surroundings at Hinuma (a) and Ishioka (b) test sites.

## 3.3 Method

In order to examine the correspondence between inundation and damage, the TM scenes for 6 August and 7 September were overlaid. TM Bands 3 (630-690 nm) and 4 (760-900 nm) were selected for the analyses, because ergonomically important parameters, such as LAI, can be measured by these red and near-infrared bands (Wiegand et al. 1979). The two test sites selected, Hinuma and Ishioka, were 250200 and 250160 pixels in size, respectively. The two-channel data for each test site formed a 4-channel, multitemporal, multispectral data set for each site.

To extract the pixels for the rice fields, a supervised classification was carried out using these multitemporal image data. Within the rice field pixels, there were several distinguishable spectral categories that ranged from undamaged (category A) to abandoned (category G) that could not be distinguished without using the multitemporal features. After this classification of damage categories, we subsampled rice field pixels taking every fifth pixel on every fifth line to reduce the data volume for statistical analyses (pixels that were not classified as rice were skipped). Representative pixels from the rice categories such as inundated or not inundated in the Hinuma scene are identified by letters in Figures. 3.6, 3.8, and 3.9 which show the results.

As detailed in Chapter 2, the actual rice yield was determined at seven test sites in the Hinuma area.. The average yield for each test site was used as the ground truth, while the TM reflectance factors for each test site were determined from the average for the 3x3 array of pixels centered over those same sites.

In addition, the relationship between PVI (based on the TWL) and yield from the TM and damage survey data was compared with the PVI and rice yield data from the experimental plots of rice grown under variable fertilization in 1987. In the experiment, spectral reflectances were measured at 660 and 840 nm with a 10 nm bandwidth using the spectroradiometer described by Shibayama and Munakata (1986). Measurements were conducted on six plots of the variety Koshihikari (the most common rice variety in the damaged area) during the middle of the grain filling stage, and the yield of each plot was also determined. The spectral-agronomic relationships found in this experiment have been reported elsewhere (Shibayama et al. 1988).

TM digital counts observed at the outer edge of the atmosphere are affected by the atmosphere itself, sun elevation, and sensor degradation. Therefore, a spectral index developed using a given scene cannot be directly applied to another scene or to other sensor systems without calibration to reflectance factors. We used the calibration method and coefficients summarized by Price (1987) to convert Landsat-5 TM digital counts into a spectral albedo, defined as the equivalent solar radiance calculated for each channel. They are comparable to spectral reflectance factors measured on the ground for the same wavelengths. Hereafter in this paper, we call them spectral reflectance.

## 3.4 Results and Discussion

### 3.4.1 Scatter Diagram

Figure 3.4 displays the scatter in red and near-infrared reflectance space of inundated rice field pixels in the test area on 6 August 1986, one day after the typhoon ended.

### 3.4.2 Turbid water line

As is the case of the soil line (Richardson and Wiegand 1977), turbid water pixels fell on a line. Using the 12 water pixels of various turbidity levels (sediment load), we determined the equation of TWL (Fig. 3.5) by regression analysis. Almost all water pixels are close to this TWL line, which passes through the origin. We needed to use this TWL instead of the soil line because rice grows out of turbid water rather than soil. We used the TWL as a base to define PVI and WTI.

### 3.4.3 WTI and PVI

A linear relation between water turbidity and spectral radiance in the red band has been reported (Khorram 1981, Lathrop and Lillesand 1986). We defined WTI as the distance from the origin along the TWL to measure the turbidity (sediment load) of the floodwater. This WTI expressed in terms of spectral reflectance should hold universally, if properly placed in TM 3 and 4 band space.

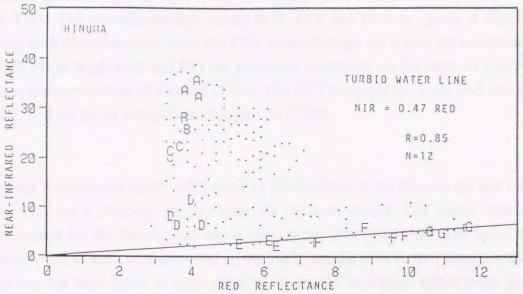



**Fig. 3.4.** Scatter diagram of red band (TM3) and near-infrared band (TM4) of paddy field pixels in Hinuma on 6 August 1986. Several ground-truthed paddies were identified with letters as follows: (A, B, C) paddies not inundated but depth of water (clear) increases from A to C; (D) canopy partially inundated with clear water; (E, F, G) paddy inundated with increasingly turbid water, respectively. The turbid water line (TWL) was determined using categories E, F and G.

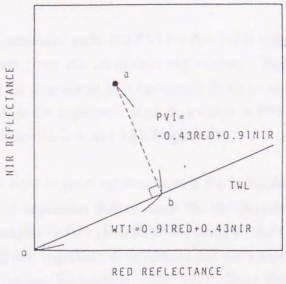



**Fig. 3.5.** Illustration of the water turbidity index (WTI) and perpendicular vegetation index (PVI). Distance from o (origin) to b is the WTI and from b to a is PVI. Defining equations for WTI and PVI are given.

Richardson and Wiegand (1977) developed the PVI in order to extract information about vegetation independently of the soil background. Here PVI is the distance from the TWL. The interrelationships among WTI, PVI, and TWL are given in Figure 3.5 along with their equations. Since the TWL passes through the origin, the coefficients in the WTI (or brightness) and PVI (or greenness) equations are the same as yielded by the *n*-space procedure of Jackson (1983). The WTI is equivalent to the soil line index (SLI) defined for by Wiegand and Richardson (1982).

Figures 3.6(a) and (b) relate WTI and PVI of rice fields in the Hinuma (a) and Ishioka (b) areas on 6 August, the day after the typhoon abated. The turbid water line determined for the Hinuma site was applied to the Ishioka test site. The turbid water pixels in the Ishioka area were also on the TWL calculated from the Hinuma data. The soil line has been found to apply globally, so that the analogous turbid water line for grayish and brownish soils in the Hinuma data should also be applicable elsewhere.

### 3.4.4 PVI and Rice Yield

We applied the PVI equation based on the TWL for the 6 August scene to TM data acquired on 7 September 1986, one month later, to see how much the rice was damaged. We also related the PVI observations to yield. Data sets from the TM scenes and an on-ground experiment (Shibayama et al. 1988) were pooled.

Figure 3.7 relates measured yield and PVI for rice fields using ground-truth data from the TM scenes and from the on-ground experiment. Yield in the inundated area ranged from zero for abandoned (not harvested) fields to almost 5 t/ha for those that were not damaged. In the experimental plots, a range in PVI was achieved by varying the amount of fertilizer (0, 2, 4, and 6 g nitrogen/$m^2$).

The two data sets were in good agreement with the principles of spectral components analysis (SCA): (a) vegetation indices sense the photosynthetic size of canopies; (b) stresses severe enough to affect yield affect the canopy; and (c) the photosynthetic size of canopies during late vegetative development and early reproductive stages and yield are related (Wiegand and Richardson, 1984, 1987). Both data were used to determine the regression equation given in the Figure (R=0.91, N=13).

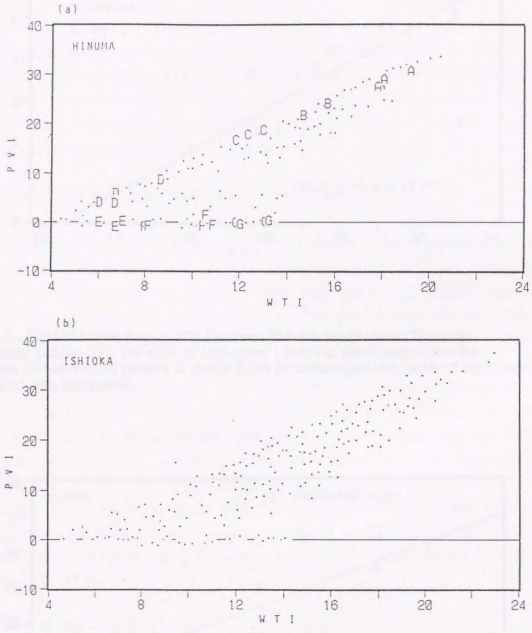



**Fig. 3.6.** Water turbidity index (WTI) vs. perpendicular vegetation index PVI for 6 August 1986, TM scene at two test sites, Hinuma (a) and Ishioka (b). The turbid water line determined for Hinuma site was applied to Ishioka test site. the categories were: (A, B, C) paddies not inundated but depth of water (clear) increases from A to C; (D) canopy partially inundated with clear water; (E, F, G) paddy inundated with increasingly turbid water respectively.

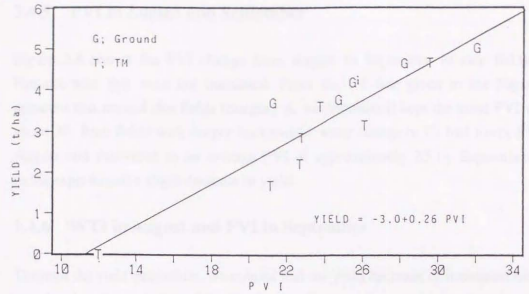



**Fig. 3.7.** Yield of brown rice vs. PVI for seven TM data sample sites (T) and six experimental paddies (G). TM scene (7 September 1986) site yields ranged zero for abandoned (not harvested) paddies to almost 5 t/ha for undamaged site; yields of experimental plots varied with fertilization.

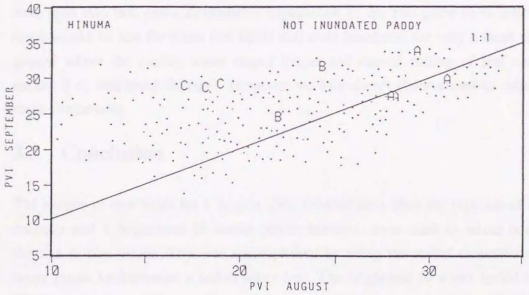



**Fig. 3.8.** PVI (7 September) vs. PVI (6 August) of paddies not inundated (most of the canopy was above water) in Hinuma area. The categories were (A, B, C) paddies not inundated but depth of water (clear) increases from A to C.

### 3.4.5 PVI in August and September

Figure 3.8 shows the PVI change from August to September of rice fields in the Hinuma area that were not inundated. From the 1:1 line given in the Figure, it is apparent that normal rice fields (category A, not inundated) kept the same PVI value of about 30. Rice fields with deeper background water (category C) had lower PVI on 6 August and recovered to an average PVI of approximately 25 by September. These fields experienced a slight decrease in yield.

### 3.4.6 WTI in August and PVI in September

Through the yield estimation, we noticed that the yield decrease of inundated rice could be related to the turbidity of the floodwater. Because damage to the rice plants was at least partially due to mud sticking to the leaves and stems, it is reasonable that the more turbid the floodwater, the heavier may be the damage. We monitored the floodwater turbidity by WTI on 6 August and the yield by PVI on 7 September. Figure 3.9 shows the relationship between WTI and PVI in the Hinuma (a) and Ishioka (b) areas with regression lines. For the Hinuma area, PVI (yield) in September was unaffected until the turbidity (WTI) in August exceeded 9, beyond which the PVI decreased linearly as WTI increased. In contrast, data for the Ishioka area indicate a linear decrease in PVI in September until the turbidity (WTI) reached 9-10, beyond which the relationship may have split into two paths. A plausible explanation of the two paths is as follows: the slope would be less for those rice fields that were inundated for only a short time and greater where the muddy water stayed longer and caused rotting of leaf and other tissues (i.e., increased damage). However, we lack direct observations to substantiate these conjectures.

## 3.5 Conclusion

TM images of rice fields for 6 August 1986 (immediately after the typhoon of 4 and 5 August) and 7 September (3 weeks before harvest), were used to relate inundation damage to rice yields. This was accomplished by using the turbid or sediment laden water pixels to determine a turbid water line. The brightness or water turbidity index (WTI) along the turbid water line and the perpendicular vegetation index (PVI) for the rice were defined using this line. Rice yield data for undamaged and damaged rice fields and in experimental rice fields were pooled and related to the PVI calculated

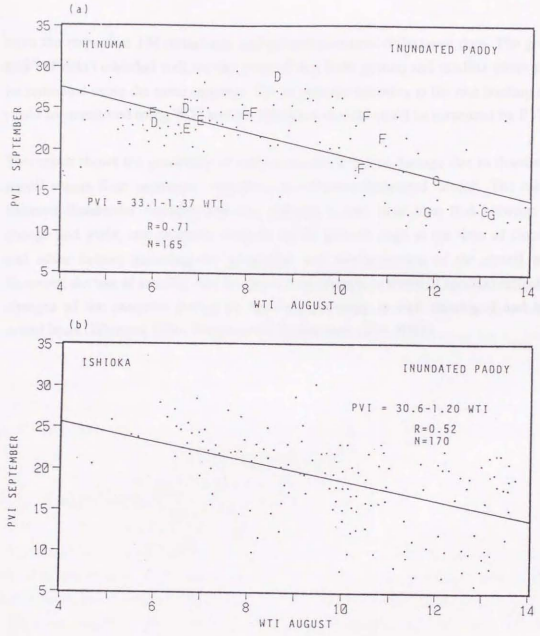



**Fig. 3.9.** PVI (7 September) vs. WTI (6 August) for paddies inundated with clear or turbid water in Hinuma (a) and Ishioka (b) areas. Regression lines for each test site are given. The data show that the more turbid the flood water in the 6 August scene, the heavier the damage to the inundated paddy rice. However the relation is not liner nor clearly determined. The categories were; (D) canopy partially inundated with clear water; (E, F, G) paddy inundated with increasingly turbid water, respectively.

from the respective TM reflectance and ground-measured reflectance data. The ground and TM data coincided well, so that yield of rice from ground and satellite views could be estimated using the same equation. The floodwater turbidity at the rice booting stage could be monitored using WTI and the reduction in yield could be measured by PVI.

This result shows the possibility of early assessment of rice damage due to flooding by runoff waters from monsoon-, typhoon-, or hurricane-associated rainfall. The relation between floodwater turbidity and rice damage is less clear than that between PVI change and yield, and probably depends on the growth stage at the time of flooding, and other factors including the geography and characteristics of the runoff areas. However, the use of satellite data to assess crop damage in terms of spectral reflectance changes of the canopies during the reproductive stage is well developed and has a sound basis (Wiegand 1984; Wiegand and Richardson 1984, 1987).

# Chapter 4
# Spectral Indices for Vegetation, Soil and Water

In this Chapter, we propose a new index to monitor wetland conditions called the Vegetation-Soil-Water Index (VSWI). The PVI is further developed to the VSWI; the VSWI can monitor vegetation, soil, and water conditions at the same time. Algorithms to determine automatically the end-member points for vegetation, soil, and water are developed by fitting a triangle to the scatter plot instead of finding the soil line. The distances between the spectrum points and the triangle edges are used as the SWI. In conventional unmixing approaches, end-member points are often determined manually and arbitrarily from the image data or from scatter plots of the data. A new algorithm that can automatically determine the end-member points has been developed. The VSWI is applied to wetland monitoring using multitemporal Landsat TM image data, and vegetation, soil, and water conditions and changes have been successfully delineated. This chapter is based on Yamagata et al. (1997a).

## 4.1 Introduction

Vegetation biomass is one of the most important parameters that can be estimated using remotely sensed data. There have been many vegetation indices developed using the red and near-infrared bands, such as the Ratio Vegetation Index (RVI), Normalized Difference Vegetation Index (NDVI), and Perpendicular Vegetation Index (PVI). The relationships between these indices and LAI and biomass have been clarified (Jackson 1983).

PVI is based on the idea of the soil line in scatter plots. It is known that the scatter plots of soils with different moisture content in red-near axis lie on a line called the soil line. Figure 4.1 shows the relationship between the soil line and PVI. On the soil line, the spectra move as the soil dries. With increasing vegetation in the target, the spectra move toward the upper left. When there is little vegetation and the background soil is appearing, the reflecting spectra move parallel to the soil line, reflecting the amount of water in the soil. This trajectory of the spectral is called the iso-vegetation line. PVI is

an index designed to suppress the effect of background soil (soil noise) to estimate consistently the amounts of vegetation by measuring the perpendicular distance of the spectra from the soil line. Figure 4.1 shows PVI as this distance.

Some modifications of the PVI have already been developed. First, the Soil Adjusted Vegetation Index (SAVI) was devised by Heute (1988) and included the nonlinear interaction of soil and vegetation reflectance. The Transformed Soil Adjusted Vegetation Index (TSAVI) was proposed by Baret and Guyot (1991), and the Modified Soil Adjusted Vegetation Index (MSAVI) by Qi et al. (1994). As described in Chapter 3, PVI was extended by Yamagata et al. (1988) to the monitoring of rice fields using the WTL, which corresponds to differences in suspended sediment in rice field water.

These modified PVIs were mainly developed for use in agricultural or forestry monitoring. However, for monitoring natural ecosystems using remote sensing, we need to monitor environmental variables as well as vegetation parameters. Particularly in wetland monitoring, we need to monitor concurrently the vegetation, soil, and water, which constitute complicated wetland ecosystems.

Unmixing is the most popular approach to estimate from spectral information the factors that constitute land cover (Inamura 1987, Ito and Fujimura 1987). In general unmixing, it is assumed that the mixel spectra can be described as linear combinations of the constituent spectra (linear mixture model), and the proportions of the constituents are estimated using statistical inference. In a case when the components in a pixel are distributed separately, the proportions can be correctly estimated using this unmixing method. However, the mixels obtained over a natural vegetation area often consist of complicated mixtures of soil and vegetation that overlap. The basic assumption of the linear unmixing approach does not always hold.

In this study, as a first step to quantitative monitoring of ground cover with complicated vegetation, soil, and water mixtures, we further develop the PVI into the VSWI, which calculates the spectral index both for the soil and water and the vegetation. The vegetation index in the PVI is used as a quantitative index related to the vegetation parameters, while the soil and water index in the VSWI can be used as the quantitative index for the soil and water parameters.

