SOFT SUPERVISED SELF-ORGANIZING MAPPING (3SOM) FOR IMPROVING LAND COVER CLASSIFICATION WITH MODIS TIME-SERIES

By

Siam Lawawirojwong

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of requirements
for the degree of

Geography - Doctor of Philosophy

2013

ABSTRACT

SOFT SUPERVISED SELF-ORGANIZING MAPPING (3SOM) FOR IMPROVING LAND COVER CLASSIFICATION WITH MODIS TIME-SERIES

By

Siam Lawawirojwong

Classification of remote sensing data has long been a fundamental technique for studying vegetation and land cover. Furthermore, land use and land cover maps are a basic need for environmental science. These maps are important for crop system monitoring and are also valuable resources for decision makers. Therefore, an up-to-date and highly accurate land cover map with detailed and timely information is required for the global environmental change research community to support natural resource management, environmental protection, and policy making. However, there appears to be a number of limitations associated with data utilization such as weather conditions, data availability, cost, and the time needed for acquiring and processing large numbers of images. Additionally, improving the classification accuracy and reducing the classification time have long been the goals of remote sensing research and they still require the further study.

To manage these challenges, the primary goal of this research is to improve classification algorithms that utilize MODIS-EVI time-series images. A supervised self-organizing map (SSOM) and a soft supervised self-organizing map (3SOM) are modified and improved to increase classification efficiency and accuracy. To accomplish the main goal, the performance of the proposed methods is investigated using synthetic and real landscape data derived from MODIS-EVI time-series images. Two study areas are selected based on a difference of land cover characteristics: one in Thailand and one in the Midwestern U.S.

The results indicate that time-series imagery is a potentially useful input dataset for land cover classification. Moreover, the SSOM with time-series data significantly outperforms the conventional classification techniques of the Gaussian maximum likelihood classifier (GMLC) and backpropagation neural network (BPNN). In addition, the 3SOM employed as a soft classifier delivers a more accurate classification than the SSOM applied as a hard classifier. Furthermore, the 3SOM-F, which applies both pure and mixed pixels during the training process, accomplishes more accurate and realistic classification results than the 3SOM-P, which applies only pure pixels in the training process. Therefore, these results suggest that the 3SOM-F could be considered the most appropriate method for land cover classification using time-series imagery. However, the results also demonstrate that there is uncertainty in the classification accuracy associated with network design architecture and internal parameter settings. As a result, the suitable neural network configuration should be investigated for the best performance of the classifier.

Additionally, two study areas, Thailand and the Midwestern U.S., are selected to investigate the performance of the 3SOM-F. All results confirmed that the classification performance of the 3SOM-F is effective even when it is applied to real landscape data in both study areas.

The proposed techniques will benefit detailed land cover classification at the regional scale. The spatial pattern of land cover classes can be valuable information for managing and understanding the environment as well as monitoring land cover change. Furthermore, the advantages of this research will contribute to various disciplines such as map updating, agricultural area estimation, cartography, and urban planning.

DEDICATION

To my parents, whose love and faith have always been unconditional.
I never would have reached this point in my life without them.

ACKNOWLEDGEMENTS

Over the past six years I have received many support and encouragement from a great number of individuals. My accomplishment is impossible without my advisor, Dr. Jiaguo Qi. He has the attitude and the substance of a genius; he continually and convincingly conveyed his spirit of adventure in regard to research. His guidance and suggestions have contributed greatly to my achievement. I am truly grateful for all the help he has provided me, the help I never would have got elsewhere. The knowledge and experiences I have received from him make me strong in my research career.

I would like to thank my committee members, Dr. Joseph Messina, Dr. Ashton Shortridge, and Dr. Sasha Kravchenko, who took part in this research for their generosity in sharing their time and ideas. I have learned so much through our conversations. I am grateful for Dr. Ashton's comments which helped me improve every detail in my research. I wish to thank Dr. Joe for providing me guidelines and suggestions for my research methodology and results and thank Dr. Sasha for her valuable statistical advice.

This research was financially supported by the Royal Thai Government, the Ministry of Science and Technology in Thailand, and the Geo-Informatics and Space Technology Development Agency (GISTDA). I am thankful for their support which helped me made all of this possible. The funding from the Graduate School, the Office for International Student and Scholars at Michigan State University (MSU), please accept my sincere gratitude for supporting me during my last year of study which allowed me to complete this dissertation.

I would like to extend my gratitude to the Department of Geography for providing me all the necessary knowledge and experiences in remote sensing and GIS as well as in geography. I

wish to thank Sharon Ruggles, Graduate Secretary of Department of Geography, for her immense help. She helped me get through obstacles during my six years at MSU. Her kindness made me my life in U.S. more pleasant. I would like to express my gratitude to my friends and staffs at CGCEO. Special thanks to Jean Lyle Lepard, Kathleen Mills, and Cameron Williams for helping me with all paper works and technical issues.

I would like to express my appreciation to Jenni Gronseth for editing and improving my English language in this dissertation. Her dedication to my work is very much appreciated. I am also grateful to Chanaichon Damsri (Prare), who helped me improve my writing skills. Their warmhearted help and assistance in my research will never be forgotten. I would like to express my special appreciation to Tanita Suepa (P'Jeap) for her support in this research. Thank you for being the best friend and a big sister during our five years of study together at MSU.

I am deeply thankful for my family, my parents, sister, brother, who always support and encourage me to continue my study in this country. Thank you for believing in me and giving me endless love and support. My dissertation is a very special memorial to my beloved mother, who passed away during my first year at MSU. This achievement is dedicated to her and I am sure that she will be glad and proud of my success.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

1.1 Research problem

Land cover data represents key environmental information for many science and policy applications and is universally used. It is also the most important terrestrial dataset. From regional to global scales, new and critical requirements for land cover information emerge from various environmental change issues. Up-to-date land cover information with highly accurate, detailed and timely results is required for the global environmental change research community to support a variety of science and policy applications (Wardlow *et al.*, 2007).

Remotely sensed data from satellite-based sensors is useful for a broad range of land cover mapping applications due to their spectral, spatial, and temporal resolutions. In addition, agricultural land cover at regional- and global- scales specifically require the ability to generate up-to-date results repeatedly and continuously. As a result, detailed regional-scale cropping patterns are needed to be mapped on a repetitive basis to characterize current land cover patterns and monitor land cover changes.

Although LANDSAT data, with multiple spectral bands and 30 m spatial resolution, provides detailed crop mapping, it still possesses quite a number of limitations. These limitations include low temporal resolution and small coverage for regional-scale mapping, data availability as well as considerable costs and time for acquiring and processing of the large number of scenes (Wardlow *et al.*, 2007).

Advanced Very High Resolution Radio Meter (AVHRR) from the National Oceanic and Atmospheric Administration's satellite is a valuable source for coarse resolution data (1 km)

with high temporal resolutions (10 to 14-day composite periods). AVHRR Normalized Difference Vegetation Index (NDVI) has been used to monitor vegetation conditions and major phenological events. However, the drawbacks of AVHRR data are the coarse resolution with possible integrated spectral-temporal response from multiple land cover types and that there are only five spectral bands (Bagan *et al.*, 2005).

Alternatively, the Moderate Resolution Imaging Spectroradiometer (MODIS) provides high-quality and scientific global coverage data with high temporal resolution (daily) and intermediate spatial resolution (250 m). MODIS is the alternative for detailed land cover mapping at a large spatial scale.

To monitor vegetation structure and function, NDVI is widely used for classifying land cover on a large spatial scale (Huete *et al.*, 1997, 2002). However, there are several limitations associated with NDVI that affect the accuracy of classification. These limitations are the sensitivity to atmospheric conditions and soil background, as well as the tendency to saturate at high biomass levels (Gao *et al.*, 2000). MODIS time-series data can produce the Enhanced Vegetation Index (EVI) at 16-day intervals which can be used for the classification of land cover. EVI was proposed to minimize the effects of the atmosphere and canopy background that contaminate NDVI and it had improved sensitivity over high biomass areas (Huete *et al*. 1997, 2002).

Furthermore, MODIS-EVI data can discriminate land cover types based on their unique phenological (seasonal) characteristics. Plant phenology is a significant factor in identifying, describing, and classifying the characteristics of different stages of periodic changes in a landscape. The satellite images combined with seasonal characteristics of different land cover types can make phenological classification possible to distinguish the differences among several

types of vegetation. Subsequently, the advantages of multi-temporal images are that they allow for phenological classifications to produce consistent and highly accurate results.

However, based on statistical assumptions, there appear to be a number of limitations associated with the application of time-series data to traditional classification methods. This is due to a number of uncertain factors such as different flight, location, and weather conditions (Bagan *et al*., 2005). The remote sensing image classification domain has been explored by scientists using statistical techniques. These techniques are based on spectral reflectance values with the assumption that the training data is normally distributed. The classical conventional spectral classifiers, such as the maximum likelihood classifier, perform well over limited areas where spectral signatures do not vary greatly from those captured in the training data. Therefore, variations in plant density, ages, and types can increase spectral confusion and decrease the accuracy of image classification.