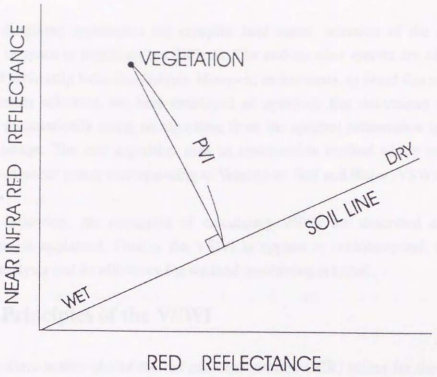



**Fig. 4.1.** Relationship between PVI and SOIL LINE on Red-NIR wavelength space.

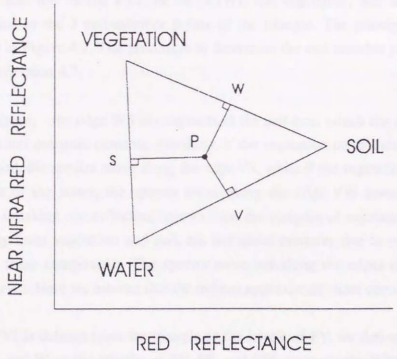



**Fig. 4.2.** Relationship between VSW indices and end-member triangle on Red-NIR scatter plot.

In the unmixing approaches for complex land cover, selection of the end-member spectra in space is important but difficult. The end-member spectra are often given or selected arbitrarily from the analysis. However, in this study, to avoid this uncertainty in end-member selection, we have employed an approach that determines end-member spectra automatically using an algorithm from the spectral information in a remotely sensed image. The new algorithm uses an optimization method which can determine the end-member points corresponding to Vegetation, Soil and Water (VSW).

In the following, the principles of calculating VSWI are described and the new algorithm is explained. Finally, the VSWI is applied to multitemporal, Landsat TM image analysis and its efficiency for wetland monitoring is tested.

## 4.2 Principles of the VSWI

If we make a scatter plot of the red and near-infrared (NIR) values for the pixels in an image, they should form a triangle, which has fairly distinct boundaries in the scatter plot. The straight line that best matches the lower right side of the scatter plot (red on the x and NIR on the y axis) is the soil line. The distance between the spectral points and the soil line is the PVI. In the VSWI, the vegetation, soil and water indices correspond to the 3 end-member points of the triangle. The principles of VSWI are depicted in Figure 4.2. The procedure to determine the end-member points is described below in Section 4.3.

In the Figure, the edge WS corresponds to the soil line, which fits the soil spectra at different soil moisture contents. Similarly, if the vegetation cover increases over that of the dry soil, the spectra move along the edge VS, while if the vegetation cover increases over that of the water, the spectra move along the edge VW towards the vertex V. (Strictly speaking, the reflecting spectra from the complex of vegetation and water, and especially from vegetation and soil, are not linear mixtures due to multiple scattering between these components. The spectra move not along the edges of the triangle but along curves. Here we assume that the indices approximate these curves.)

As the PVI is defined from the triangle as the length of PV, we define the index values for V, S, and W as the lengths of PV, PS, and PW, respectively. When the spectra are located outside the triangle, some of the VSWI give negative values.

## 4.3 Algorithm for End-member Points

Ideally end-member spectra for calculating VSWI should be determined by spectral measurements of pure targets beforehand. However, because the spectra observed by the satellite sensors are easily changed due to radiation, and atmospheric and vegetation conditions, etc., it is very difficult to use measured end-member spectra for analysis of remotely sensed image data. Also, there is no pure vegetation spectrum as it changes with species and season.

End-member spectra have often been determined manually from the image or scatter plots of the spectral data. However, this selection process is arbitrary and has much ambiguity; there is no guarantee of its reproducibility and it is not suitable for quantitative analysis.

### 4.3.1 Red-NIR Scatter Function

Let RED(i,j) and NIR(i,j) be the digital values of the red and near-infrared bands of the image data, respectively. Here, i and j are the coordinates of the image and RED and NIR take integer values between 0 and 255, because the TM image has an 8-bit dynamic range.

If we make a scatter plot by putting RED on the x-axis and NIR on the y-axis for pixels in the image, the scatter plot becomes grid data containing frequency distributions for each grid coordinate. We will call this distribution H(x,y), where x and y stand for red and near-infrared intensity, respectively.

H(x,y) often has a triangular shape and most of the pixel values are inside this triangle, but some of the points lie outside it. We can eliminate these points by making a threshold to eliminate low-frequency points. By increasing the threshold value gradually, we can determine the a value that eliminates 5% of the pixels. If the frequency is lower than this threshold value, the data at that point are eliminated.

Then we calculate the convex structure of the distribution H(x,y), and set the value inside the structure to 1 and outside to 0. This new binary value distribution: C(x,y) often has a triangular shape. In general, the structure is a polygon with a minimum edge length that contains all points of interest.

### 4.3.2 Estimation of Initial Values

Considering the characteristics of red-NIR reflection of V, S, and W, we can assume that the end-member points corresponding to these, Ev, Es, and Ew, have the following properties:

a) Ev is located distant from the soil line in the upper left position.
b) Es is around the maximum value for red.
c) Ew is around the minimum value for NIR.

By checking the above conditions in the distribution of C(x,y), we can determine the initial values for the end-member points.

### 4.3.3 End-member Point Determination

We consider another binary triangle distribution T(x,y) which has a value of 1 inside and 0 outside the triangle determined by the estimated end-member points. To determine the T(x,y) so as to approximate the C(x,y), we first define the areas S1 and S2 as follows.

$$S1 = \{\text{area of } C(x,y) == 1 \text{ and } T(x,y) == 0\}$$
$$S2 = \{\text{area of } C(x,y) == 0 \text{ and } T(x,y) == 1\}$$

The S1 is the area inside the convex polygon distribution C(x,y) and outside the triangle T(x,y), whereas S2 is the area outside C(x,y) and inside T(x,y). If the distributions T(x,y) and C(x,y) coincide, both S1 and S2 are zero. The amount and type of discrepancies are known from values of S1 and S2. By integrating S1 and S2 as

$$F=S1^2+S2^2$$

we can assess the matching of the T(x,y) and C(x,y) by the value of F.

Since F is a function of Ev, Es, and Ev, we can determine the optimal location of the end-member points by minimizing F. We use a non-linear optimization algorithm (the Simplex method) to search for optimal end-member points that minimize F, starting

from the initial estimate points.

## 4.4 Application to Multitemporal Landsat TM Data

In order to test the effectiveness of VSWI, we applied this index to monitoring wetlands. In the analysis, multitemporal Landsat TM scenes acquired on 3 dates in 1986 and 1991 over the Kushiro wetland were used.

### 4.4.1 Study Area and Image Data

The study site, the Kushiro wetland, is the largest in Japan with 18000 ha. Most of the wetland is fen, covered with reeds and sedge grasses, while some parts are sphagnum moss-covered bog areas.

The image data used for the analysis are from Landsat TM scenes acquired on 28 June, 9 September, and 27 October in 1986 and 26 June, 29 August, and 1 November in 1991 as shown in Figure 4.3. All images in the figure are displayed by assigning false colors (red for the near-infrared band, green for the mid-infrared band, and blue for the red band).

In this figure, red means vegetation, blue corresponds to bare land or urban areas, and black represents water bodies. In June, the wetland vegetation was at the beginning of its growing season: sedge was already growing but reeds were still brown; alder trees were developing leaves. In August, the vegetation in the wetlands was at its maximum growth: the reeds had grown to around 3m in height. In early November, most of the vegetation except the sphagnum moss was dead.

The characteristics of the Landsat TM data are depicted in Table 4.1. The ground resolution of the TM sensor is 30 m except the thermal band, which is considered to be sufficient for monitoring wetland vegetation distribution and land use in the surrounding region. The image analyzed has 1024 x 1024 pixels which includes almost all the wetland area. The TM sensor has 7 bands, among which especially the near-infrared (TM4) and the mid-infrared band (TM5,6) are known to be effective for discriminating among wetland vegetation types (Yamagata et al. 1995).

### 4.4.2 Assessment of VSW Index Color Composite

The pixel values are plotted on the red-NIR axes in Figure 4.4. The 6 scatter plots correspond to the 6 TM scenes in Figure 4.3. Each scatter plot has the same scale, and the density in the plots is depicted in log scale.

The VSW end-member points determined by the algorithm described are also shown in Figure 4.4. The Figure shows that the end-member points were consistently extracted by this algorithm as the vertex points of the triangle, irrespective of the seasonal changes in the wetland area. However, the size of the triangle changed in the autumn scenes because the vegetation end-member point was not distinct.

The equations of the lines that correspond to the three edges of the triangle were calculated and formulas were obtained for measuring distances between the spectra and the lines; from these distances, the VSWIs were calculated.

Calculated VSWIs were scaled at 0- 255 (integer values) and form a color image for each scene. Figure 4.5 shows the VSWI color maps. In this figure, V, S, and W indices are displayed as green, red, and blue, respectively. The maximum values of the colors in this figure correspond to the values of the VSW Indices shown in Table 4.2. The index is allowed take a negative value, but here all negative values were set to 0 for convenience. By using this Table, we can read the value of the VSW indices of the image pixels. From the VSWI color maps, the change in wetland condition can be seen.

## 4.5 Discussion on VSWI

One of the advantages of VSWI is that it allows easy interpretation of surface conditions using a color composite. It is not easy for non-remote sensing researchers to directly interpret raw false color TM scenes. We can actually see the seasonal difference and 5-year change from the color composite of the VSWIs (Figure 4.5).

However, VSWI is an index value just like the PVI. It has little meaning without being calibrated against some physical parameters such as biomass, LAI, soil moisture, water content, etc., by on-ground measurements. VSWI can be used as a quantitative measure for environmental monitoring, once these calibrations are conducted on a reflectance basis.

**Table 4.1.** Characteristics of Landsat TM bands.

| BAND | ($\mu$ m) | | IFOV(m) |
|---|---|---|---|
| 1 | 0.45–0.52 | Green | 30x30 |
| 2 | 0.52–0.60 | Green | 30x30 |
| 3 | 0.63–0.69 | Red | 30x30 |
| 4 | 0.76–0.90 | Near IR | 30x30 |
| 5 | 1.55–1.75 | Mid IR | 30x30 |
| 7 | 2.08–2.35 | Mid IR | 30x30 |
| 6 | 10.4–12.5 | Thermal | 120x120 |

**Table 4.2.** Minimum and maximum values of VSW indices for R, G, B level of Fig. 4.6.

| Scene | Vmax | Smax | Hmax |
|---|---|---|---|
| 28/06/86 | 137.78 | 69.64 | 98.70 |
| 09/09/86 | 113.28 | 42.91 | 56.21 |
| 27/10/86 | 61.47 | 48.12 | 87.85 |
| 26/06/91 | 133.35 | 73.99 | 105.49 |
| 29/08/91 | 131.32 | 69.03 | 93.79 |
| 01/11/91 | 72.75 | 34.77 | 49.23 |

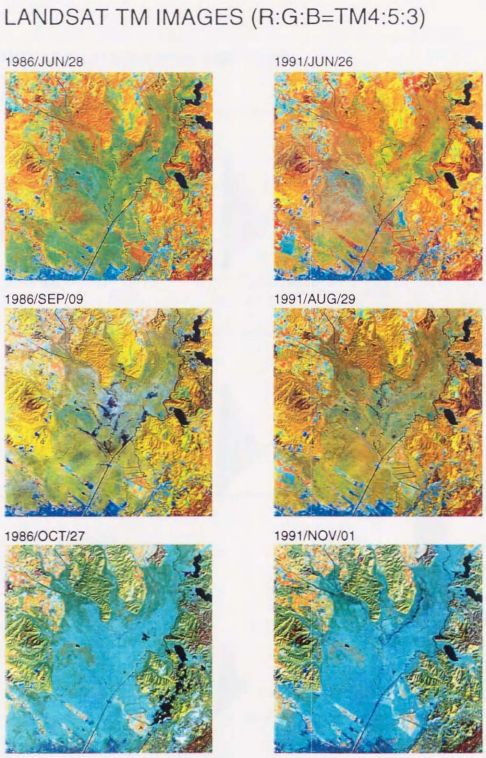



**Fig. 4.3.** Landsat TM false color images of Kushiro wetland. (R, G, B = TM4, TM5, TM3)

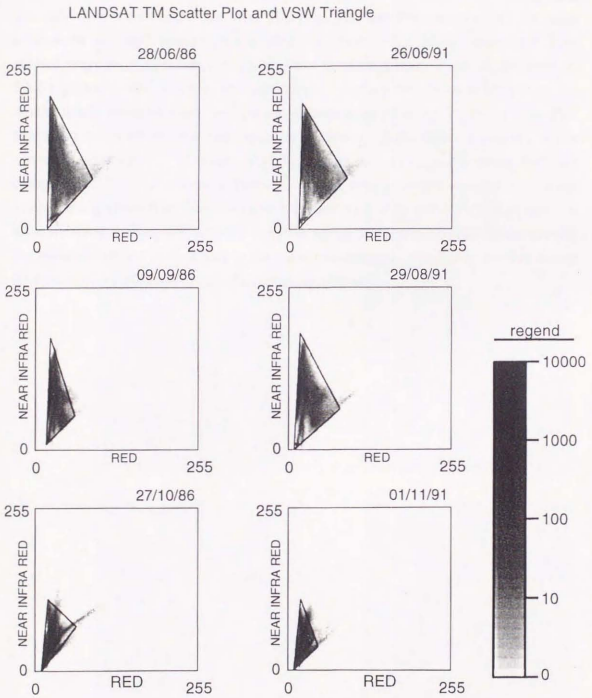



**Fig. 4.4.** Scatter graphs of TM images on Red-NIR space overlaid with the determined VSW end-member points. The each scatter graph correspond to the scene in Fig. 4.3 respectively.

Moreover, remotely sensed data are not free from atmospheric effects, sensor degradation, etc. If we need to compare data acquired at different times, atmospheric and radiometric corrections are vital. However, the data that are necessary for these corrections are unfortunately not available in most cases. Many researchers have studied ways to compensate for this problem by some empirical calibration method. VSWI performs this function to some extent, because the end-member points are automatically selected from the spectral signature distributions in the scenes. This automatic end-member selection works as a kind of standardization process of the observed pixel values. Although VSWI takes on another meaning when there are completely different end-member points in a scene, even in scenes acquired in different years the algorithm will find the same end-member points and VSWI will have the same meaning without any further calibration whenever most of the land cover remains the same and only a limited part in the scene has changed. This is an another strong advantage of VSWI in detection of environmental change.

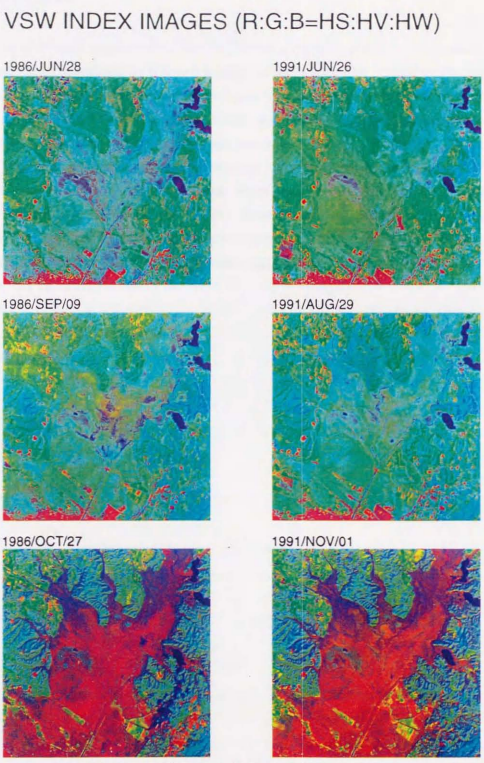



Fig. 6 Distribution of VSW indices. (R, G, B=Soil, Vegetation, Water index)

**Fig. 4.5.** Distribution of VSW indices. (R, G, B = Soil, Vegetation, Water index)

## 4.6 Conclusion

Algorithms to determine automatically the end-member spectral points for vegetation, soil, and water were developed by fitting a triangle to the scatter plot instead of finding the soil line. The distances between the spectral points and the triangle edges were used as the new index, VSWI. In conventional unmixing approaches, end-member points are often determined manually and arbitrarily from the image data or a scatter plot of the data. The VSWI was applied to wetland monitoring using multitemporal, Landsat TM image data, and the vegetation, soil, and water conditions and changes were successfully delineated. Because the VSWI determines the end-member points automatically, it also works as a standardization method for spectral data, which facilitates use of VSWI for detection of environmental change.

# Chapter 5
# Unmixing Spectral Image Data

In this chapter, a new approach for spectral unmixing by the subspace method is proposed and tested using hyperspectral image data. By the subspace method, unmixing is calculated as the projection of each unknown pixel vector onto the subspace of each class. This method is more stable against noise in the data than are conventional methods and it works effectively as a feature-extraction and data-reduction procedure as well. The performance of this method is tested by an unmixing experiment using a hyperspectral Compact Airborne Spectral Imager (CASI) image acquired over the Kushiro wetland in NE Japan. Unmixing for 7 wetland vegetation classes is also conducted by the least squares, quadratic programming, and orthogonal subspace projection methods. Finally, the results of the unmixing experiment are compared and evaluated with regard to wetland vegetation monitoring. This chapter is based on Yamagata (1996c).

## 5.1 Introduction

In wetland landscapes, various vegetation types are continuously distributed. Remotely sensed spectral data over wetland areas are spectral mixtures of several vegetation types. These images consist of mixels that have to be analyzed using spectral unmixing procedures to estimate the state of each of the constituents (Settle and Drake 1993).

Conventional statistical unmixing methods such as least squares use a linear mixing model. In this linear model, the mixed spectral vector is assumed to be a sum of the class spectral vectors that constitute the mixel. By solving this linear mixing model using predetermined class vectors, we can estimate the proportion of each class within the pixel.

However, the computational complexity increases substantially as the number of image channels increases, and the least squares solution becomes unstable due to the high

autocorrelation between the channels. In conventional approaches, it is necessary to reduce the number of spectral dimensions in the problem as a preprocessing step for unmixing (Malinowski 1991).

Hyperspectral sensors are a recent development in remote sensing and have been used for environmental monitoring (Kramer 1992). Hyperspectral imaging is recognized as an effective means for estimating vegetation parameters (Gong et al. 1994). However, to unmix the very large number of channels in hyperspectral imagery, an algorithm that can unmix several spectral classes in a fast and stable manner must be established. A number of unmixing methods, which incorporate modern signal processing and neural network methodologies, have been explored recently (Harsanyi and Chang 1994, Benediktsson et al. 1995).

Unmixing by the subspace method (Oja 1984) utilized in this paper is a new approach, based on a fundamentally different principle for pattern recognition. The subspace method first assigns a specific subspace in the high-dimensional spectral space to each vegetation class, instead of fitting a mixel model with a predetermined number of spectral dimensions. Unmixing is then performed by measuring the projection length of the mixel vector onto the subspace of each class. In addition, the subspace method has the remarkable characteristic that it unifies the process of feature extraction and unmixing, which are usually separate processes in conventional methods.

In this Chapter, the principles of the new unmixing approach by the subspace method are explained, along with experimental results derived from CASI data to compare this new method with conventional approaches.