In addition, the traditional classifications apply *"hard classification"* in which the output for each pixel comprises only the code of the class that has the highest strength of membership (Zhang & Foody, 2001). This technique assumes that the study area is unique and the internally homogenous classes are mutually exclusive. Such techniques cannot represent geographic phenomena and may lead to an inaccurate classification. This problem is heightened in the areas where the classes exist as continua rather than as a mosaic of discrete classes. This is due to the fact that classes in the real world are not typically separated by sharp boundaries (Zhang & Foody, 2001). These are called mixed pixel problems where the classification cannot be identified by a single homogeneous category in one pixel. This is because the image may contain more than one land cover class, particularly in coarser spatial resolution images (Foody, 1996b).

In order to deal with these limitations, the effective algorithm of an Artificial Neural Network (ANN) can be considered for use with *"soft classification"*. This alternative approach differs significantly from the traditional ones in the ways that an ANN is able to learn, store information, and react. The downside is that some of its processing elements are destroyed or impaired. Additionally, the main advantages of ANN techniques are that it is able to generate its own rules by learning from the examples. It does not require prior knowledge about the statistical characteristics of class data. Also, it is easy to combine multi-source data together. These important abilities of ANNs—to learn from input data and to generalize and predict unknown patterns based on the data source—can provide accurate output for image classification.

The *Self-Organizing Map (SOM)* based on ANNs is a robust approach because it provides topology-preserving mapping from a high-dimensional input space onto a low-dimensional map space. In addition, SOM can remove the problem of the local optimum in the learning process which is found in other techniques, such as Fuzzy C-means and MLP (Liu *et al.*, 2010).

However, most applications of SOM focus on the unsupervised pattern recognition, spatial information extraction, and ecological modeling (Li, 2007). Few studies apply this technique to supervised classification of remotely sensed images. In addition, when the SOM is associated with a supervised classification, a majority voting technique is usually used to determine which class each output neuron belongs to. However, this technique may lead to the problem of unlabelled neurons, in addition to causing unclassified pixels in the final map (Li & Eastman, 2006). In order to increase the effectiveness of SOM for image classification, it is necessary to improve this method.

Therefore, the main goal of this research is to provide an effective image classification algorithm for land cover classification of coarse resolution, MODIS-EVI time-series images. A self-organizing map (SOM) is improved upon to provide a supervised SOM (SSOM) and a soft supervised SOM (3SOM) in order to increase efficiency and accuracy of classification. In addition to spectral values, this research applies phenological information according to characteristics of MODIS-EVI time-series data to enhance the capability of land cover classification, which distinguishes this method from other currently used methods.

The proposed method will be beneficial at regional scales for detailed land cover classifications and change detections. Furthermore, the benefits of this method will contribute to various disciplines such as map updating, land cover monitoring, cartography, and urban planning.

1.2 Research objectives

In this research, *the long term goal* is to provide an effective image classification algorithm for high efficiency and accurate land cover classification of coarse resolution, MODIS-EVI time-series images.

In order to achieve the goal of this research, *the objectives of this study* are:

1) To improve the self-organizing map (SOM) to provide a supervised SOM (SSOM) and a soft supervised SOM (3SOM) in order to classify land cover by using MODIS-EVI time-series images.
2) To determine the appropriate input data needed for land cover classification by comparing time-series and phenology images

3) To determine the suitable neural network architecture and internal parameter values of the neural network-based classifiers applied in this research.
4) To identify the appropriate classifiers by comparing the accuracy of SSOM to the Gaussian maximum likelihood classifier (the statistically-based classifier) and the backpropagation neural network (the neural network-based classifier) regarding the applicability of the MODIS-EVI time-series images.
5) To investigate the advantages of soft and hard classification methods by comparing the accuracy of SSOM and 3SOM.
6) To investigate the advantages of 3SOM by comparing the accuracies derived from the fully-soft classification to the partially-soft classification.
7) To quantify the uncertainty of the classification accuracy of the SSOM classifier associated with the input data, training data, and the classifier.
8) To apply the identified classification using real world, remotely sensed landscape data.

Chapter 2

Background and Literature Reviews

2.1 Phenology from remotely sensed imagery

Phenology (from the Greek "*to show*" or "*to appear*") is the study of periodic biological events in the animal and plant worlds influenced by the environment, particularly temperature changes driven by weather and climate. The phenological events are those involved in the plant and animal life cycles which have been changed by seasonal and interannual variations in climate. Seasonality is a special terminology concerning phenological events. The meaning is similar to non-biological events, such as the timing of the fall formation and spring break up of ice on the fresh water lakes. Phenological principles are regarded as the observation of phenological seasons which can be explained in two terms: a phenological calendar and a phenological season. The phenological calendar is the occurrence date of various phenophases and their sequences in the annual cycle, whereas the phenological season represents the characteristics of different stages of the phenological landscape. Remote sensing is an essential key to studing seasonal and interannual seasonal characteristics of phenology across broad spatial and temporal scales (Schwartz, 2003).

Plant phenology has become an emerging indicator of landscape and environmental changes, and its responses to global environmental changes (Houghton *et al*., 1990). Plant phenology data extracted from remote sensing technologies can be studied for the spatial transition of phenological data from points to coverages in different time frames (Zhang *et al*., 2006).

2.1.1 The utilization of satellite data in phenology detection

Phenology can be used to identify, describe and classify different types of vegetation. Satellite time-series data with coarse resolution contain indispensable information on seasonal vegetation dynamics from regional to global scales as they provide consistent measures of vegetation greenness and activity (estimated by means of vegetation index or NDVI) at high temporal frequency over extended periods.

The wide range of uses for remotely sensed data are being increasingly recognized for land use/land cover classification purposes. The various temporal, spatial and spectral resolutions have heightened the importance of remote sensing to classify land use/land cover (Merry *et al*., 2000). The annual cycle of vegetation phenology inferred from remote sensing can identify phenological phases or growing seasons at annual time scales. (Zhang *et al*., 2003) Furthermore, remotely sensed data can be utilized to explore the changes of land use/land cover, particularly agricultural crops.

Time-series of NDVI derived from e.g. NOAA/AVHRR, SPOT/VEGETATION, or TERRA/MODIS spectral measurements, can be utilized to gather information on seasonal vegetation development. NDVI data are strongly correlated with the photosynthetic activity of plants. Myneni *et al*. (1997) mentioned that the timing of seasonal rise and fall in NDVI provide significant changes in the length of the active growing season. This information contributes to the analysis of the functional and structural characteristics of the global and regional land cover. Long time-series of NDVI data can also provide information on shifts in the spatial distribution of bioclimatic zones by indicating variations in large-scale circulation patterns of land use and agricultural crop changes.

2.1.2 Land cover classification utilizing phenological modeling

The satellite images, which provide a phenological approach, due to their multi-temporal capabilities can define land cover classes in terms of vegetation timing, duration and intensity of photosynthetic activity. Thus, differences in these phenological characteristics, combined with seasonal characteristics of different land cover types, can make phenological classification possible to distinguish the differences among types of crops (Roehrig, 2005). Regarding the remote sensing based phenology practice, there are various noteworthy examples of this research as follows.

In a study of how phenological differences in tasseled cap indices can improve deciduous forest classification, Dymond (2002) suggested that both the use of phenological information in satellite data and the use of vegetation indices have improved classification. Moreover, these sources of information also provide effective change detection of land cover classification.

The advantages of multi-temporal images are not only highly beneficial for phenological classification, but also result in higher classification accuracy in all classes. These advantages provide high potential for classification, particularly in the areas where vegetation or land use rapidly changes (Agrawal, 2006).

Furthermore, crop classifications that utilize vegetation indices (VI) reduce atmospheric effects and strongly influence the visible and near-infrared reflectance. Temporal vegetation indices are applied to differential crop types to develop crop classifications and products for crop conditions including potential yield maps. Doraiswamy (2007) developed the crop classification in the U.S. Corn Belt by utilizing MODIS imagery. The method of this research has been successfully applied for operational crops to yield prediction for Iowa and Illinois, and will be

expanded to the rest of the U.S. Corn Belt. Consequently, it is affirmed that vegetation phenology of time series VI data has been used widely in land cover classification.

In addition, Doraiswamy (2007) indicated that the 250 m resolution MODIS 8-day composite surface reflectance data (MOD09) are suitable for developing within-season crop classifications in the United States. Crop parameters were developed by the MODIS crop classification model at mid-season to predict grain yields. The corn and soybean crop classification utilizing MODIS data provides an overall accuracy of 75 – 80% of the LANDSAT classification.

Dalstra (2008) also found that EVI is more sensitive to forest and vegetation health classes by applying a multi-temporal remote sensing classification methodology. The supervised maximum likelihood classification of EVI derived from MODIS multi-temporal imagery data provides highly accurate classification results with an overall increase in accuracy of 5%.

Land cover types can be varied from agricultural practices, so land cover features are considered complex. In the case of variation, plant density, ages, and types can enhance spectral confusion for classification, particularly for annual crops. The phenological cycles of different fields generate important shifts between fields or a lot of overlaying signatures between these classes. These problems cannot be addressed by implementing ordinary classification methods such as maximum likelihood. Therefore, with image time series, the crop cycle can be monitored and identified with the discrimination against classes with high accuracy classification (Simonneaux, 2007).

Leite *et al*. (2008) also supported this concept. In an article of crop type recognition based on Hidden Markov Models of plant phenology, their research indicates that a multi-temporal crop classification technique utilizing satellite imagery containing plant phenology with

the Hidden Markov Model (HMM) provides significant potential comparing to a mono-temporal maximum likelihood classification approach. This research identifies different agricultural crops by analyzing the crop specific temporal profiles of spectral features over a sequence of medium resolution satellite images (LANDSAT images). The results show a remarkable superiority of the HMM model of multi-temporal classification with an average accuracy of no less than 93% in the identification of the correct crop.