## 5.2 Unmixing by the Subspace Method

### 5.2.1 Statistical Unmixing Methods

Conventional statistical unmixing methods assume that the mixel spectral vector is a weighted mean of the class spectral vectors that constitute the mixel. Within each mixel, there are several mixed classes whose proportions correspond to the weights of the model. These weights are estimated by the unmixing method. In a remotely sensed image with $p$ channels, $K$ land cover classes exist in the image and the proportion of class $\omega_i$ is $f_i$. A linear mixel model assumes that the observed $p$ dimensional vector

**r** is expressed as

$$\mathbf{r} = \mathbf{Mf} + \mathbf{n} = \sum_{i=1}^{p} f_i \mathbf{m}_i + \mathbf{n} \qquad (5.1)$$

Where $\boldsymbol{M}$ is a $p \times p$ matrix with class spectral vector $\mathbf{m}_i$ as column vector, $\mathbf{f}$ is a vector with $f_i$ as components, and $\mathbf{n}$ stands for the noise vector.

Statistical unmixing methods include unmixing by least squares, factor analysis and singular value decomposition (Malinowski 1991, Settle and Drake 1993). Unmixing by the subspace method does not assume a linear statistical model.

### 5.2.2 Principles of the Subspace Method

The basic idea behind the subspace method is that the class spectral vector lies mainly in a small class-specific subspace instead of within the entire dimension of the spectral space. If the class subspace is determined from the training sample of each class, class member values can be calculated by the projection of the mixel spectral vector onto the class subspaces that is determined from training samples (Watanabe 1969, Kohonen 1977).

The 3 ways of calculating subspace by the subspace method are the algebraic, statistical, and learning subspace methods (Oja 1984). In this paper, a statistical subspace method called the CLAFIC (CLAss-Featuring Information Compression) algorithm is used. This method is known to be fast and effective in cases where the volume of training data is moderate.

### 5.2.3 Enhanced CLAFIC Method

The CLAFIC algorithm determines the class subspace in order to maximize the projection of the class vector on the corresponding class subspace. However, by maximizing the projections for all classes at the same time, the separation between similar classes decreases.

In order to avoid this drawback, we have employed the enhanced CLAFIC algorithm, which maximizes the projection on the class subspace to which the training vector belongs and also minimizes projection on the other subspaces at the same time.

In the enhanced CLAFIC method, the class subspace $\mathbf{L}^i$ which corresponds to land cover classes $\boldsymbol{\omega}^{(i)}$ $(i=1,...,K)$ is determined so as to maximize the expected projection of vector $x$ that belongs to the classes $\boldsymbol{\omega}^{(i)}$. It also minimizes the expected projection of vector $x$ that belongs to other classes $\boldsymbol{\omega}^{(j)}$ $(j \neq i)$. The problem here is to determine the subspace $\mathbf{L}^i$ to satisfy these conditions at the same time as formulating the next minimization problem:

$$\sum_{\substack{j=1 \\ j \neq i}}^{K} E(x^t P^{(i)} x | x \in \omega^{(j)}) - E(x^t P^{(i)} x | x \in \omega^{(i)}) \tag{5.2}$$

where $P^{(i)}$ is the projection matrix to the $\mathbf{L}^i$.

The first term of equation (5.2) is the expected projection of sample vectors that do not belong to class $\boldsymbol{\omega}^{(i)}$ and the second term is the expected projection of vectors that do belong to class $\boldsymbol{\omega}^{(i)}$. Using expression (5.2), we can determine the subspace $\mathbf{L}^i$ that minimizes the first term of (5.2) and maximizes the second term.

The projection matrix $P^{(i)}$ is expressed using orthogonal normal bases $u_1^{(i)},...,u_{p^{(i)}}^{(i)}$ of subspace $\mathbf{L}^i$ as

$$P^{(i)} = \sum_{k=1}^{p^{(i)}} u_k^{(i)} u_k^{(i)t} \tag{5.3}$$

By substituting equation (5.3) into (5.2) and rewriting (5.2) using base vector $u_k^{(i)}(k=1,...,p^{(i)})$, we get

$$\sum_{\substack{j=1 \\ j \neq i}}^{K} \sum_{k=1}^{p^{(i)}} E((x^t u_k^{(i)})^2 | x \in \omega^{(j)}) - \sum_{k=1}^{p^{(i)}} E((x^t u_k^{(i)})^2 | x \in \omega^{(i)}) \tag{5.4}$$

Calculating the expectation first, (5.4) becomes

$$\sum_{\substack{j=1 \\ j \neq i}}^{K} \sum_{k=1}^{p^{(i)}} u_k^{(i)} Q^{(j)} u_k^{(i)} - \sum_{k=1}^{p^{(i)}} u_k^{(i)} Q^{(i)} u_k^{(i)} \tag{5.5}$$

where $Q^{(i)}$ is the correlation matrix of class $\omega^{(i)}$, which is defined as

$$Q^{(i)} = E(xx^t \mid x \in \omega^{(i)}) \tag{5.6}$$

By combining (5.5) with the normal condition of bases $u_1^{(i)},...,u_{p^{(i)}}^{(i)}$,

$$u_k^{(i)t} u_k^{(i)} = 1 \quad , \quad k = 1,\ldots, p^{(i)} \tag{5.7}$$

Using the Lagrange multiplier method, minimization of (5.2) is transformed to the minimization of the next term:

$$\sum_{k=1}^{p^{(i)}} u_k^{(i)t} (\sum_{\substack{j \neq i \\ j=1}}^{K} Q^{(j)} - Q^{(i)}) u_k^{(i)} - \sum_{k=1}^{p^{(i)}} (\lambda_k^{(i)} u_k^{(i)t} u_k^{(i)} - 1) \tag{5.8}$$

Taking the derivative of this term with respect to the base vectors $u_k^{(i)}(k=1,...,p^{(i)})$, we obtain the necessary condition for a minimizing solution:

$$(\sum_{\substack{j \neq i \\ j=1}}^{K} Q^{(j)} - Q^{(i)}) u_k^{(i)} = \lambda_k^{(i)} u_k^{(i)} \quad , \quad k = 1,\ldots, p^{(i)} \tag{5.9}$$

From equation (5.9), it is known that the solutions for the base vectors $u_k^{(i)}$ $(k=1,...,p^{(i)})$ of $\mathbf{L}^i$ are the eigenvectors of the next matrix:

$$Q = \sum_{\substack{j \neq i \\ j=1}}^{K} Q^{(j)} - Q^{(i)} \tag{5.10}$$

In addition, by setting the $i$th eigenvalue of $Q$ as $\lambda_k^{(i)}$, (5.8) becomes

$$\sum_{k=1}^{p^{(i)}} u_k^{(i)t} Q u_k^{(i)} = \sum_{k=1}^{p^{(i)}} \lambda_k^{(i)} u_k^{(i)t} u_k^{(i)} = \sum_{k=1}^{p^{(i)}} \lambda_k^{(i)} \tag{5.11}$$

So, in order to minimize (5.8), we can select the eigenvectors that correspond to the minimum $p^{(i)}$ eigenvalues as the ortho-normal base of $\mathbf{L}^i$. Here, the dimension $p^{(i)}$ of the subspace is the parameter used to adjust the mean projection length on the classes.

Because the subspace $\mathbf{L}^i$ is uniquely determined from the base vectors $u_k^{(i)}$ $(k=1,...,p^{(i)})$, the above procedure determines the subspaces which minimize the enhanced CLAFIC criterion (5.2).

### 5.2.4 Unmixing by the Subspace Method

Once the class base vectors $u_k^{(i)}$ $(k=1,...,p^{(i)})$ are determined as the eigenvectors corresponding to the eigenvalues of the correlation matrix, the projection matrix $P^{(i)}$ is calculated from equation (5.3). The length of the projection of the observed mixel spectral vector $\boldsymbol{x}$ on the class subspace $\mathbf{L}^i$ is calculated as

$$x^t P^{(i)} x = \sum_{k=1}^{p^{(i)}} (x^t u_k^{(i)})^2 \tag{5.12}$$

This projection length expresses how much of the mixel vector belongs to the class $\boldsymbol{\omega}^{(i)}$. By a natural extension of the membership values, we interpret this projection as a measure of the class component contained in the mixel vector and define the unmixing in each class as the projection on the class subspace calculated by (5.12).

## 5.3 Unmixing Experiment using CASI Data

In order to check whether unmixing by the subspace method works effectively for hyperspectral images, we conducted an unmixing experiment using a 288-channel CASI and compared the results with those of conventional statistical unmixing methods.

**Table 5.1.** Characteristics of CASI spectral mode and image acquisition condition.

| Specifications of CASI sensor | |
|---|---|
| Band width | 1.8 nm |
| number of bands | 288 channels |
| Band range | 410.3—923.7 nm |
| Image size | 39 pixels, 489 lines |
| Dynamic range | 12 bit |

| Image acquisition conditions | |
|---|---|
| Altitude | 3000 m |
| Velocity | 200 km $h^{-1}$ |
| Ground resolution | 3.7 m (along swath)<br>12.6 m (along flight) |
| Date | 31 Aug 1993 |
| Time | 11:25—11:30 |
| Weather | Clear |

**Table 5.2.** Base vectors of 7 subspaces which correspond to 7 classes.

| CLASS | PC | EIGEN VALUE | EIGEN VECTOR | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | 478.5nm | 549.5nm | 620.7nm | 692.4nm | 764.5nm | 836.8nm | 909.2nm |
| YOSHI | 1 | -129.52 | -0.0173 | 0.0408 | -0.0053 | -0.0256 | -0.2552 | -0.0381 | -0.0236 |
| | 2 | -2.25 | -0.0458 | 0.0249 | 0.0457 | 0.0010 | 0.0130 | -0.0167 | 0.0501 |
| | 3 | -1.87 | -0.1464 | 0.0076 | -0.0140 | 0.0200 | 0.0235 | -0.0437 | -0.1009 |
| HANNOKI | 1 | -337.47 | -0.0464 | 0.0763 | -0.0082 | -0.0107 | -0.1042 | -0.0525 | -0.0185 |
| | 2 | -2.30 | -0.0232 | 0.0450 | -0.0110 | 0.0637 | -0.0131 | -0.0323 | -0.0378 |
| | 3 | -2.17 | -0.0278 | -0.0173 | 0.0466 | -0.0402 | 0.0049 | 0.0233 | 0.0411 |
| MIZUGOKE | 1 | -31.42 | -0.0505 | -0.0599 | 0.0166 | 0.0468 | -0.0654 | 0.0206 | 0.0239 |
| | 2 | -2.72 | -0.0846 | 0.0278 | -0.0390 | 0.0552 | 0.0410 | 0.0267 | 0.0267 |
| | 3 | -2.11 | 0.0137 | 0.0181 | 0.0297 | 0.0590 | 0.0366 | 0.0139 | 0.2041 |
| ISOTSUTSUJI | 1 | -6.98 | 0.1075 | -0.0568 | -0.0076 | -0.0470 | -0.0205 | 0.0005 | 0.0539 |
| | 2 | -1.74 | 0.0567 | 0.0022 | 0.0107 | 0.0081 | 0.0722 | 0.1964 | -0.1394 |
| | 3 | -1.30 | 0.1149 | 0.0396 | -0.0430 | -0.0078 | -0.0041 | -0.0110 | -0.0146 |
| SUGE | 1 | -146.98 | 0.0509 | -0.1038 | 0.0076 | 0.0328 | -0.1448 | 0.0094 | 0.0263 |
| | 2 | -3.00 | -0.0570 | -0.0223 | 0.0409 | 0.0101 | -0.0545 | 0.0682 | 0.0515 |
| | 3 | -2.08 | 0.0419 | 0.0229 | 0.0808 | 0.0391 | -0.0189 | 0.0644 | -0.1118 |
| WATER | 1 | -604.14 | 0.0856 | 0.0215 | -0.0516 | -030621 | -0.0015 | -0.0108 | 0.0171 |
| | 2 | -1.39 | -0.0654 | 0.0132 | 0.0252 | -0.0124 | -0.0353 | 0.0296 | 0.0476 |
| | 3 | -0.89 | -0.0256 | -0.0013 | -0.0372 | -0.0115 | 0.0527 | -0.0566 | -0.1631 |
| ROAD | 1 | -24656.21 | 0.0620 | 0.0711 | 0.0939 | 0.0836 | -0.0173 | -0.0293 | -0.0239 |
| | 2 | -47.39 | -0.0146 | 0.0529 | 0.0654 | -0.1158 | -0.0247 | 0.0633 | -0.0073 |
| | 3 | -5.04 | -0.0215 | -0.0564 | 0.0342 | 0.0676 | 0.0031 | -0.0031 | 0.0074 |

### 5.3.1 Study Site

The spectral image used for our analysis was a CASI image acquired over the Kushiro wetland. The CASI spectral sensor can measure a spectrum from 470 to 920 nm with a 1.8 nm band width. The specification and the data acquisition conditions for the CASI sensor are shown in Table 5.1. The image was acquired at an altitude of 3000 m from a Cessna 404 aircraft. The ground resolution was 12.6 m. Each pixel in the image contains the mean spectral radiance of the ground target.

A selection of 7 bands from the original CASI image (spaced every 40 channels) is shown in Figure 5.1. The first 4 channels are in the visible spectrum and the others are in the near infrared. In the center of Figure 5.1 is Lake Akanuma and the artificial dike across the area is clearly visible. There are various wetland plant species in this study area, especially reeds, sedges, and sedum, overlapping and continuously distributed over sphagnum moss.

Before the analysis, the CASI image was corrected for geometric distortion caused by the roll of the aircraft and the digital numbers were converted to radiance values (Babey and Soffer 1993).

### 5.3.2 Unmixing

The spectral characteristics of the 7 land cover classes used for unmixing are shown in Figure 5.2. All the classes are wetland vegetation communities (except for the road and water classes). The spectral differences between these vegetation classes are difficult to discriminate using a common remotely sensed image with a small number of bands. So far, wetland vegetation classification has not been intensively studied with respect to overlapping and continuously changing vegetation distributions, due to the lack of established methods (Yamagata 1995). However, from wetland ecosystem conservation planning and the global warming model perspectives, wetland vegetation classification has become an urgent research theme.

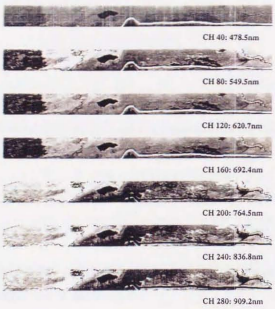



**Fig. 5.1.** Selected 7 channel image of original CASI data.

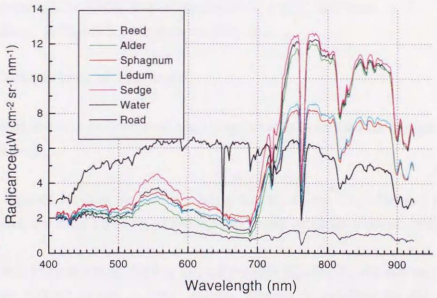



**Fig. 5.2.** Training spectra of 7 classes used for unmixing.

### 5.3.3 Procedure of Unmixing

The process of unmixing by the subspace method applied to CASI imagery was as follows:

1) Nine pure pixels (end-member points) for each unmixing class were selected as the training data based on knowledge from field surveys.

2) Using training vectors, the class correlation matrix $Q$ was calculated by equation (5.6).

3) The eigenvalue problem using the class correlation matrix $Q$ was solved to determine the subspaces for each class.

4) The projection of pixel vectors of the CASI image on the class subspace was calculated using equation (5.12).

5) The projection (component of unmixing) for each class was normalized to (0, 1) and mapped to an image.

The calculated subspace (base vectors) is shown in Table 5.2. The base vectors consist of 288 dimensions. However due to restriction of space, only those in the 7 channels corresponding to Figure 5.1 are shown in the Table.

### 5.3.4 Comparison of Methods

As mentioned above, 3 conventional unmixing methods were used for comparison with the new method:

1) Least Squares Method: Assuming a linear mixing model, proportions of each class in a mixel are determined by a least squares model using the training data.

2) Quadratic Programming: Adding a condition that the proportions add up to unity in a linear mixing model, a least squares solution is obtained by the quadratic programming method.

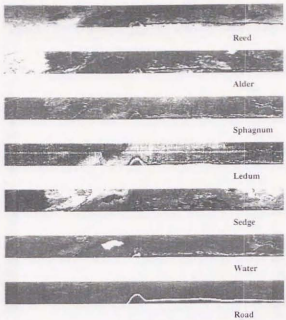



**Fig. 5.3.** Unmixing by subspace method.

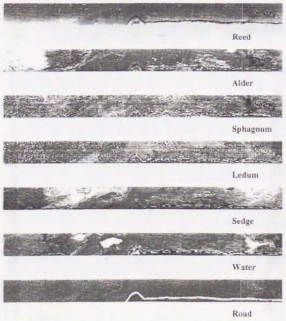



**Fig. 5.4.** Unmixing by least squares method.

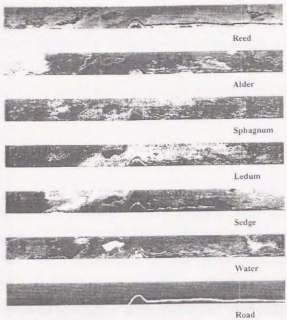



**Fig. 5.5.** Unmixing by quadratic programming method.

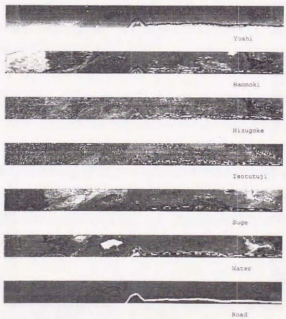



**Fig. 5.6.** Unmixing by orthogonal subspace projection method.

3) Orthogonal Subspace Projection Method: First, the projection of the mixel vector onto the orthogonal complement space spanned by the class vectors of the other classes is computed. The inner product of this projected vector and the class vector are then calculated (Harsanyi and Chang 1994).

### 5.3.5 Results of Unmixing

The result of unmixing by the subspace method applied to the CASI image of Kushiro Mire is shown in Figure 5.3. The results of unmixing by conventional least squares, quadratic programming, and orthogonal subspace projection methods are shown in Figures 5.4, 5.5, and 5.6, respectively. Here the unmixed vegetation classes are Yoshi (*Phragmites*: Reed), Hannoki (*Alnus*: Alder), Mizugoke (*Sphagnum*: Sphagnum Moss), Isotsutsuzi (*Ledum*), Suge (*Carex*: Sedge).

We compared the accuracy of quantitative classification of unmixing by the subspace method with that of the other methods, and investigated the correspondence between them and the actual vegetation distribution from field surveys. The following results were obtained:

1) The subspace method highlighted the reed area mixed with sedge as Sedge class (Figure 5.3).