2.2 Land cover classification in remote sensing

Land cover is one of the most fundamental geographical variables and it plays an important role in geographical inquiry, particularly in resource planning and environmental management (Foody, 1996b). In addition, the changing patterns of land cover reflect the changes in economic, social, and environmental conditions. Monitoring such changes can be important for national and international policymakers, particularly for coordinated actions in environmental fields and model building such as climate and hydrological models (Bernard *et al*., 1997). Subsequently, land use and land cover maps are essential required for scientific research (Atkinson, 2005).

However, providing land cover maps with important, informative, and accurate data is both difficult and expensive (Atkinson, 2005). The quality of the land cover data which are currently used in scientific research is considered inadequate because they land cover data may be spatially incomplete, out-of-date, or inaccurate (Atkinson, 2005). Therefore, frequent up-dating with an accurate classification, especially at regional and global scales, is essential for land cover mapping.

Remotely sensed data have been used to map land cover in various spatial and temporal scales (Foody, 1996b). The wide range of using remotely sensed data is being increasingly applied to land use/land cover classification. With various temporal, spatial, and spectral resolutions, remote sensing is capable of providing land cover information at various scales. Furthermore, multi-temporal images can be used to monitor changes in land cover over time. For these reasons, remote sensing provides great value for land cover mapping and monitoring (Atkinson, 2005). Therefore, it has long been the goal of remote sensing research to improve the accuracy and reduce the time required for image classification.

The accuracy and value of the land cover maps derived from remote sensing depend on a range of factors related to the data sets and methods used. For example, the accuracy of maps derived from conventional supervised image classification techniques is a function of the factors related to the training, allocating, and testing stages of the classification (Foody, 1996b). The conventional techniques for image classification from remotely sensed imagery focus on hard classification (both supervised and unsupervised approaches) in which each pixel is allocated to one class (Atkinson, 2005). The supervised and unsupervised approaches are the basic principles for image classification. The *supervised classification* uses training sites to acquire the spectral signatures of each land cover class in each spectral band. Next, training statistics are used to allocate pixels of unknown class membership to a class in accordance to specific decision rules, and then the quality of the classification is evaluated (Foody, 1996a). The *unsupervised method* analyzes an image in an n-band space in order to group pixels according to a given criteria, then associate such groups with a known land cover class, e.g., the k-means cluster analysis (Bardossy & Samaniego, 2002).

The classical conventional spectral classifiers, such as the maximum likelihood classifier, perform well over limited areas where the spectral signatures do not vary greatly from those captured in the training data. Therefore, variation in plant densities, ages, and types can increase spectral confusion and decrease the accuracy of image classification. Other conventional hard classification techniques, such as minimum-distance and parallelepiped techniques, use the same principle by assigning each pixel to a single class. In reality, many pixels in an image may represent more than one land cover on the ground. To allocate a mixed pixel to a single land cover class not only provides an unrealistic result, but also leads to an inaccurate representation of land cover (Thornton *et al*., 2006).

According to Foody (1996b), mixed pixels are a major problem in land cover mapping applications. This is because the conventional image classification techniques assume that all the pixels within the image are pure, that is, they represent an area of homogeneous cover of a single land cover class. This assumption is often untenable with pixels of mixed land cover composition, which is abundant in an image. For example, while a mixed pixel must contain at least two classes, the classification procedures are generally used to produce a land cover map that force allocation into one class. The relationships between the sensor's spatial resolution and the fabric of the landscape, especially near the boundaries of two or more discrete classes, are also the consequences of the mixed pixel problem. In addition, mixed pixels will also occur where the land cover classes are continuous and inter-grade gradually with many areas of mixed class compositions rather than discrete classes (Foody & Cox, 1994).

In addition to classification methods, the accuracy of land cover maps derived from remotely sensed data depends on the nature of the land cover classes and the spectral and radiometric resolutions of the remotely sensed data (Bardossy & Samaniego, 2002). This is

because the spectral signature of land cover types may vary from microclimatic variations during the growing season, slope and aspect of the ground, or the heterogeneity of materials.

Atkinson (2005), Watanachaturaporn (2005), and Thornton *et al*. (2006) mentioned two primary causes of mixed pixels in image classification:

1) The frequency of sampling afforded by the sensor's spatial resolution is less than or equal to the frequency of spatial variation in land cover. The spectral measurement in this case will be a combination of individual object spectra, particularly sensors such as AVHRR and MODIS.
2) A proportion of pixels will be mixed where the spatial resolution is fine relative to the frequency of variation in land cover because some pixels inevitably straddle boundaries between land cover objects. In this case, mixed pixels may also be presented at the boundaries of the two classes due to the linear features and the presence of small classes within the larger classes. Another case is when materials are combined into a single mixture (e.g., water and soil).