2) Comparison of Figures 5.3 and 5.5 shows that the subspace and quadratic programming methods delineated accurately the ground pattern of the Sedge class.

3) Only quadratic programming (Figure 5.5) delineated the Moss and *Ledum* classes, which are spectrally very similar (Figure 5.2). This result may be due to a constraint of quadratic programming, i.e. it tries to enhance the subtle spectral differences between classes to increase membership difference.

4) The Alder class was accurately delineated only by quadratic programming (Figure 5.5).

5) Water and Road classes were delineated accurately by all methods.

### 5.3.6 Evaluation of Unmixing Methods

Based on the results obtained above, the unmixing methods can be evaluated as follows:

1) Spectrally distinct classes such as Road, Water and Sedge (Figure 5.2) are well unmixed by the subspace method (Figure 5.3).

2) Spectrally similar classes such as *Ledum* and Moss (Figure 5.2) are unmixed sufficiently only by quadratic programming (Figure 5.5).

3) The results achieved by the orthogonal subspace projection method (Figure 5.6) are exactly the same as those achieved by the least squares method (Figure 5.3).

4) Quadratic programming (Figure 5.5) provides the most accurate pattern of unmixing across all classes. However, it is the most time-consuming to implement. The subspace method uses a very fast algorithm and unmixing is performed through a simple inner product calculation suitable for parallel processing.

## 5.4 Conclusion

A new approach for the spectral unmixing problem by the subspace method is proposed and applied to wetland vegetation unmixing using hyperspectral imagery. For hyperspectral imagery, unmixing by the subspace method is superior to that done by conventional methods in numerical stability and computation speed. The results of the unmixing experiment showed that unmixing by subspace is accurate except for classes that are spectrally very similar. In the near future, the number of sensor channels and the size of the image area will rapidly increase. The fast and stable unmixing algorithm based on the subspace method will be most useful for such data. Further, we need to improve the separability between the spectrally very similar classes by developing the present approach further.

# Chapter 6
# Feature Selection for Classification

In this chapter, by applying feature-selection methods, effective spectral band combinations for wetland vegetation classification were investigated using airborne MSS (Multi Spectral Scanner) image data. By maximizing the Jeffries-Matusita distance and maximizing the classification accuracy of the test data, near-infrared, mid-infrared, and the green bands were selected. The maximum likelihood and minimum distance methods were used as the classification algorithms. The highest classification accuracy when using 7 bands was attained by the minimum distance method. Using this band combination, classification of the whole MSS image was conducted. This chapter is based on Yamagata et al. (1995).

## 6.1 Introduction

Most of the remote sensing studies so far have focused on classification using remotely sensed data. There are few studies on the selection of effective bands for classifying wetland types. Moreover, most of the classification work has been done for marsh or swamp areas. There are very few investigations in bog (*Sphagnum* moss-covered) areas. In this study, effective spectral bands are selected by feature-selection methods for wetland classification. Selection is conducted based on airborne MSS data acquired over a bog in Kushiro Mire in Japan. Based on the selected bands, a vegetation classification map for the bog area is produced for the first time.

For band selection, we used two criteria of selection: maximizing the Jeffries-Matusita (JM) distance, which measures the separability between classes, and minimizing the error rate of test data classification. The optimal combination of bands was selected by evaluating all combinations by these criteria. Although there are many types of separability measures between spectral classes, the JM distance is superior to the other methods as an indicator of classification accuracy.

Band selection was also conducted using the divergence measure. However, as the result was almost same as that with the JM distance, in this study we describe only the results obtained by the JM distance.

We compared the results of band selection by changing the classification algorithm to determine how classification accuracy affected the band selection. We tested the minimum distance and maximum likelihood methods in selecting bands using as a criteria the minimizing of test data classification errors.

Moreover, to decrease the dependence of the selected result on the training data, we used a cross-validation method in which we changed the training and test data repeatedly with 20 different combinations. The selected bands were weighted for their effectiveness by evaluating all the band selection results for the 20 cases. Finally, the whole airborne MSS image was classified using the bands that showed the highest accuracy with the test data.

## 6.2 Data and Methods

### 6.2.1 Study Area and Wetland Classification

The Kushiro mire (Figure 6.1) is the largest wetland in Japan, located at the eastern part of Hokkaido Island. The wetland was formed by peat accumulation during the last 4000 years. The depth of the peat is 1-4 m. The average temperature is low (around 5.6 degrees Celsius), and annual precipitation averages 1100 mm. In the wetland, there are many endangered species of plants and animals. Kushiro Mire was designated as a national park in 1987 and registered as a Ramsar convention site; it is recognized as an important wetland for migrating birds.

Most of the wetland belongs to the fen class of mires, covered with reeds and sedge grass, while some parts belong to the bog class, covered with *Sphagnum* moss. Here, fens are defined as peat-producing wetlands influenced by soil nutrients from water flowing through the system. Fens are found in boreal and tundra regions, including the area of permafrost in the extreme north, while bogs are peat-producing wetlands in moist climates, where organic

material has accumulated over long periods. Their main feature is ombrotrophy, which means that water and nutrient input into the system is entirely through precipitation. There is no mineral input through soil water to the system. During peat formation, the bog area has risen above the land surface. Bogs are extremely acidic and nutrient deficient. Typically, the major vegetation component is *Sphagnum* moss (Aselmann and Crutzen 1989).

### 6.2.2 Airborne MSS Data

The specifications of the airborne MSS data used in the analysis are described in Table 6.1. Channels 1 and 2 are blue bands; channel 3 is a green band; channels 4 and 5 are red bands; channels 6,7, and 8 are near-infrared bands; and channels 9 and 10 are mid-infrared bands.

The chemical composition of leaves affects their spectral properties in the visible (400-700 nm) and shortwave infrared (700-2500 nm) regions. Absorption by photosynthetic pigments (chlorophyll, xanthophyll, and carotene) dominates the visible wavelengths. Each of the pigments has an absorption maximum in the 300-500 nm region. However, only chlorophyll absorbs in the red wavelengths. Principal absorption peaks of extracted chlorophyll-a occur at 430 and 660 nm and those of chlorophyll-b at 455 and 640 nm. Vegetation exhibits high reflectance in the near infrared (700-1300 nm) and high absorption in the middle infrared (1300-2500 nm). The near-IR wavelengths are greatly influenced by cellular structure and refractive index discontinuities within the leaf. Minor water absorption features near 960 and 1200 nm vary significantly in shape and depth and may be related to both cellular arrangement within the leaf and hydration state. The mid-infrared region is dominated by leaf water absorption and has been related to plant water content. The region intermediate to the water absorption maxima at 1450 and 1940 nm may be strongly influenced by cell structure, morphology, and tissue constituents. (Wessman 1994)

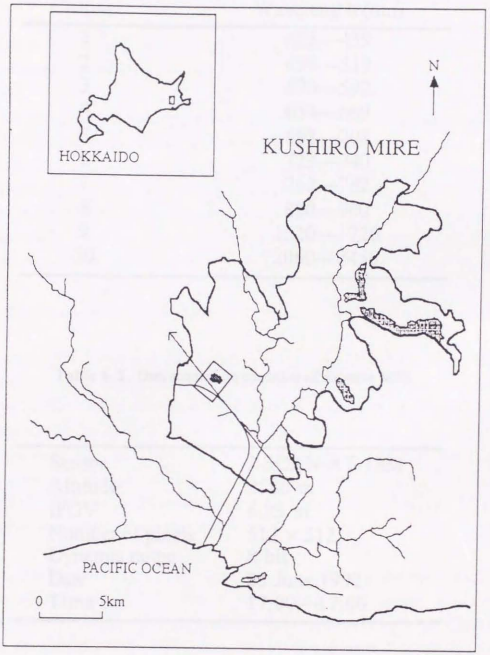



**Fig. 6.1.** Location map of Kushiro mire and study area.

**Table 6.1.** Spectral bands of airborne MSS..

| Channel | Wavelength (nm) |
|---|---|
| 1 | 425—439 |
| 2 | 499—519 |
| 3 | 570—592 |
| 4 | 654—669 |
| 5 | 688—708 |
| 6 | 723—740 |
| 7 | 762—782 |
| 8 | 820—900 |
| 9 | 1520—1720 |
| 10 | 2060—2450 |

**Table 6.2.** Data acquisition condition of airborne MSS.

| | |
|---|---|
| Sensor | J-SCAN-AT-18M |
| Altitude | 2500 m |
| IFOV | 6.25 m |
| Number of pixels | $512 \times 512$ |
| Dynamic range | 8 bit |
| Date | 23 June 1992 |
| Time | 17:00—17:06 |

Data acquisition conditions for the airborne MSS are described in Table 6.2. The data were acquired by the optical scanner sensor with an IFOV of 6.25 m carried in a Cessna 404 aircraft at an altitude of 2500 m on 23 June 1992. The image obtained was corrected for the radiance increase along the swath and for geometrical distortion. We then discarded 35% of the pixels on both sides to avoid the influence of the look-angle effect. Data of 512 pixels in the center area were used for the analysis.

When the data were acquired, the wetland vegetation was at the beginning of the growing stage, the *Sphagnum* moss was already active, but most of the vegetation above it was still moribund. There was no rainfall in the previous days and the surface of the vegetation was dry. The all-channel image of the MSS data is shown in Figure 6.2. Because the image was acquired at 5 p.m., the radiance was rather low in channels 6 and 7 and some sensor noise was also observed in these channels. We judged that it was not a problem to include these channels when selecting bands for classification.

### 6.2.3 Band Selection by JM Distance

An effective band combination for classifying wetland vegetation types can be selected by maximizing the mean Jeffries-Matusita (JM) distance between the classes. Considering the band combination with the maximum separability between classes as the best combination for classification, the optimal combination is found by exhausting all combinations and calculating the mean JM distance each time. In order to check the effect of the training data on the results, 20 different combinations of training data were used to select the band combinations. The selected 20 optimal combinations were then evaluated by weighting the bands with the number of times each was selected.

Here, the JM distance between two classes i and j is defined as the probabilistic distance between two distributions as described in Chapter 1.6.3. The minimum value for JM is 0 and the maximum value is 1414. The smaller the JM distance, the more difficult it is to separate the classes, while a JM distance of 1414 means perfect separation between the classes.

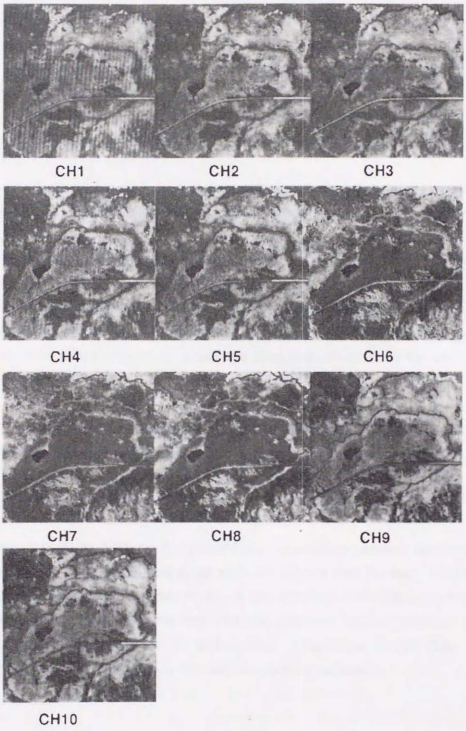



**Fig. 6.2.** Images of all airborne MSS channels.

The procedure used to select n effective bands using the JM distance is as follows:

1) Select 6 areas (5x5 pixels) for each class.

2) Assign 3 areas to training areas and the other 3 to test areas. (There are 20 different combinations of assignment)

3) List all combinations of the chosen bands out of the 10 available bands.

4) Calculate JM distances for all band combinations and all class pairs and average over the class pairs. From this we obtain the mean JM distance for all band combinations.

5) Select the band combination that gives the highest mean JM distance.

6) To evaluate the total effectiveness of a band for classification, the band receives a score (weight) according to the number of times it is selected during the selection process from n=1 to n=10. Thus, if a band is selected first and continues to be selected, it receives a weight of 10.

7) Weights of the bands are summed over all 20 combination of training and test data, from which the bands are ranked.

### 6.2.4 Band Selection by Accuracy of Test Data

In band selection using JM distance, the probabilistic separability between the classes is used. In classifying the whole MSS image, the same classifier is used for data which might have different spectral properties from those of the training data. Thus, there is no guarantee that a selected band using training data also gives the optimal combination for classifying real images. To compensate for this problem, we must use the test data, which were untouched previously, for assessing the real classification accuracy.

Therefore, we conducted band selection by using as a criteria tothe maximization of test data classification accuracy. Then, in order to check the dependency of the selection results on the classifier, we compared band selection using the maximum likelihood and minimum distance methods.

The band selection procedure for maximizing the classification accuracy was conducted in the same way as that for the JM distance method: band selection against the 20 combinations of training and test data was performed, weighted according to the number of times selected, and ranked according to effectiveness.

## 6.3 Results of Band Selection

### 6.3.1 Spectral Characteristics of Vegetation

The 6 training and test regions, each consisting of 5 x 5 pixels were selected for each vegetation class in the bog area. The regions were determined based on a field survey, aerial photographs, and the color composite image of MSS data. The mean spectra of the vegetation classes are shown in Figure 6.3.

From this figure, it is obvious that the spectral pattern changes considerably in the visible bands (1,2,3,4, and 5), near-infrared bands (6,7, and 8) and mid-infrared bands (9 and 10). The high reflectance in the near-infrared bands is due to growth of the sedges in the area, while the high reflectance in the mid-infrared bands is due to dead reeds.

To detect spectral distribution differences between the vegetation classes, pixel values in each region for each vegetation class were plotted using near-infrared channel 8 as the x-axis and the mid-infrared channel 9 as the y-axis (Figure 6.4). The near-infrared and mid-infrared bands are known to be related to the biomass of the vegetation and the water content of the surface, respectively. From this scatter plot, it is seen that the major vegetation classes such as Alder (Hannoki), Sedge (Suge), Reed (Yoshi) and *Sphagnum* moss (Mizugoke) are spectrally separable along these 2 axes. However, the distributions of vegetation types in the bog area such as *Sphagnum* (Mizugoke), *Ledum* (Isotututji), and pools (Titou) overlap.

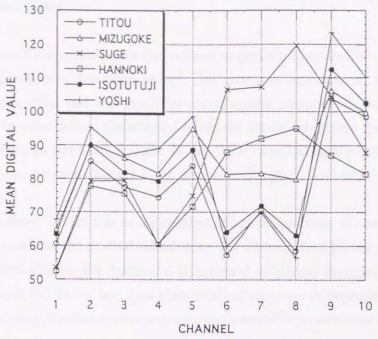



**Fig. 6.3.** Spectral characteristics of vegetation classes.

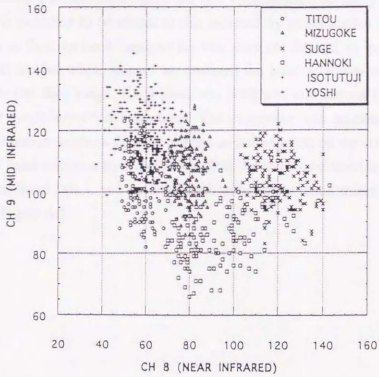



**Fig. 6.4.** Scatter diagram of vegetation classes for CH8 versus CH9.

### 6.3.2 Selected Band

The optimal band combination for the wetland vegetation classification was selected by the two criteria procedures described above. The results of selection by the JM distance are shown in Table 6.3a, and the result by classification accuracy is shown in Table 6.3b, respectively. In these tables, the selected bands are displayed against the training data used, in order of selection from left to right. In the Table, a minus sign means that the band was selected once but was not selected at the next selection point with additional channels.

It is shown in these Tables that as the number of bands increases, the order of selection of a band is more stable in the case of selection by JM distance than by test data accuracy. This is due to the fact that the JM distance is determined only from the statistical distribution of the training data, while the test data classification accuracy is dependent on both training and test data; hence, the latter selection is more susceptible to statistical fluctuation.

### 6.3.3 Assessment of Band Selection

When we need to classify all the pixels in the image, the selected band combination using training data does not necessarily give the best classification accuracy. The optimal band combination is expected to be closer to that selected by classification accuracy using test data. In order to find the band combination that does not depend on training data and that can work well for the whole image, we assessed the band selection results by averaging them over the test data range. Each band was weighted and ranked for its effectiveness following the procedure described above. The procedure was repeated by changing the classification method to check the dependency of the selection on the classification method. The results of band selection by JM distance (JM), by test data classification accuracy using maximum likelihood (ML), and by classification accuracy using minimum distance (MD) are shown in Figure 6.5.