This mixed pixel problem is highlighted when using coarse satellite images such as MODIS data. With coarse resolution images, a large number of pixels may be mixed at the scale of measurement. These mixed pixels reflect the composite spectral responses of the classes within them (Xu *et al.*, 2005). Foody (1996b) also indicated that the proportion of mixed pixels generally increases with a coarsening of the spatial resolution of the sensing system. However, land cover at regional and global scales is required to access land cover change. Coarse spatial resolution sensor data is a possible approach for land cover mapping, although the large

proportions of mixed pixels in these coarse resolution data can lead to significant errors in the estimation of land cover change over time.

Therefore, the existence of mixed pixels leads to the development of several approaches for soft classification in which each pixel is allocated to all classes in varied proportions (Atkinson, 2005). The techniques for soft classification applied to remotely sensed imagery have been referred to as spectral unmixing, spectral decomposition, fuzzy classification, and sub-pixel classification.

The basic principle for soft classification is that the strengths of class membership derived in the classification should be related to its land cover composition (Foody, 1996b). The soft classification process decomposes a collection of class component spectra or endmember into a collection of corresponding fractions or abundances. The proportion of each class or endmember within the pixel is indicated by the abundances (Watanachaturaporn, 2005). A wide range of soft classifiers has been developed for land cover classification such as the linear mixture model (Food & Cox, 1994), maximum likelihood classification (Foody, 1996a, 1996c; Zhang & Foody, 2001; Eastman & Laney, 2002; Ibrahim *et al.*, 2005), fuzzy c-means classification (Foody, 1996a, 1996c; Dai *et al*., 2010), and neural networks (Foody, 1996a, 1996c; Zhang & Foody, 2001; Ibrahim *et al.*, 2005). These techniques provide more informative and potentially more accurate results than the hard classification.

Linear mixture models (LMM) are the most widely used soft classifiers. A LMM is developed based on the assumption that a pixel contains several different classes. The spectral signature of each class is taken to be a multi-dimensional Gaussian distribution. Consequently, this technique is considered as a statistical model. This technique is appropriate when the combination is linear and class components in a pixel appear in spatially segregated patterns. If

the classes are in an intimate association or the spectral mixture is nonlinear (e.g., spectral measurement from a beach), the use of LMM may not be appropriate (Watanachaturaporn, 2005).

The conventional techniques, such as maximum likelihood classification (MLC), can also be used to soften classifiers. This technique depicts the partial and multiple class memberships of each pixel, and assumes that the data follow a Gaussian distribution (Xu *et al.*, 2005; Watanachaturaporn, 2005). However, MLC, a probabilistic classifier, may not always be appropriate for all applications. This approach highly depends on an assumption of the distribution of data. Unfortunately, classes often display non-normal distributions, which can be difficult to correct. Additionally, the size of the training set used to characterize class appearance for the classification is often too small to reliably characterize class appearance (Mather, 1987).

Fuzzy classification techniques are more attractive as the concept of a pixel having a degree of membership to all classes is fundamental to fuzzy-sets-based techniques (Foody, 1996b). Fuzzy c-means (FCM) clustering has been most widely used in remote sensing soft classification. This method is an unsupervised approach where the class membership from FCM has been found to be related to the class composition of a pixel (Watanachaturaporn, 2005). However, this approach is based on the probabilistic constraint that class membership of a pixel across the classes sum to one. In addition, Dai *et al*. (2010) mentioned that the significant problem of this technique is that the probabilistic membership resulting from FCM does not always correspond to the degree of belonging or compatibility of data points with the class prototypes; therefore, the algorithm has considerable trouble in noisy environments.

Amongst classifiers, one particularly attractive approach, which is becoming increasingly popular in remote sensing, is the use of artificial neural networks (ANN). An ANN is a non-

parametric technique, which has been shown to generally be capable of classifying data as or more accurately than conventional classifiers (Foody, 1996b).

2.3 Artificial neural network for remotely sensed image classification

An artificial neural network (ANN) is a simplified version of actual biological neuron cell, with the desire of superior abilities over conventional serial processors in cognitive tasks (Yang, 2005). Kohonen (as cited in Yang, 2005) defined the ANN as "The artificial neural networks are massively parallel interconnected networks of simple (usually adaptive) elements and their hierarchical organizations which are intended to interact with the objects of the real world in the same way as the biological nervous systems do." The process of an ANN is comprised of two characteristics of the human brain: the ability to *learn* and to *generalize* from limited information (Hewitson and Crane, 1994).

2.3.1 The concept and the process of ANN for image classification

The conceptual function of an ANN is operated as a '*black box*' approach, which has a great capacity in predictive modeling. The unknown situation of all characters is served as the input to train an ANN, and the identification (prediction) is then generated (Lek and Guegan, 1999). A black box with input and output performs certain functions to map the input with the output. The first step is to run the untrained net in a random state to represent a random function and then train the net to learn some mapping relationship between the input and output. This step is accomplished by applying learning algorithms to process data from the sample of known input and output and modifying the internal function performed by the net to find a relation between the input and output. Then the training samples are calculated in the learning process until the net can be applied in a similar manner to the further unknown data (Hewitson and Crane, 1994).

According to Yang (2005), there are three basic elements in a simple neuron model as follows: 1) a set of synapses with weight vector connects the input vector to the neurons, 2) the weighted input signal is summed and 3) passes through the transfer characteristic or the activation function. A threshold function adjusts the weighted input signal level before passing through the transfer characteristic. Therefore, an ANN has input paths, output paths, and connecting weights.

A typical ANN for image classification is a *multilayer perceptron (MLP) neural network* (Mill et al., 2006; Lek & Guegan, 1999). The architecture of this model composes of a set of nodes, which are usually partitioned into different layers (input, output, and hidden layers) and fully connected together if two nodes are within neighboring layers (Yang, 2005 & Ke et al., 2008). Generally, for image classification, the number of nodes in the input layer is determined by the number of input bands, the number of output nodes is dependent upon the number of land cover classes in the classification scheme, and the number of hidden nodes is related to the optimal design of ANN (Mill et al., 2006). Each node can have incoming weight connections from the previous layer and outgoing weight connections to the next layer (Ke et al., 2008). The multilayer feed forward networks with a sufficient number of hidden nodes between the input and output units have a "*universal approximation*" property; in other words, they can approximate virtually any function of interest to any desired degree of accuracy (Li, 2007).

The training process is the most important step for an ANN and the objective of training is to achieve the proper weights both for the connections between the input and hidden layers, and between the hidden and the output layers for the classification of the unknown pixels. *The back-propagation learning algorithm* is generally used to train the network (Schalkoff et al., 1992; Lek & Guegan, 1999; Li, 2007; Li, 2008; Ke et al., 2008). The *back-propagation neural*

network (BPNN) is a layered feed forward neural network, in which the non-linear elements (neurons) are arranged in successive layers, and the information flows unidirectionally, from input layer to output layer, through the hidden layers.

The neural network training in this algorithm is started at the input layer. Training pixels are fed through the network and network outputs are compared with the target outputs, which are known for training pixels. The error, if any, is then propagated backward through the network to the input layer with the weights for relevant connections corrected via a relation equation. The training data are then entered again and the process repeated until the overall error is minimized or declines to an acceptable level. Li (2007) mentioned that MLP can process both hard and soft classifications. In the case of hard classification, the input pattern is assigned into the class that is associated with the neuron that has the highest activation level. In terms of soft classification, the membership of an input pattern belonging to each potential class is expressed as the activation level of the output layer neurons.

2.3.2 Advantages and capabilities of ANN

Nelson and Illingworth (1991), Hrycej (1992), Villmann et al. (2003), and Yang (2005) stated that the capabilities of the ANN approach can be defined in several characteristics which make ANNs attractive, as follows:

- *Adaptive learning*: An ability to learn how to do tasks based on the data given for training or initial experience. Adaptability is one of the most significant features of an ANN and the capability for self-adjustment. The adaptability process consists of three aspects: example-based learning, generalization capability, and format- free input data. In addition, an ANN can automatically adjust their connection weights or

network structures (number of nodes or connection types), to optimize their behavior as controllers, predictors, pattern recognizers, and decision makers.

- *Self-Organization*: An ANN can create its own organization or representation of the information that it receives during the period of learning time.
- *Real Time Operation*: An ANN computation may be carried out in parallel and special hardware devices, which are designed and manufactured to take advantage of this capability.
- *Fault Tolerance via Redundant Information Coding*: Partial destruction of a network leads to the corresponding degradation of performance. However, some network capabilities may be retained even with major network damages.
- *Generalization:* Generalization is the ability of the network to respond to input that it has not seen before. Although the input is partial, incomplete, fuzzy, ambiguous, or contains partially corrupted data, an ANN can deal with these situations using the characteristics of intuition, prediction and statistical pattern reconstruction. With the generalization approach, an ANN is obviously appropriate for real-world data. In these cases an ANN is capable of using similar inputs or situations for output or inferences that are also similar.
- *Parallel processing:* Since ANN implementation is considerably difficult to speed up in a single processing unit, the only alternative solution is to distribute computationally expensive tasks to work in parallel. This property of ANNs makes the inherent parallelism of virtually all neural network algorithms able to be updated simultaneously.