**Table 6.3a)** Selected bands by maximizing mean JM distance.

| Training Data | | | Selected Channel (from left) |
|---|---|---|---|
| 1 | 2 | 3 | 8 9 3 -3 5 10 1 3 6 4 2 7 |
| 4 | 5 | 6 | 8 3 9 2 10 6 1 5 4 7 |
| 2 | 4 | 6 | 8 3 9 2 10 1 -2 5 2 6 4 7 |
| 1 | 3 | 5 | 8 9 5 2 10 4 1 6 3 7 |
| 1 | 2 | 4 | 8 3 9 10 1 5 -10 10 6 4 2 7 |
| 3 | 5 | 6 | 8 9 3 10 2 6 5 4 1 7 |
| 1 | 2 | 5 | 8 10 3 9 -10 10 6 1 5 4 2 7 |
| 3 | 4 | 6 | 8 3 9 5 6 2 -5 10 5 1 4 7 |
| 1 | 2 | 6 | 8 3 9 5 10 1 6 4 7 2 |
| 3 | 4 | 5 | 8 3 9 2 10 6 1 4 5 7 |
| 1 | 3 | 4 | 8 3 9 4 2 -4 6 10 1 4 5 7 |
| 2 | 5 | 6 | 8 10 3 9 -10 10 2 5 6 1 4 7 |
| 1 | 3 | 6 | 8 3 9 5 6 1 10 2 4 7 |
| 2 | 4 | 5 | 8 10 3 9 -10 2 10 6 1 5 4 7 |
| 1 | 4 | 5 | 8 10 3 9 -10 2 -3 10 3 6 5 1 4 7 |
| 2 | 3 | 6 | 8 9 3 5 10 1 6 2 4 7 |
| 1 | 4 | 6 | 8 3 9 5 6 2 10 1 7 4 |
| 2 | 3 | 5 | 8 9 3 10 1 6 5 4 2 7 |
| 1 | 5 | 6 | 8 4 3 -4 9 4 10 2 5 6 7 1 |
| 2 | 3 | 4 | 8 3 9 2 6 1 10 4 5 7 |

**Table 6.3b)** Selected bands by maximizing classification accuracy of test data.

| Training Data | | | Selected Channel (from left) |
|---|---|---|---|
| 1 | 2 | 3 | 4 -4 5 6 4 -5 10 2 -4 5 -6 8 -2 4 6 3 2 7 9 1 |
| 4 | 5 | 6 | 8 5 3 -5 9 5 6 -8 10 8 7 4 1 2 -7 7 |
| 2 | 4 | 6 | 8 5 3 -5 9 6 7 5 1 10 -1 2 4 1 |
| 1 | 3 | 5 | 10 3 8 -10 9 6 7 10 4 5 2 1 |
| 1 | 2 | 4 | 8 3 5 -5 9 10 7 4 1 2 -3 -4 5 6 -7 3 4 7 |
| 3 | 5 | 6 | 6 3 -6 8 9 7 2 -3 5 6 -7 1 -2 10 -1 2 4 1 7 -10 10 3 |
| 1 | 2 | 5 | 10 3 8 9 1 6 -8 -1 4 7 5 -6 8 1 2 6 |
| 3 | 4 | 6 | 10 5 8 -10 9 3 4 10 1 -3 6 7 2 3 -4 4 |
| 1 | 2 | 6 | 8 3 9 5 2 6 4 -6 10 1 7 6 |
| 3 | 4 | 5 | 8 9 3 -3 5 6 2 1 -2 3 10 2 4 -10 10 7 |
| 1 | 3 | 4 | 6 5 3 -5 9 5 10 4 -6 8 2 6 -10 10 7 1 |
| 2 | 5 | 6 | 8 9 3 1 -3 5 7 -1 2 3 6 -7 10 -2 4 7 1 2 -7 7 |
| 1 | 3 | 6 | 8 9 3 5 6 2 -2 4 10 2 -4 7 1 4 |
| 2 | 4 | 5 | 8 9 3 -3 4 6 5 1 3 -6 10 7 2 6 -10 10 |
| 1 | 4 | 5 | 8 3 6 -8 -6 8 9 6 10 7 4 5 2 1 |
| 2 | 3 | 6 | 10 3 8 -10 9 4 6 5 10 7 2 1 |
| 1 | 4 | 6 | 8 9 3 -9 10 6 9 -10 7 10 2 1 -4 5 -7 4 7 |
| 2 | 3 | 5 | 6 4 3 -4 -6 8 9 4 2 -4 7 -2 5 10 4 6 -7 1 2 -4 4 7 |
| 1 | 5 | 6 | 8 5 10 3 -5 6 5 9 -10 1 10 2 7 4 |
| 2 | 3 | 4 | 8 9 3 -9 10 5 6 4 2 -2 7 9 1 2 -7 7 |

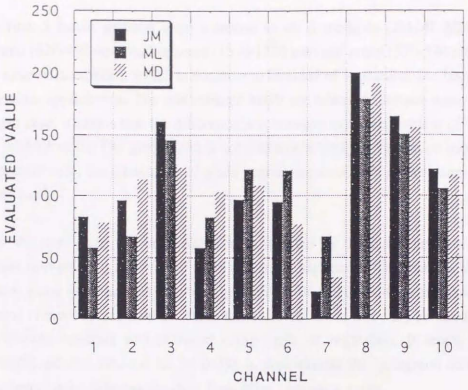



**Fig. 6.5.** Evaluated value of band selected by JM distance method, Maximum Likelihood method and Minimum Distance method.

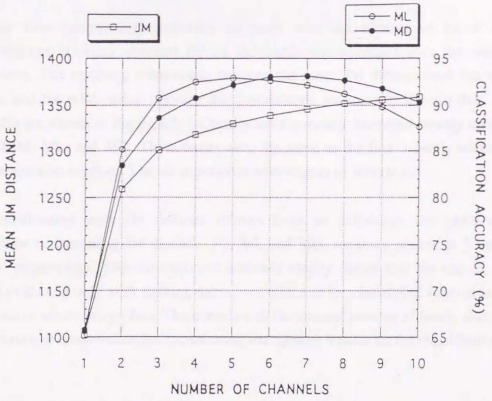



**Fig. 6.6.** Relationship between number of bands and classification accuracy.

The first 3 bands selected were common to all 3 methods (JM,ML,MD), that is near infrared (820-900 nm), mid infrared (1520-1720 nm) and green (570-590 nm) in that order. The near-infrared band, which is sensitive to biomass of vegetation, has been used in many vegetation applications. The mid-infrared bands are related to surface water content; in the present case, it seems that the difference was between the water content of live sedges and that of dead reeds. The green band is not commonly used for vegetation monitoring, but in the present case, the reflectance of green vegetation at its early growth stage seems distinct in this band.

After the third band, different bands were selected by the three methods. Among them, another mid-infrared band (2060-2450 nm) has a high score. This band is also sensitive to surface water content, and is often combined with other mid-infrared bands to delineate mineral components. However, these mid-infrared bands also vary in sensitivity to the 3-dimensional structure and chemical components of vegetation. It seems that this mid-infrared band was selected for its ability to discriminate the *Sphagnum* moss class, which has a completely different structure from other vegetation types.

### 6.3.4 Relationship with Classification Accuracy

To see how classification accuracy increases with the number of bands used, the classification accuracy obtained for all the bands was averaged over the training data selections. The resulting relationship between the mean JM distance and the number of bands, and the relationship between the classification accuracy of the test data using ML and MD are shown in Figure 6.6. Classification accuracy increased rapidly up to 3 bands for all JM, ML, and MD. These bands were the same as the first 3 bands selected above. This result also supports 3 bands superiority with respect to accuracy.

The relationship with JM distance follows from its definition: the additional bands contribute to increasing JM distance. For ML and MD, accuracy peaked at 5 bands and 7 bands, respectively. This saturation of accuracy clearly shows that the use of additional bands evaluated only with training data is not effective for classifying vegetation types for test data or whole image data. The existence of the optimal number of bands also shows the importance of cross-validation for selecting the optimal feature set for classification.

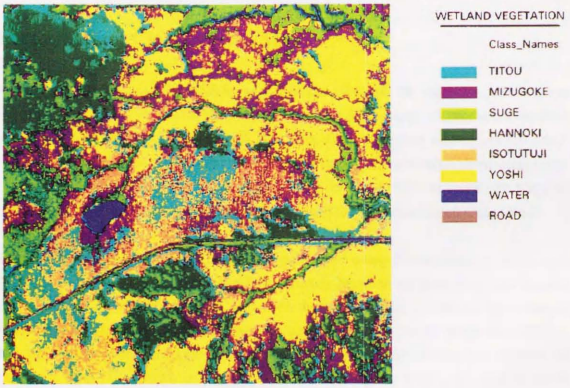



**Fig. 6.7.** Wetland vegetation classification map produced from airborne MSS.

Moreover, the results show that as the number of bands increases, the accuracy of MD is superior to that of ML. This means that ML is not always the optimal classification method when the distribution of vegetation is not normal. MD is simpler but more robust than the more rigid ML, especially when there are not enough training data to estimate the parameters of the classifier.

## 6.4 Vegetation Classification Map

The 7-band combination (2,3,4,5,8,9, and 10) that showed the highest classification accuracy with MD was used with the data of the whole MSS image to produce a wetland vegetation map of the bog area (Figure 6.7). Here, the 6 classified classes are the small pools (Titou) in *Sphagnum* moss-dominated areas, the *Sphagnum* moss-dominated areas (Migugoke), the sedge-dominated areas (Suge), the alder tree forest (Hannoki), the areas of *Ledum* bushes (Isotutuji) over the *Sphagnum* moss, and the reed-dominated areas (Yoshi).

Finally, classification of the MSS image was conducted using the consistently selected 3 bands (3,8, and 9). There were not many differences between the results of this 3-band classification and those using 7 bands. However, the proportion of *Sphagnum* moss area in the classification based on 3 bands decreased from 28% to 21%, while those of sedge and *Ledum* were somewhat increased. These differences are probably due to the absence of mid-infrared band (10), which is effective in discriminating the *Sphagnum* moss-covered areas.

## 6.5 Conclusion

As an application of feature-selection methods, an effective spectral band combination for classifying wetland vegetation was determined based on airborne MSS image data. We used as criteria of maximizing the Jeffries-Matusita distance and maximizing the classification accuracy of the test data. Three bands were consistently selected: near infrared (820-900 mm), mid-infrared (1520-1720 mm), and green (570-590 mm), in that order. As the classification algorithm, the maximum likelihood and the minimum distance methods were used, and the same 3 bands were selected using both methods. The classification accuracy of the test data also increased up to the same 3 bands. The highest

classification accuracy was attained when using 7 bands by the minimum distance method. Using the 7-band combination, the whole MSS image was classified into 6 wetland vegetation classes.

A conventional vegetation map in the wetland was produced from interpretation of aerial photographs. It was very difficult to discriminate among the complicated vegetation types in the bog area from these photographs. However, the above analysis it is showed that classification of wetland vegetation is possible using the spectral signatures of the vegetation types, especially in the near-infrared and mid-infrared regions. The MSS image data used in this study were acquired in the early growing stage of the vegetation. Completion of the wetland vegetation classification using these image data only is difficult. Combination of these data with images acquired in the middle of the growing stage to obtain more accurate classification results is necessary (Oguma and Yamagata 1986).

# Chapter 7
# Classification using Multitemporal Image Data

In this Chapter, a multitemporal approach for classification is investigated from several aspects to establish a methodology for wetland vegetation classification. The seasonal changes in wetland vegetation happen so quickly that it is difficult to classify the vegetation using only a scene acquired on a single date. In the experiment described here, we use multitemporal Landsat TM data acquired in June, August, and November, to classify the major vegetation types, such as reeds, sedges, alder trees, and *Sphagnum* moss, in the Kushiro wetland. From the sampling measurements of biomass and spectral reflectance of several vegetation types, it is clarified that each wetland vegetation type has a distinct growth pattern and temporal spectral signature. Using this temporal information, it is possible to classify the vegetation types with high accuracy. An accurate, wetland vegetation classification map is produced as the result of supervised classification using multitemporal Landsat TM images This chapter is based on Yamagata et al. (1996b).

## 7.1 Introduction

In the wetland area, there are many valuable species of plants and animals. Wetlands have been recognized as one of the most important ecosystems in terms of species diversity and its conservation. In order to conserve wetland ecosystems, accurate monitoring of the state of wetlands, especially vegetation, is vital. However, it is difficult to study wetland vegetation due to water inundation. There has not been enough research on the spatial distribution of wetland vegetation. Remotely sensed images from satellites or airborne platforms seem to be the only sources for monitoring wetland vegetation distribution.

Several studies on mapping wetland vegetation have been already conducted using aerial photographs in the Kushiro wetland. However, this approach has several limitations: 1) The

vegetation types are not distinct in color photographs. 2) Seasonal changes of vegetation prevent accurate mapping using aerial photographs acquired on a single date. 3) Geometrical distortion caused by the lens and aircraft motion, and radiometric distortion from atmospheric effects are too difficult to correct, making it difficult to produce a continuous mosaic of aerial photographs. It is also difficult to produce a classification map using a single classifier. Classification is often conducted manually, but depends too much on human interpretation, such that assessment of accuracy is difficult.

In this study, multitemporal Landsat TM data are used to produce an accurate wetland vegetation classification map. There are several advantages to this approach: 1) We can make use of spectral signatures, especially in the near-infrared and red bands. 2) By using multitemporal image data, seasonal growth patterns that depend on the vegetation type can be used for classification. 3) Since the whole wetland (in the case of the Kushiro wetland) fits in one Landsat image, there in no need to make a mosaic of several scenes.

## 7.2 Study Area

The study area was the Kushiro wetland, described earlier. Because of the rapid development in the surrounding agricultural areas, nutrient rich water has been flowing into the wetland and has dramatically changed the vegetation distribution. Among other changes, alder trees have increased dramatically in number during the last 50 years.

Wetland vegetation classification so far has been conducted using spectral signature differences observed in remotely sensed image data. However, the seasonal changes of the vegetation occur rapidly, and the growing season varies from species to species. The spectral signature differences between seasons are much greater than those between vegetation types.

## 7.3 Seasonality of Wetland Vegetation

By knowing how the wetland vegetation changes seasonally, we can determine the optimal period for image acquisition using remote sensors. We conducted sampling measurements during the growing season for the most typical vegetation types in the Kushiro mire: sedge

and reeds. These types have nevertheless not been successfully discriminated in conventional classification maps. We conducted measurements in the test area of the above-ground biomass for these types, 8 times from May until September. Sampling measurements were conducted in two 1 x 1 m areas at each site of sedges and reeds (Figure 7.1). It can be seen from the Figure that the sedges begin to grow in late May and reach maximum size around early August, after which they begin to wither. The reeds begin to grow in late June, reach their maximum size around late August, and then wither. Thus, there is a distinct difference in the temporal pattern of growth between sedges and reeds

In both test sites, there were several other plant species. Vegetation types are often specified by the dominant species, such as reeds or sedges. Therefore, it is necessary to classify the wetland vegetation types using the growth pattern of the dominant species.

From the results of our seasonality survey, the following are relevant to timing image acquisition. 1) In late June, we can see the difference between growing sedges and dead reeds. 2) In late August or early September, the vegetation difference between sedges and reeds is difficult to distinguish in terms of biomass. 3) The combination of the early season (from late May until early June) and late season (from late August until early September) is effective to discriminate the growth pattern differences between sedges and reeds.

## 7.4 Seasonality of the Spectral Signature

When we combine multitemporal remotely sensed data, we can use both the seasonal growth information and the spectral signature. For this purpose, we took spectral measurements of several wetland vegetation types using a spectral radiometer, in order to see how the spectral signatures change temporally. The measured vegetation types were *Sphagnum* moss, reeds and sedges. The spectral measurements were conducted in the early growing stage (23 July 1992), and in the late growing stage (31 August 1992).

**Table 7.1.** Characteristics of Landsat TM sensor.

| Channel | Wavelength (μm) | Band | Resolution (m) |
|---|---|---|---|
| 1 | 0.45—0.52 | Blue | 30 × 30 |
| 2 | 0.52—0.60 | Green | 30 × 30 |
| 3 | 0.63—0.69 | Red | 30 × 30 |
| 4 | 0.76—0.90 | Near IR | 30 × 30 |
| 5 | 1.55—1.75 | Mid IR | 30 × 30 |
| 6 | 2.08—2.35 | Mid IR | 30 × 30 |
| 7 | 10.4—12.5 | Thermal | 120 × 120 |

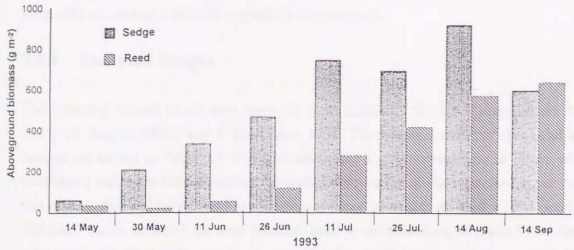



**Fig. 7.1.** Seasonal changes in above ground biomass of reed and sedge test site.

A spectral radiometer (Opt-research MSR-7000) was used to measure spectral reflectance in the wavelengths of 400-2500 nm. The measurements were made from 1.5 m vertically above the canopy. The measured spectral radiance was converted to reflectance divided by the spectral radiance of a standard white board, measured at the same time. Because the data were noisy at wavelengths around 1400 nm and longer than 1800 nm, we eliminated data in these regions, and the all data were smoothed using a median filter of 15 nm width.

The measured spectral reflectances for vegetation types at the 2 dates are shown in Figure 7.2. In June, the reeds were still dead and their spectral pattern was flat. In August, when the reeds reached maximum size, strong absorption by chlorophyll in leaves was observed in the red band, as well as strong reflection in the near-infrared band. Growth of sedges had begun by June, although there were still dead plants visible. The difference of the spectral pattern for sedges between the June and August was not as large as that of the reeds. The spectral change of sphagnum moss during this period was minimal. We found that the spectral reflectance of this moss had a special pattern at wavelengths of around 970 and 1180 nm. Although we could not measure the spectral reflectance for the alder trees, we expected that their spectral signature would be very distinct because they are actively growing in both June and August.

## 7.5 Classification using Multitemporal Image Data

By using the temporal pattern of spectral signatures of these wetland vegetation types, we were able to conduct a wetland vegetation classification.

### 7.5.1 Analyzed Images

The remotely sensed image data analyzed were Landsat TM scenes acquired on 26 July 1991, 29 August 1991, and 1 November 1991. The characteristics of the Landsat TM images are shown in Table 7.1. The TM sensor has a ground resolution of 30 m, which is considered sufficient for delineating the spatial distribution of the main land cover types in this region. The whole region of the Kushiro wetland is imaged in 1024 x 1024 pixels. The TM has 7 bands, in which the near-infrared (band 4) and mid-infrared (band 5) are the most effective bands for discriminating among wetland vegetation types (Yamagata et al. 1995).

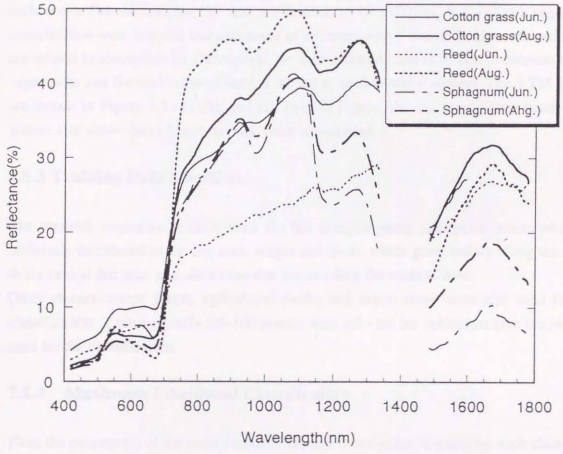



**Fig. 7.2.** Spectral reflectance characteristics of 3 wetland vegetation types in June and August.

### 7.5.2 Registration of Image Data

The 3 Landsat TM scenes were geometrically corrected using a 1:50000-scale topographic map onto the 20-m grid cell data. Registration was performed using first-order Affine transformation, and the r.m.s error of the correction was within one pixel. Three bands from each scene (bands 3,4, and 5) that are known to be effective for wetland vegetation classification were selected and composed as an image with 9 channels. Here, the red bands are related to absorption by chlorophyll; the near-infrared band is related to biomass of the vegetation; and the mid-infrared band is related to surface water content. The 3 TM scenes are shown in Figure 7.3 (a), (b), and (c). In each Figure, the red color shows band 4, the green color shows band 5, and the blue color shows band 3.

### 7.5.3 Training Data Selection

The wetland vegetation classes used for the analysis were *Sphagnum* moss, which is uniformly distributed in the bog area, sedges and reeds, which grow mainly along the rivers in the central fen area, and alder trees that are invading the wetland area.
Other classes--water, forest, agricultural fields, and urban areas--were also used for the classification. Training pixels (25-100 pixels) were selected for subclasses (see below) and used for the classification.

### 7.5.4 Maximum Likelihood Classification

First, the parameters of the multitemporal, spectral distribution function for each class were estimated using training data. Several subclasses are assumed to belong to each class. For example, in agricultural fields there is grassland, bare land, upland fields, etc. For each wetland vegetation type, subclasses were defined according to differences in growth patterns, water conditions, etc. The estimated subclass distribution functions were used to classify the pixels of the image. These classifications were allocated among 7 classes. However, the subclasses were used in the classification, since several land cover subclasses (or land conditions) correspond to the each major land cover class, and also to avoid a distribution function that is skewed rather than normal..

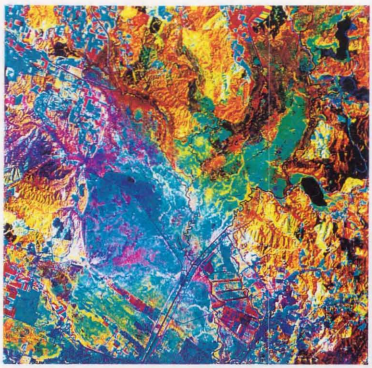


**Fig. 7.3(a)** Landsat TM scene in June 26th. (RGB = 453 with histogram equalization)

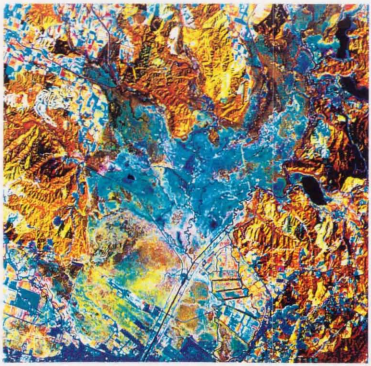


**Fig. 7.3(b)** Landsat TM scene in August 29th. (RGB = 453 with histogram equalization)

**Fig. 7.3(c)** Landsat TM scene in November 1st. (RGB = 453 with histogram equalization)

Moreover, in selecting training areas, we carefully chose positions such that the training sites were scattered around the whole image scene. We also repeated the process of selecting (deleting and adding) the training data and classifying them so that any misclassification arising from interference among them was minimized. This process was repeated until there was no further increase in classification accuracy. The resulting classification map of the Kushiro wetland is shown in Figure 7.4.

### 7.5.5 Assessment of the Classification Results

Based on the vegetation classification based on ground truth, we made the following assessment:

Because the conventional vegetation classification map of the Kushiro wetland was produced from a mosaic of aerial photographs, the sedges and reeds were misclassified in the areas surrounding the bog. In the present study, these were correctly classified in this region. However, part of the inundated area in the wetland was misclassified as dark forest. Nevertheless, several wetland-related researchers who evaluated the satellite data-derived map felt it was accurate enough for actual use in wetland management.

## 7.6 Conclusion

Because the seasonal changes in wetland vegetation happen quickly, it is difficult to classify wetland vegetation using a scene obtained on a single date. We have experimented to established a way of classifying wetland vegetation types using multitemporal remotely sensed image data. In the experiment, we used multitemporal Landsat TM data acquired over the Kushiro wetland in June, August, and November, to classify the major vegetation types, including reeds, sedges, alder trees, and *Sphagnum* moss. Evaluation of a biomass sampling measurement and spectral reflectances of several vegetation types demonstrated that the wetland vegetation types have distinct temporal growth patterns both in biomass and spectral signature. With this temporal information, we were able to classify the vegetation types with a high accuracy. We produced a vegetation classification map of the Kushiro wetland accurate enough to be used for wetland management purposes.

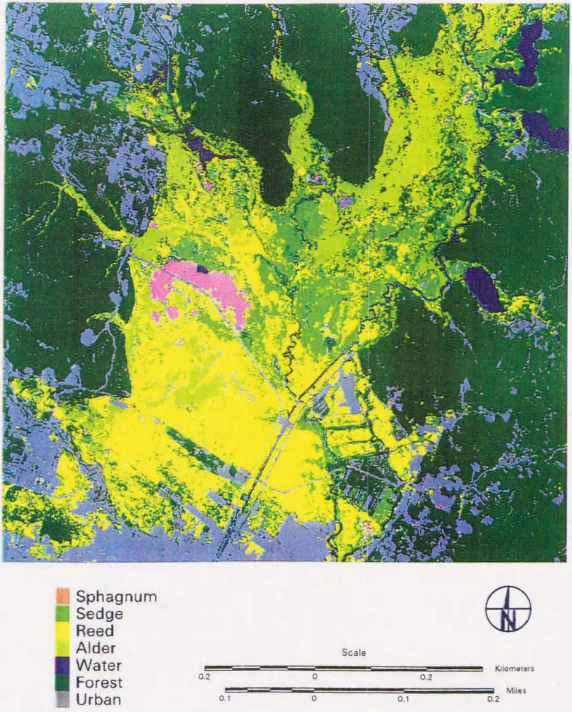



**Fig. 7.4.** Vegetation classification map of the Kushiro mire using multitemporal Landsat TM data.

In the TM scene in August, the central part of the wetland, where sedges usually grow, was totally inundated by floodwater. Because of the high percentage of cloud cover in the wetland area, satellite images can often only be obtained after heavy rain or a typhoon. So, there is a bias in the data in that flooding observed in the remotely sensed image is greater than that in normal wetland conditions, especially for scenes acquired in the middle of summer.

It has been shown that wetland vegetation types such as reeds, sedges, alder trees, and *Sphagnum* moss can be successfully delineated using Landsat TM data. However, the detailed vegetation types in the bog area are difficult to discriminate among due to the lack of both ground resolution and spectral resolution. To classify these vegetation types, we need to use a sensor with higher ground resolution and with more spectral bands. In Chapter 8, we show an example of such vegetation classification using hyperspectral image data.

# Chapter 8
# Classification using Spectral Image Data

In this chapter, wetland vegetation classification is conducted using a high-resolution spectral image, registered on a digital elevation model produced from ground measurements to analyze the relationship between vegetation and elevation. Compact Airborne Spectral Imager (CASI) data were acquired over a *Sphagnum* moss-covered bog. Vegetation classification is performed with an unsupervised classification (k-means) method, and then validated by a vegetation survey. Because the ground resolution was high enough to detect spatial changes in the bog vegetation types, we could discriminate even the *Sphagnum* moss class, which had never before been delineated using satellite sensors such as those on the Landsat TM. The results show the good performance of wetland vegetation classification in the bog area using high-resolution remotely sensed image data. This chapter is based on Yamagata et al. (1996a).

## 8.1 Introduction

Because wetland vegetation has a complex structure, it is not easy to discriminate among vegetation types using only a few spectral bands. As mentioned in Chapter 7, there are limitations on the usefulness of aerial photographs for making a vegetation map, and it is difficult to classify wetland vegetation types even using satellite sensors such as those on the Landsat TM. In this study, to monitor detailed wetland vegetation types, we tested the classification of vegetation using CASI data over the Akai wetland.

CASI sensors can obtain a spectral image in the visible and near- infrared wavelength region (400-900 nm). In supervised classification, it is possible to discriminate among vegetation types by selecting the training data of each class for input into the classifier.

However, as the number of channels increases, the number of the training samples necessary to estimate the parameters of the classifier increases rapidly. For hyperspectral data, such as from CASI, the number of channels often exceeds 100. However, it is difficult to obtain sufficient good training data, say 10000 observations for each class. Particularly for complex wetland vegetation types, the classes that can be classified are often not clear at the beginning of the analysis. In this analysis, we conducted a clustering (unsupervised classification), and the clusters were assigned to vegetation classes based on ground-truth information to produce a vegetation classification map.

The water flow in a wetland area is often determined by the difference between its elevation and that of the surrounding land. An elevation inclination inside the bog area is formed by the dome-shaped growth of *Sphagnum* moss, and these subtle elevation differences are related to water content and vegetation distribution in the wetland. In this analysis, by overlaying elevation data, the CASI image is displayed in 3-dimensional space, and the correspondence between vegetation and elevation is studied.

## 8.2 Data Analyzed

The data used for the analysis were from the CASI image acquired over the Akai wetland on 2 June 1993. The conditions of data acquisition are shown in Table 8.1. At the time of acquisition, the weather was fine. Most of the vegetation in the wetland was at the beginning of the growth stage, and plants were low in height. The image data have ground resolution of 2 m, appropriate for the nature of wetland vegetation change in the area. Before the analysis, geometric correction and calibration of the radiance data were conducted (Babey and Soffer 1992).

The spectral bands used for observation are shown in Table 8.2. Channel 1 is the blue band, channels 4 and 5 are red bands, channels 6 and 7 are between the red and near-infrared bands, and channels 8 and 9 are in the near-infrared band. The spectral signature of each vegetation type is shown in Figure 8.1. The spectral radiance features shown in this image were obtained from spectral-mode observation by the CASI sensor. Details of the vegetation types will be described later.

**Table 8.1.** The condition of CASI image acquisition

| | |
|---|---|
| Sensor | CCD (CASI) |
| Altitude | 7,000 ft (1,600 m above ground) |
| IFOV | 0.0690° (along track)<br>35.40° (swath) |
| Ground resolution | 1.9 × 2.2 m |
| Image size | 512 × 512 pixel |
| Dynamic range | 12 bit |
| Acquisition date | 2 June 1994 |
| Acquisition time | 11:00 a.m. |

**Table 8.2.** Spectral Channels of CASI

| Channel | Spectral bands (nm) |
|---|---|
| 1 | 460—470 |
| 2 | 550—560 |
| 3 | 595—605 |
| 4 | 635—645 |
| 5 | 655—665 |
| 6 | 675—685 |
| 7 | 695—705 |
| 8 | 715—725 |
| 9 | 745—755 |
| 10 | 825—835 |

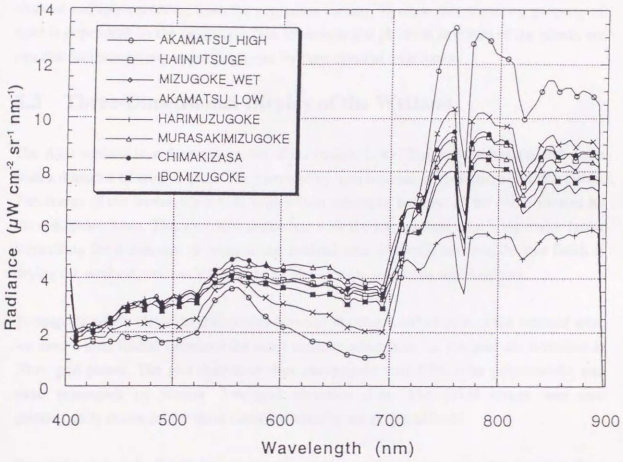



**Fig. 8.1.** Spectral radiance of each vegetation class.

The spectral reflectance from the vegetation has a steep rise between the red and the near-infrared band and has high value in the near-infrared region. This is due to chlorophyll absorption of the red light, while the near-infrared light is strongly (50%) reflected by plant leaves. Moreover, plant leaves are transparent (50%) to the near-infrared light. Actually we observe multiple scattering from the vegetation canopy. Because this scattering property of light is dependent on the vegetation type (chemical and physical structure of the plant), we can discriminate among vegetation types by their spectral reflectance.

## 8.3 Three-Dimensional Display of the Wetland

The Akai wetland is at the western side of Inawashiro Lake. The wetland has a round shape with a diameter of around 1 km. The surrounding area consists of mountains and rice fields. The center of the wetland is a little higher than the edges because of the dome formed by the sphagnum moss. This elevation difference with the surrounding area is the major factor controlling the movement of water in the wetland area. Drainage between the rice fields is drying the wetland, and has allowed many pine trees to invade its southern part.

To map the subtle elevation differences between the inside and outside of the wetland area, we used a total station (measure the exact distance using laser to) measure the elevation at 50-m grid points. The grid data were then interpolated with fifth-order polynomials, and were resampled to provide 2-m grid elevation data. The CASI image was also geometrically corrected for these elevation data for use in the analysis.

Figure 8.2 shows the CASI data displayed in 3-dimensional space using the elevation data. This is a bird's eye view from the south in which the elevation difference is enhanced 5 times. The image is displayed in false colors: red is assigned to the near-infrared band 10, green is assigned to the red band 7, and blue is assigned to green band 3. The image clearly shows the trees invading the southern part of the wetland.

Figure 8.3 shows a CASI image inside the wetland area, overlaid with the detailed elevation contour lines and their measured values. The center part of the wetland is seen to be around 2 m higher than the edges, with a slope to the southeast. The blue color corresponds to the inundated area.

**Fig. 8.2.** 3-D view of CASI false color image using detailed DEM.

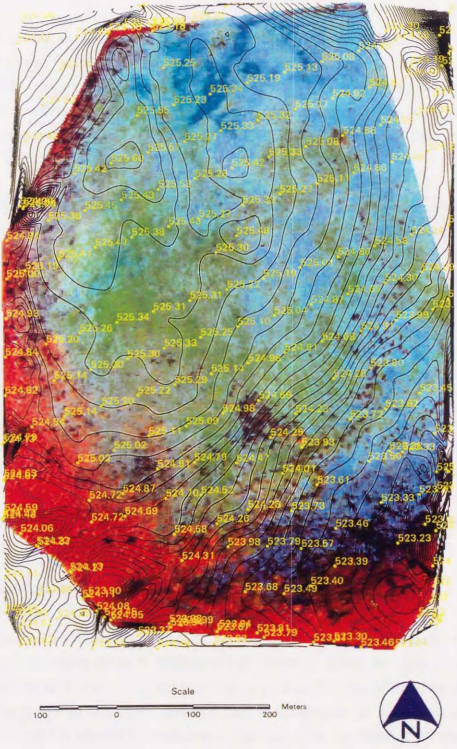



**Fig. 8.3.** False color CASI image of Akai mire with detailed DEM.

## 8.4 Wetland Vegetation Classification

The CASI image was clustered into spectral classes by unsupervised classification using the k-means method as follows. 1) The k-means method was repeated until the cluster means converged into 20 clusters. 2) Based on ground truth from the vegetation survey at 50-m grid points and interpretation of aerial photographs, the cluster classes were assigned to several vegetation species. 3) The cluster classes were then recoded into a stable community of 8 vegetation classes. The results are shown in Figure 8.4.

The delineated vegetation classes were 1) Ibomizugoke: small reed grasses growing over *Sphagnum* moss (Ibomizugoke), 2) Harimizugoke: *Sphagnum* moss (Harimizugoke) growing in the inundated area, 3) Murasakimizugoke: small bushes (Haiinutuge) growing over *Sphagnum* moss (Murasakimizugoke), 4) Haiinutuge: bushes (Haiinutuge) uniformly distributed over the *Sphagnum* moss (Ibomizugoke), 5) Chimakizasa: several grasses (Chimakizasa) mixed with bushes (Haiinutuge), 6) Mizugoke: inundated area covered with *Sphagnum* moss, 7) Akamatu (low): small red pine trees with some alder trees, and 8) Akamatu (high): large red pine trees.

These vegetation classes are the result of classification using spectral signatures that might change according to the growth stage of the vegetation. For example, although *Sphagnum* moss could be observed in this image because the other vegetation was still very small, the classification of *Sphagnum* moss type would become difficult with the growth of the other grasses. To perform more accurate classification, which considers this seasonal change, we need to use multitemporal remotely sensed image data as described in Chapter 7.

## 8.5 Relationship between Vegetation and Elevation

From the comparison of the contour lines and the vegetation classification results, we can investigate the relationship between vegetation and elevation. Along the slope from the center to the southeast, it is seen that the vegetation type changes from Murasakimizugoke to Ibomizugoke, while in the lower northern part, Harimizugoke is dominant. This is because Harimizugoke is adapted to inundation in that area. Near the center of the wetland,

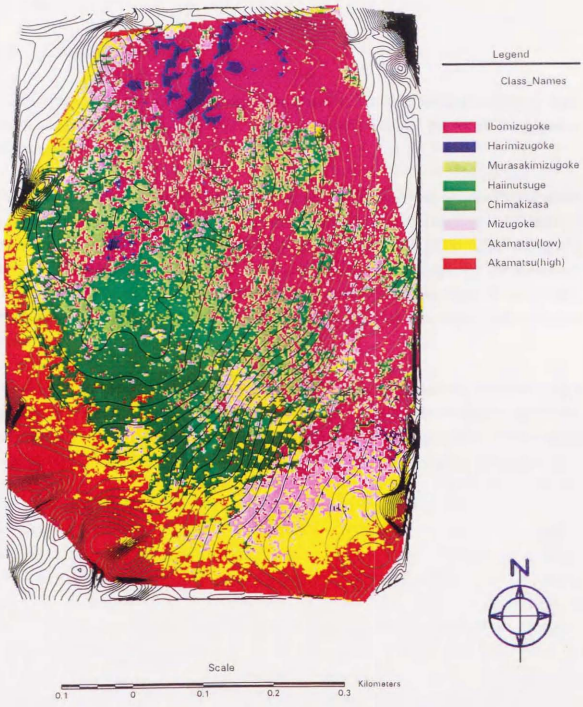



**Fig. 8.4.** Vegetation classification map using CASI image.

the slightly higher elevation keeps the surface dry and several shrubs and grasses are seen to grow there. In the southern part, although the elevation is lower, the area is dry because of drainage between the rice fields. This dryness is why many red pine trees are invading this area.

## 8.6 Conclusion

In this study, we have attempted vegetation classification in a wetland bog using high-resolution CASI data, overlaid with a digital elevation model produced from ground measurements.

The CASI data were acquired at the beginning of the Akai growing season. Classification was performed through unsupervised classification by the k-means method, validated by a vegetation survey. Because the ground resolution was high enough to detect the spatial changes of the bog vegetation types, we could discriminate among even the *Sphagnum* moss types, which has never been classified using satellite data. The result shows the high performance of wetland vegetation classification in the bog area using high-resolution image data.

The analysis of elevation effects shows the close relationships between vegetation types and the small elevation differences within the wetland. It was shown that the invading trees, the *Sphagnum* moss type, and the growth of the grasses in the central part of the wetland were strongly influenced by water inundation level, which is determined by elevation.

# Chapter 9
# Classification using the Gaussian Process

The Gaussian process is developed from Bayesian neural networks with an infinite number of nodes in the hidden layer. It is also a Bayesian model-averaging approach that integrates a model's predictions with the a posteriori probabilities of the parameters. In this Chapter, the basic theory of the Gaussian process for classifying satellite remote sensing data is introduced and tested using multitemporal LANDSAT TM, JERS1, and ERS1 SAR data. The accuracy of the classifications is compared with the maximum likelihood and Bayesian neural network methods; the Gaussian process outperforms the other methods for classifying LANDSAT/TM, JERS1/SAR, and ERS1/SAR data, and especially performs well with sensor fusion data. This chapter is based on Yamagata (1997c).

## 9.1 Introduction

Recent rapid developments in remote sensing techniques have made it possible to use time series of remotely sensed image data acquired by various satellite sensors for environmental monitoring. However, when it comes to classifying the remotely sensed image data of natural environments such as wetlands, the number of training samples that are available for classification is often limited due to difficulty in conducting ground surveys (Yamagata and Yasuoka 1996). When we try to classify a high-dimensional image that contains much information, the classification accuracy often decreases because the training data for constructing the classifier is not sufficient (Hoffbeck and Landgrebe 1996). Therefore, it is becoming more important to develop a classification method that can perform better than conventional methods even with a limited amount of training data.

In order to perform accurate classification using fewer training data, feature-selection and feature-reduction methods have often been used as a preprocessing step in the classification (Young and Fu 1986, Fujimura and Kiyasu 1996, Koller and Sahami 1996). Feature reduction transforms the variable space to new feature variables that are

considered to be effective by some evaluation function, while in feature selection, only the subset of features that are effective for the classification is selected. In both cases, the reduction in the number of parameters contributes to increasing the accuracy of the parameter estimation. Also, elimination of redundant information helps to increase classification accuracy.

In feature selection, classification accuracy for the feature subset is necessary. Cross-validation and probabilistic distance are often used as evaluation methods (Michie et al. 1994). In the cross-validation method, the training data are partitioned into several random subsets; each subset is classified by the classifier learned from the other subsets; and the accuracy of these classifications is then averaged to assess the accuracy of the model. In the probabilistic distance method, instead of actually assessing the classification accuracy, probabilistic distance between the classes such as the Jeffries-Matusita (JM) distance is used to measure the separability of classes. In both cases, the effective feature combination is found by exhausting the combinations of the feature subsets, by forward, backward, or stepwise selection methods.

In learning a classifier with few training data, there is the problem of overfitting. Overfitting occurs in the learning of a classifier with many parameters. When overfitting occurs, the classifier model overfits the training data, and loses applicability to other data. In order to avoid overfitting, we need to employ cross-validation to determine the model parameters by their performance against training data. There are other ways to avoid overfitting by using evaluation functions for model selection. These include information criteria such as AIC (Akaike Information Criteria) and NIC (Network Information Criteria) which have a penalty term that linearly increases with the number of model parameters (Akaike 1974). These information criteria are often used to avoid overfitting in selecting features and in determining the number of hidden layers in neural networks (Ripley 1996).

To some extent, it is possible to avoid degradation of classification accuracy when using high-dimensional remotely sensed images by feature selection and feature reduction. In these methods, only one model is selected and used later for classification of the real data. However, the result of model selection depends on the training data used. Since the training data are only a sample of the whole distribution, there is uncertainty in the model selection process. There is no guarantee that the selected model gives the best classification for the whole dataset. To account for this uncertainty, we need to employ

the Bayesian approach to model selection to achieve more consistently high classification accuracy.

In the Bayesian approach, model prediction is performed by integrating model estimations with the a posteriori probability of the model as the weight, when the a priori estimates by the model and the training data are given. In general, this calculation is very difficult both analytically and numerically. The a posteriori probability of a model is often calculated by some approximation method. The Bayesian approach using approximation can be divided into Bayesian model selection and Bayesian model averaging.

In Bayesian model selection, the a posteriori probability from a model is approximated by the Laplace method. Bayesian Information Criteria (BIC) are used to select the model. If the uncertainty of the training data is not considered in the model selection process, more complex models are preferred. However, using the BIC, which have larger penalties for increasing model complexity, can compensate for this drawback (Raftery 1995).

In Bayesian model averaging, Bayes factor, the likelihood of the validity of the training data, is calculated for the candidate models, which are selected by Bayesian model selection. Then, the model predictions are integrated using the Bayes factor as a weight to obtain the prediction of the Bayesian model (Madigan and Raftery 1994).

Bayesian model selection is applied to such problems as determining the graphical structure of the causality between variables (Buntine 1996), and optimizing the nodes of neural networks (Ripley 1996). Bayesian model averaging is used e.g. for probabilistic inference using Bayesian networks (Heckerman 1995). Several studies of Bayesian model averaging have been conducted using weight-decay parameters in the learning process of neural networks (Ripley 1995).

Bayesian neural networks developed by Neal (1996) constitute a genuine Bayesian approach that does not use approximation in calculating a posteriori probability. When the weights between the inputs, hidden layers, and outputs are given together with the a priori estimations, and the training data are available, integration using the a posteriori probability is calculated using the Markov-Chain Monte Carlo method. In Bayesian neural networks, Bayesian model averaging is conducted over the entire parameter

space. Because no model selection is included in the process, there is no danger of overfitting or underfitting, which arise from inappropriate model complexity. In this approach, there is also no need for early stopping of the learning process, or weight-decay to avoid overfitting. Therefore, for the learning process in Bayesian neural networks, there is also no need to partition the training data into test datasets for cross-validation. We can make use of all the training data for the learning process of the classifier.

The Gaussian process was developed from the Bayesian neural network approach. It is known that a neural network can approximate any function, if the number of nodes in the network increases. Unfortunately, such neural networks have so many parameters to estimate that it is impossible to learn these models with a limited number of training data points. However, such neural networks, with an infinite number of hidden layers, have been shown to converge to the Gaussian process with only a few parameters (Neal 1995, 1996).

In the Gaussian process, the model can be determined by the parameters on the covariance function of the output variables. It is possible to estimate these parameters with fewer training data than in neural networks. Moreover, the functional form of the Gaussian process is much clearer than that of neural networks, so it is easy to give them the appropriate a priori probabilities according to the problem. We can consider the Gaussian process as the successor to Bayesian neural networks, that can make use of training data information (Neal 1997).

In this study, we theoretically formulate a new classification approach using the Gaussian process and test this approach experimentally. In the classification experiment, wetland vegetation types are classified using remotely sensed images such as Landsat TM, JERS1 SAR, and ERS1 SAR data. We tested the new classification approach using the Gaussian process by comparing its accuracy with that of conventional methods, such as Bayesian neural networks and the maximum likelihood method.

## 9.2 Classification Method using the Gaussian Process

We assume that there are $n$ samples of training data from $K$ classes; the training data are expressed in pairs of $p$-dimensional feature vectors and the class label as $\{(x^{(1)},t^{(1)}),(x^{(2)},t^{(2)}),\ldots,(x^{(n)},t^{(n)})\}$, where $x^{(i)} \in \mathbf{R}^p$, $t^{(i)} \in \{0,\ldots,K-1\}$; the variable

$x^{(i)}$ is normalized beforehand to mean 0 with variance 1. The objective of the classifier is to label the class of a newly observed sample vector, that is to predict $t^{(n+1)}$ against the $x^{(n+1)}$.

However, it is difficult to directly model the classification process using the Gaussian process. First we employ the multiple logistic model, which is the same as the Softmax method often used in neural networks, to convert the classification problem to a regression problem (Ripley 1996).

Thus, we introduce latent variables $y_0^{(i)},\ldots,y_{K-1}^{(i)}$ to model the probability that the *i*th training sample belongs to the class k in the following equation:

$$P(t^{(i)}=k)=\exp(-y_k^{(i)})/\sum_{l=0}^{K-1}\exp(-y_l^{(i)}) \tag{9.1}$$

Then, assuming that the latent variable $y_k^{(i)}$ follows the stochastic process which has a covariance function:

$$Cov[y_k^{(i)},y_k^{(j)}]=\eta_k^2\exp(-\sum_{u=1}^{p}\rho_{ku}{}^2(x_u^{(i)}-x_u^{(j)})^2)+\delta_{ij}J \tag{9.2}$$

where, $\eta_k$, $\rho_{ku}$ is the hyperparameter, $\delta_{ij}$ is a delta function that is 1 only when $i=j$ and otherwise is 0, and J is the Jitter term that corresponds to the observational noise and has the effect of stabilizing the matrix calculations for estimation of parameters. We also assume that the arbitrary number of samples in this stochastic process follows a multiple Gaussian distribution.

Then, it can be shown from multivariate statistical inference theory (Giri 1977) that the K latent variables $y_k^{(n+1)}(k=0,\ldots,K-1)$ of the $n+1$th sample independently follow the Gaussian distribution which has mean *E* and variance *Var*:

$$E[y_k{}^{(n+1)}|y_k{}^{(1)},\ldots,y_k{}^{(n)}]=\mathbf{k}_k{}^{\mathrm{t}}\mathbf{C}_k{}^{-1}\mathbf{y}_k \tag{9.3}$$

$$Var[y_k{}^{(n+1)}|y_k{}^{(1)},\ldots,y_k{}^{(n)}]=v_k-\mathbf{k}_k{}^{\mathrm{t}}\mathbf{C}_k{}^{-1}\mathbf{k}_k \tag{9.4}$$

where, $\mathbf{C}_k$ is the covariance matrix of training data given by (9.2), $\mathbf{y}_k = [y_k^{(1)},\ldots,y_k^{(n)}]^t$ is the set of latent variables of training data, $\mathbf{k}_k$ is the covariance vector between $y_k^{(n+1)}$ and $y_k^{(i)}, i = 1,\ldots,n$ , while $v_k = Cov[y_k^{(n+1)}, y_k^{(n+1)}]$, respectively.

Then, using this probability distribution function, we can estimate the latent variable for each class. Moreover, using (9.1), we can estimate the probability that the unknown sample belong to each class.

Here, if we use the conventional statistical approach, the parameters in (9.2) can be estimated by the maximum likelihood method; then by calculating the inverse of the covariance matrix of each class, we can determine the classifier. However, in the Bayesian approach, the prediction is performed by integrating the model estimation with the a posteriori probability as weight for all parameter space:

$$P(y_k^{(n+1)}|x^{(n+1)},(x^{(1)},y_k^{(1)}),\ldots,(x^{(n)},y_k^{(n)}))$$
$$= \int P(y_k^{(n+1)}|x^{(n+1)},\theta)P(\theta|(x^{(1)},y_k^{(1)}),\ldots,(x^{(n)},y_k^{(n)}))d\theta \qquad (9.5)$$

where $P(y_k^{(n+1)}|x^{(n+1)},\theta)$ is the probability distribution of $y_k^{(n+1)}$, calculated using (9.3) and (9.4) for the specific parameter $\theta$.

Then the a posteriori probability $P(\theta|(x^{(1)},y_k^{(1)}),\ldots,(x^{(n)},y_k^{(n)}))$ of $\theta$ is proportional to the product of the a priori probability $P(\theta)$ and the likelihood $L$ by Bayes' theorem as,

$$L((x^{(1)},y_k^{(1)}),\ldots,(x^{(n)},y_k^{(n)})|\theta) = \prod_{i=1}^{n} P(y_k^{(i)}|x^{(i)},\theta) \qquad (9.6)$$

where $P(y_k^{(i)}|x^{(i)},\theta)$ is the probabilistic distribution of $y_k^{(i)}$, calculated using (9.2), (9.3), and (9.4) for specific $\theta$ and $x^{(i)}$.

Therefore, in order to calculate the integral (9.5), it is necessary to integrate the latent variable $y_k^{(i)}$ over all $\theta$ , by repeatedly calculating the likelihood $L$ using training data. At each repetition, we need to calculate the inverse of the covariance function. This integral calculation is not possible analytically or numerically using conventional methods. Here, we employed the Markov-Chain Monte Carlo method to approximate this integral.

In Markov-Chain Monte Carlo method, the integral calculation is approximated by summation of the sampling points of parameter $\theta^{(i)}, i = 1,\ldots,N$. generated for the Markov-Chain process, and which has a posteriori probability $P(\theta|D)$ as the stationary distribution of the process. The integral is calculated as

$$\int P(y_k^{(n+1)}|x^{(n+1)},\theta)P(\theta|D)d\theta \cong \frac{1}{N}\sum_{i=1}^{N} P(y_k^{(n+1)}|x^{(n+1)},\theta^{(i)}) \tag{9.7}$$

For rapid convergence of the right hand side of (9.7), it is better to use only the search points

$\theta^{(i)}, i = 1,\ldots,N$ that give high $P(\theta|D)$, rather than a take random walk in parameter space. Thus, we use a hybrid Monte Carlo method to determine the search points.

Finally, we can use (9.1) and the estimated latent values $y_k^{(n+1)}$ from this integral calculation to obtain the probabilities that the unknown sample belongs to each class. Then we can determine the class label for the unknown sample by choosing the class that gives the maximum probability.

## 9.3 Experiments and Assessment of Accuracy

In order to evaluate the classification accuracy using high-dimensional remotely sensed image data by the Gaussian process, we conducted a classification experiment. In the experiment, we used Landsat TM, JERS1 SAR, and ERS1 SAR data for classifying wetland vegetation types. The test site was again the Kushiro wetland. The image acquisition conditions and the characteristics of the sensors are listed in Table 9.1.

Landsat TM is an optical sensor. As mentioned earlier, past investigations have shown that the near-infrared band 4 is related to vegetation biomass and the mid-infrared bands 5 and 7 are related to surface water content. In this study, all band data, except those from thermal band 7, which has different ground resolution and is susceptible to environmental fluctuations, were used for the analysis.

**Table 9.1.** Sensor specification of satellite sensors used for the experiments.

| LANDSAT/TM | | | |
|---|---|---|---|
| Channel | Wavelength | Band | Resolution |
| 1 | 0.45 ~0.52 $\mu$ m | Blue | 30m |
| 2 | 0.52 ~0.60 $\mu$ m | Green | 30m |
| 3 | 0.63 ~0.69 $\mu$ m | Red | 30m |
| 4 | 0.76 ~0.90 $\mu$ m | Near Infra Red | 30m |
| 5 | 1.55 ~1.75 $\mu$ m | Mid Infra Red | 30m |
| 6 | 10.4 ~12.5 $\mu$ m | Thermal | 120m |
| 7 | 2.08 ~2.35 $\mu$ m | Mid Infra Red | 30m |

| | ERS1/SAR | JERS1/SAR |
|---|---|---|
| Wavelength | 5.66 cm | 23.5 cm |
| Polarization | VV | HH |
| Resolution | 30 x 30 m | 18 x 18 m |
| Frequency | 5.3 GHz | 1.275 GHz |
| Band | C | L |
| Incident Angle | 23° | 38.5° |
| Swath | 100km | 75 km |

The JERS1 SAR and the ERS1 SAR are Synthetic Aperture Radar (SAR) sensors that measure reflection from a surface of microwaves emitted from the satellite platform itself. Thus, they are called active microwave radar sensors. The JERS1 SAR uses a microwave band of 24 cm. In this band, most of the microwaves are scattered by rough surface features and tree trunks, while the ERS1 SAR uses the 6-cm band, which is scattered and reflected by leaves and small branches of trees in wetlands.

Using this spectral signature of wetland vegetation, it is possible to discriminate among the vegetation types from the remotely sensed image data. The seasonal change in these characteristics during vegetation growth is another valuable tool for classification as described in chapter 7.

In this study, using various remotely sensed datasets produced by changing the combinations of sensors, and using multitemporal data and speckle noise filters to process the SAR image data, the accuracy of the classification methods was assessed.

All the image data used in this analysis were geometrically corrected to 20-m grid data before the training data were selected, using topographic maps overlaid on each other.

The wetland vegetation classes used for classification were Hannoki (alder trees), Hannoki/Yoshi (mixture of alder trees and reeds), Titou (small pools in the *Sphagnum* moss area), Mitugasiwa (*Menyanthes trifoliata*), Yoshi (reeds), Yoshi/Suge (mixture of reeds and sedges), and Suge (sedges). These 7 vegetation classes are all named after the dominant species in each class. Unlike agricultural fields, there are no distinct boundaries between the classes. In selecting training data and test data, the aggregated area for a class was selected from image data based on ground truth observations in the field and interpretation of aerial photographs. We selected 100 pixels as training data and 100 pixels as test data for each class. The training data were fixed during the analysis, and the same test data were used to assess the classification methods.

In the following section, we describe the results of the classification experiments separately for 3 cases according to the type of sensors used. The classification methods used were common to all cases: maximum likelihood (ML), Bayesian neural networks (BNN), and the Gaussian Process (GP). In ML, all features are used for the classification; the covariance matrix for each class is estimated using the training data, which are also classified by ML. In classification using Bayesian neural networks and

the Gaussian process model, we employed the algorithms of Neal (1997).

### 9.3.1 Classification using Landsat TM scenes

Classification was conducted using 13 Landsat TM scenes separately, and their classification accuracy was compared. Figure 9.1 shows this comparison. In this Figure, the horizontal line shows the season of data acquisition, while the vertical line shows the classification accuracy as % misclassification.

For the results of the Bayesian neural network and the Gaussian process, the standard error of the accuracy is also depicted as bars. The accuracy of the Gaussian process was significantly higher than that of the ML in 9 out of 13 cases--more than 5% higher in 6 cases-- and was also higher than the Bayesian neural network in 7 cases.

While the Bayesian neural network was less accurate than ML in 4 cases, the Gaussian process was less accurate in only one case. Thus, the Gaussian process consistently outperformed the other methods used here to classify Landsat TM scenes.

### 9.3.2 Classification using JERS1 and ERS1 SAR

Using 4 multitemporal scenes of JERS1 SAR and ERS1 SAR, with 4 different kinds of speckle noise filters, classification accuracy was assessed for the 3 above methods. The acquisition dates for the SAR data were, JERS1 SAR: 3 April 1993, 21 May 1993, 30 June 1993, and 13 August 1993; and ERS1 SAR: 29 April 1993, 3 June 1993, 12 August 1993, and 21 October 1993. The speckle noise filters applied were: original image without filter (ORG), frost filter (FRS), lee filter (LEE), maximum posterior filter (MAP), and median filter (MED). Moreover, for the MAP and MED filters, which showed the highest classification accuracy, we also conducted classification experiments using an image composed of both JERS1 and ERS1 SAR data.

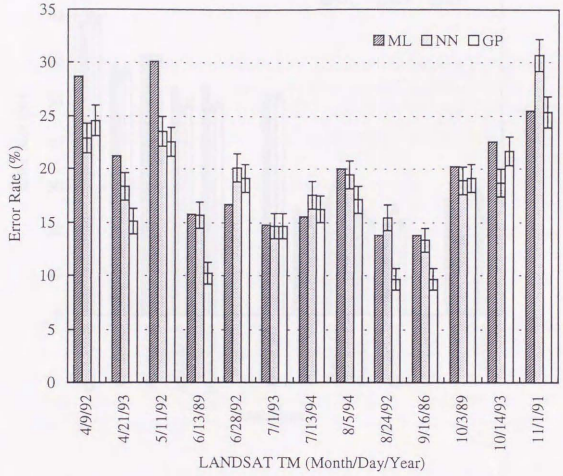



**Fig. 9.1.** Classification accuracy comparison for each LANDSAT/TM data.

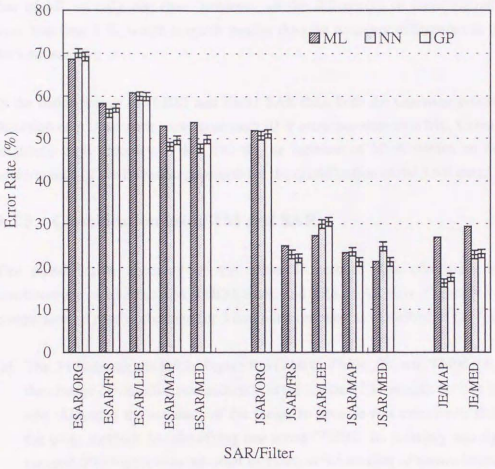



**Fig. 9.2.** Classification accuracy comparison for JERS1/SAR and ERS1/SAR data with speckle noise filters.

The comparison of the classification results obtained for these datasets is shown in Figure 9.2. The Gaussian process was significantly more accurate than the ML in 6 out of 10 cases. The Bayesian neural networks were significantly less accurate than the ML in 6 out of 10 cases. The Gaussian process showed lower classification accuracy than that of ML in only one case. However, all the differences in classification accuracy were less than 3 %, which is much smaller than the accuracy differences in the case of the Landsat TM.

In the fusion image of JERS1 and ERS1 SAR data, both the Gaussian process and the Bayesian neural network were more than 10 % more accurate than ML. Considering the relatively high accuracy of ML (70 %), an increase of 10 % means an outstanding performance of the Bayesian approach for the classification of the SAR data.

### 9.3.3 Classification using TM and SAR

The classification accuracy of the same 3 methods was compared for several combinations of Landsat TM, JERS1 SAR, and ERS1 SAR data. Figure 9.3 shows the comparison of results obtained for 3 combination cases as described in the following.

(a) The 3 results on the left in Figure 9.3 (TM06, TM06_08, and TM06_08_11) show the change of classification accuracy as the Landsat TM images are overlaid one by one. Although the accuracy of the Gaussian process was even lower than those of the other methods for classifying one scene (TM06), its accuracy was significantly (around 5%) higher than the other methods as the number of scenes increased.

(b) Using 13 Landsat TM scenes, NDVI (Normalized Difference Vegetation Index) defined as $NDVI = (TM4 - TM3) / (TM4 + TM3)$ was calculated for each scene and the calculated NDVI values were composed into an image with 13 channels. The results of the classification using these image data are shown in the middle of Figure 9.3 (NDVI). The image was further overlaid with a combination of JERS1 and ERS1 SAR images filtered with MAP. The result using this sensor fusion image is also shown in the Figure (NDVI_JEMAP).

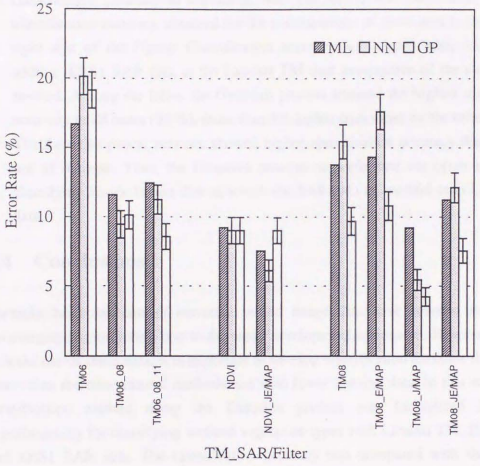



**Fig. 9.3.** Classification accuracy comparison for various combinations of LANDSAT/TM, JERS1/SAR, ERS1/SAR data.

(c) Although the SAR sensors can monitor the earth's surface through clouds, observation using optical sensors such as that on the Landsat TM is often prevented by cloud cover. Thus, the combination of a Landsat TM image with multitemporal JERS1 SAR and ERS1 SAR images is a very promising approach. Here, we combined a Landsat TM scene acquired in August 1992, which showed the highest classification accuracy as one scene, with the JERS1 and ERS1 SAR data. The classification accuracy obtained for the combinations of these data is shown in the right side of the Figure. Classification accuracy was significantly increased by adding JERS1 SAR data to the Landsat TM data irrespective of the classification method. Among the latter, the Gaussian process attained the highest classification accuracy in all cases (96 %), more than 5% higher than those by the other methods. The Bayesian neural network showed higher classification accuracy than ML in 1 out of 4 cases. Thus, the Gaussian process outperformed the other methods in classifying sensor fusion data in which the SAR data is overlaid onto Landsat TM data.

## 9.4 Conclusions

Recently, huge amounts of remotely sensed image data have become available for environmental monitoring due to the rapid development of sensors. However, in order to make use of these data, it is important to develop classification methods that perform better than the conventional methods and with fewer training data. In this study, a new classification method using the Gaussian process was formulated and tested experimentally for classifying wetland vegetation types with Landsat TM, JERS1 SAR, and ERS1 SAR data. The classification accuracy was compared with that of other methods, the ML and the Bayesian neural network.

The Gaussian process was developed from Bayesian neural networks as the limiting case when the number of hidden layers becomes infinite. By transforming the classification problem into a regression problem using the Softmax method with latent variables, the Gaussian process could be used for modeling the latent variables. Then, to predict the class label from the data, it was first necessary to predicate the distribution of the latent variables by integrating the model predictions with the a posteriori probability of the parameters as weights. The integral calculations were performed with the Markov-Chain Monte Carlo method.

The experiments using Landsat TM scenes showed that the Bayesian neural network and the Gaussian process methods outperform ML. Classification accuracy was increased by around 5 %. Classification results using the JERS1 SAR and the ERS1 SAR showed that the Gaussian process performed better than the other methods, but the accuracy increase was less than 3 %. Using the combination of Landsat TM and JERS1 SAR data, the Gaussian process outperformed the other methods and attained the best classification accuracy of all. From this result, we can conclude that the Gaussian process works well for sensor fusion data, especially for a combination of optical and radar sensors.

In general, the Bayesian approach integrates model predictions with the a posteriori probability as weight. It uses no model selection process, but is free from overfitting and underfitting problems; it can make use of the information in the training data for classification. The Gaussian process is a flexible Bayesian approach that can take many functional forms with fewer parameters to perform classification. However, at the moment, the biggest drawback of the Gaussian process is computation time. In the process of integrating the model space, the present algorithms need to invert the covariance matrix, which requires computational time parallel to the third order of the number of samples at all model space points. To calculate the Gaussian process with several hundred training data points, it takes 10 times the computational time required by the Bayesian neural network.

# Chapter 10
# Concluding Remarks and Future Directions

The preceding 9 chapters have described new remote sensing techniques and their application to wetland monitoring. Here, I review what has been accomplished in this thesis and describe the implications of the new techniques for monitoring wetlands and other complex ecosystems. Finally, I discuss future work that might be developed from these results.

## 10.1 Major Contributions of this Thesis

There are three major original contributions in this thesis. The first is the development of new spectral indices--PVI, WTI and VSWI--that are specifically designed for wetland monitoring. Second is the development of a new unmixing procedure using the subspace method. This new method is useful in unmixing high-dimensional remotely sensed data, for which it has better performance than conventional unmixing techniques. Third is the formulation of the Gaussian process modeling approach as a classification method; the good performance of the new method has been verified by experiments using optical and radar sensor fusion data.

In addition to these contributions, several extensive empirical studies on the classification of wetland vegetation have also been conducted, using multitemporal remotely sensed data. Through these results, I have established remote sensing techniques for wetland monitoring.

The conclusions of each chapter can be summarized as follows:

In Chapter 2, water inundation in rice fields, which are artificial wetlands, was analyzed to delineate the relationship between flooding and the influence of inundation on the vegetation. More specifically, TM images were subjected to several multivariate analyses in order to assess the extent of flood damage to the rice and to analyze the relationship between the flood and the damage. Owing to the relationship that we found

between floodwater turbidity and the actual decrease in rice yield, it was shown that the rice yield decrease could be estimated not only from the TM image data acquired shortly before the harvest, but also from the image acquired immediately after the flood.

In Chapter 3, the new spectral indices, PVI and WTI, were developed for flood inundation monitoring and vegetation change in wetland areas. These spectral indices were based on rice field flood - damage analysis. In the analysis, TM rice-field images acquired during flooding and one month later were used to relate inundation damage to rice yield. This was accomplished by using the turbid water pixels to determine a turbid water line. The WTI along the turbid water line and the PVI for paddy rice were defined using this line. The relationship was delineated between floodwater turbidity at the rice booting stage, which could be monitored using WTI, and the reduction in yield, which could be measured by PVI.

In Chapter 4, the PVI was further developed to the VSWI that can monitor vegetation, soil and water at the same time. An algorithm to determine automatically the end-member spectral points for vegetation, soil and water was developed by fitting a triangle to the scatter plot instead of finding the soil line. The distances between the spectrum point and the triangle edges were used as the new index. In conventional unmixing approaches, end-member points are often determined manually and arbitrarily from the image data or from a scatter plot of the data. The new algorithm automatically determines the end-member points. The VSWI was applied to wetland monitoring using multitemporal Landsat TM image data. Vegetation, soil, and water conditions and changes were successfully delineated.

In Chapter 5, a new approach to the unmixing problem, the subspace method, was developed and applied to wetland vegetation using hyperspectral image data. Unmixing by the subspace method was superior to conventional methods for hyperspectral imagery in numerical stability and computational speed. The results of an unmixing application showed that unmixing by subspace is accurate except for classes that are spectrally very similar.

In Chapter 6, by applying feature-selection methods, an effective spectral band combination for wetland vegetation classification was determined, based on airborne MSS image data. By using as standards the maximizing of the Jeffries-Matusita

distance and maximizing the classification accuracy of the test data, the near-infrared, infrared, and green bands were selected. The maximum likelihood and minimum distance methods were used as the classification algorithms. The best classification accuracy was attained using 7 bands by the minimum distance method. Using this 7-band combination, the whole MSS image data was classified.

In Chapter 7, we established a way to classify wetland vegetation types using multitemporal remotely sensed image data. Because seasonal changes in the wetland vegetation happen rapidly, it is difficult to classify using a scene obtained on a single date. In an experiment, we used multitemporal Landsat TM data acquired over the Kushiro wetland in June, August, and November, to classify the major vegetation types--reeds, sedges, alder trees, and *Sphagnum* moss. Distinct temporal growth patterns of wetland vegetation both in biomass and spectral signatures were apparent from measurements of biomass and spectral reflectance of several vegetation types. By using this temporal information for the vegetation types, it was possible to classify them with a high degree of accuracy.

In Chapter 8, vegetation classification was conducted in a wetland bog using high-resolution CASI data. The image was overlaid with a digital elevation model produced from ground measurements. CASI data were acquired at the beginning of the growing season over the Akai wetland using spectral bands known to be effective for wetland vegetation classification. The classification was performed using unsupervised classification by the k-means method, and then validated by a vegetation survey. Because the ground resolution was high enough to detect spatial change in the bog vegetation types, we could discriminate among even the *Sphagnum* moss types that had never before been delineated using satellite sensors such as Landsat TM. The results showed the possibility of wetland vegetation classification in bog areas using high-resolution remotely sensed image data.

In Chapter 9, a new classification method to classify wetland vegetation types using Landsat TM, JERS1 SAR, and ERS1 SAR data based on the Gaussian process was formulated and tested experimentally. The classification accuracy was compared with those of two other methods, maximum likelihood and Bayesian neural networks. The Bayesian neural networks and the Gaussian process method outperformed the maximum likelihood method in experiments using Landsat TM scenes. Classification accuracy was increased by around 5 %. The classification results using the JERS1 SAR

and the ERS1 SAR data showed that the Gaussian process method performed better than the other methods, but the accuracy increase was less than 3 %. Classification results using a combination of Landsat TM and JERS1 SAR data showed that the Gaussian process method outperformed other methods and attained the best classification accuracy overall. We conclude that the Gaussian process method works well for sensor fusion data, especially for combination of optical and radar sensor data.

## 10.2 Applications for Wetland Monitoring

The advanced remote sensing techniques developed in this thesis can be applied to actual wetland monitoring in several ways:

1) The Perpendicular Vegetation Index (PVI) is useful for monitoring vegetation biomass over inundated wetland areas. Firstly, because the vegetation biomass in paddy rice is directly related to yield (as shown in Chapter 2), the PVI can be used to estimate rice yield. This rice yield estimation is possible not only in the case of flood damage but also under normal growth conditions. As rice is the major grain feeding most of Asia, this estimation will be very important for global food supply prediction. Secondly, measurements of aboveground biomass of wetland vegetation are important for estimating rates of carbon accumulation in wetlands. It has recently become known that vegetation biomass in wetlands, including paddy fields, is related to rates of methane emission. These estimations of the accumulation and emission of greenhouse gasses (carbon dioxide and methane) will, in the future, provide very important parameters for modeling global warming. And thirdly, changes in vegetation growth rate in wetland areas are often caused by human disturbance, such as agricultural development in surrounding areas or tourism. The monitoring of vegetation biomass in wetlands make it possible to determine the extent of disturbance. Knowledge of vegetation changes in wetland areas is vital for planning appropriate wetland management.

2) The Water Turbidity Index (WTI) is useful for delineating the distribution of water turbidity or suspended sediments in flood waters. Because nutrients in a wetland are contained in the soils that are transported by flooding, the distribution of suspended sediment provides information on how the nutrients are distributed in the wetland. As the spatial distribution of nutrients in wetland areas has never been monitored using other methods, this approach using WTI will be very challenging and may

make possible analysis of the relationships between nutrient distributions and vegetation changes in wetland areas.

3) The Vegetation-Soil-Water Index (VSWI) is useful for monitoring wetland complexes using satellite-borne optical sensors. Because wetland areas are composed of vegetation, soil (or sediment), and water, this three-component index is the most natural one for evaluating the state of a wetland. This index is applicable not only to high resolution sensors, such as those on Landsat, but also to global coverage images, such as that prepared by NOAA. Although the Normalized Difference Vegetation Index (NDVI) has been used universally as the spectral index of choice for NOAA data, VSWI will be an alternative index that is more suitable for imaging wetland areas. Once the VSWI is calculated, we can proceed to quantitative studies like vegetation biomass estimation or to classification of the image using VSWI as a new signature space.

4) Unmixing using hyperspectral data is a promising approach for monitoring wetland areas. Because wetlands are among the most complex ecosystems in the world, extracting information on each wetland constituent is the ideal approach to study them. Particularly in bog areas, wetland ecosystems are very sensitive to environmental change. Even a subtle vegetational difference, which reflects the water flows beneath, might be an ecologically important factor for managing an area. Moreover the extracted spatial distribution of each vegetation type tells of the history of the vegetational succession in the past.

## 10.3 Implications for Monitoring Complex Ecosystems

The methods developed in this thesis can be further applied to monitoring of other complex ecosystems, such as forests, inland waters, etc. Of course, spectral indices are basically designed for specific targets of observation; the model structure as well as the parameters must be adjusted according to the application. However, the basic ideas of unmixing and classification are directly applicable to other fields by changing the end-member spectral points and the training classes according to the objective of the study. The monitoring of complex natural ecosystems is possible for several specific parameters:

1) The VSWI is applicable to forest monitoring. In this case, the vegetation end-member corresponds to the most densely vegetated area, the soil end-member corresponds to the background dry soil and the water end-member corresponds to the background humid soil rather than water. Because the VSWI is valid only when an unobstructed view of the background soil is apparent in some places in a scene, forest monitoring using VSWI will be appropriate in the sparsely vegetated semi-arid and northern tundra regions. However, there appear to be many different problems in applying this index to forests due to the different environmental conditions found in them, such as the effects of slope, the look-angle of the sensors, shadows, etc.

2) Unmixing of spectral classes using hyperspectral data is a promising approach in forest and water quality monitoring. Because the spectral signatures of forest species and algae in inland waters are different because of differing chemical components, it is, in principle, possible to discriminate quantitatively among these vegetation types with unmixing. Forest species unmixing is rather difficult because of the three dimensional structure of forests and the different spectral signatures of the components of trees such as trunks, branchs, and leaves, whereas the unmixing of algal type in water is more feasible. Moreover, if the spectral signature is observed in the mid-infrared bands, it will be possible to delineate from the spectral information even the chemical components of the target (lignin, etc.). This approach will cause a revolutionary change in the use of remote sensing in ecosystem monitoring. The unmixing method developed in this thesis may play an important role in the handling of hyperspectral data covering bands from the visible to the mid-infrared.

3) Advanced classification algorithms, such as the Gaussian process, are expected to be used extensively in classifying complex ecosystems using all sorts of remotely sensed data ranging from airborne images to global coverage imagery. Collecting ground truth data in complex ecosystems such as wetlands is difficult, which limits the availability of training samples. The Bayesian model is a promising approach to compensate for the uncertainty of limited training data and still achieve highly accurate classification. Although band selection will continue to be an important criterion for the design of future satellite sensors, present band selection algorithms are inadequate to handle the combinatorial explosion in calculations when hundreds of bands are needed for classification. The Bayesian approach, especially when combined with the Gaussian process, would be a prominent candidate for future classification algorithms for complex ecosystems.

## 10.4 Directions for Future Research

The performance of the methods inevitably depends on the application field as well as the remotely sensed data used. In order to obtain good performance in the specific application, we must search for optimal methods for each type of remotely sensed data; this thesis is a contribution to this search. However, the studies conducted in this thesis are not sufficient. We still need to investigate other remote sensing techniques for other types of sensors.

In the near future, huge amounts of new types of remotely sensed data will become available with the launching of the Japanese ADEOS II platform and the American EOS platform. These new data will provide us with global coverage as well as high-dimensional optical and radar band information. The potential of the new techniques from this thesis will be challenged by these data. Furthermore, more evaluations, on more datasets, and comparison with the best alternative methods, will be required before any conclusions can be drawn for global monitoring of the complex ecosystems.