Several studies show the advantages of ANNs as described below (Foody, 1996a; Carpenter *et al*., 1997; Mills et al., 2006; Bagan et al., 2005):

1) ANNs make no a priori assumption about data distributions. Consequently, it is able to learn nonlinear and discontinuous data samples.
2) ANNs can readily accommodate ancillary data such as textural information, slope, aspect, and elevation.
3) An ANN is typically more accurate than conventional classifiers; an ANN can improve classification accuracy by 10-30% compared to conventional classification techniques.
4) ANN architectures are quite flexible and can be adapted to improve performance on particular problems.
5) ANNs have been proved to be successfully applied to land cover mapping from satellite remote sensing data, both hard and soft classifications.

In conclusion, ANNs offer a number of advantages such as requiring less formal statistical training, the ability to implicitly detect complex nonlinear relationships between dependent and independent variables, the ability to detect all possible interactions between predictor variables, and the availability of multiple training algorithms (Tu, 1996). An ANN is a parallel distributed processor that has a natural tendency for storing experiential knowledge. It can provide suitable solutions for problems, which are generally characterized by non-linear ties, high dimensionality noisy, complex, imprecise, and imperfect or error prone sensor data, and lack of a clearly stated mathematical solution or algorithm. A key benefit of an ANN is that a

model of the system can be built from the available data (Seetha et al., 2008). Li & Eastman (2006) and Foody (1997) demonstrated that ANNs have been of considerable interest for the classification of remotely sensed imagery because of their freedom from assumptions about the form and distribution of input data, their ability to generate non-linear decision boundaries, and their ability to generalize inputs as well as to learn complex patterns.

2.3.3 Limitations of ANN

Although there are several benefits obtained from ANNs, the major limitations for image classification were addressed by Seetha et al. (2008). They mentioned that a backpropagation learning algorithm is the optimization tool for neural network training but this technique has several problems such as premature convergence and efficiency of differential operation. Additionally, ANNs have been claimed to be a difficult technique for understanding the structure of the algorithm. Although an ANN has the advantages mainly of more tolerance to noise inputs and the representation of boolean functions apart from others, many attributes may result in over-fitting. ANNs follow a non-parametric approach for image classification but the selection of the non-linear boundary is efficient when the data have only few input variables in the ANN. The accuracy of results and training speed in the neural networks depends on network structure, momentum factor, learning rate, and convergence criteria. These optimal parameters can only be determined by experimentation. Tu (1996) criticized that the black box nature, greater computational burden, proneness to over-fitting, and the empirical nature of model development are the major disadvantages of ANNs.

According to Foody (1996b), there may be problems associated with training ANNs, particularly in relation to over training and training time. However, an ANN, once trained, may

classify data extremely rapidly because the classification process may be reduced to the solution of a large number of extremely simple calculations, which may be performed in parallel.

2.3.4 Application of ANN classification in remote sensing

Many studies of ANNs in remote sensing focus on the recognition of land cover classes using both supervised and unsupervised classification. ANNs show high performance capacity for incorporating different types of data. Key et al. (1989) and Maslanik et al. (1990), for example, studied the ability of ANNs to classify merged images of summer arctic data from 5-channel Advanced Very High Resolution Radiometer (AVHRR) and 2-channel Scanning Multi-channel Microwave Radiometer (SMMR). They found that ANNs are easier to use and to interpret than other approaches. They also suggested that the understanding of the physical process reflected in the under-investigated data is necessary to effectively design an ANN and to interpret its results. Heermann and Khazenie (1990) revealed the suitability ANNs for the classification of multi-spectral and multi-source remote sensing data. Although the result shows that ANNs did not improve the accuracy, the accuracy of an ANN classification increases as the absolute (not the percentage) size of the training dataset increases. Benediktsson et al. (1990) compared the performance of ANNs to those of a variety of statistically-based classifiers for the classification of multi-source remote sensing and geographic data. They stated that ANNs are distribution free; therefore, they could use the ancillary data without worrying about ranking or weighting them. Civco and Wang (1994) developed ANN techniques to process LANDSAT TM data from two acquisition dates, two channels of illumination data, and a measure of image texture to derive more accurate land use and land cover information. They claimed that using the enhanced ANN technique is more accurate than using single-date LANDSAT TM data due to its ability to handle multi-spectral, multi-temporal, multi-source spatial data more efficiently than

parametric statistical methods. Bischof et al. (1992) compared the results of ANN classifications of seven-band LANDSAT TM data, on a pixel by pixel basis, to those of the maximum likelihood (ML) classifier. They found that ANN outperformed the ML classifier. The ANN achieves an 85.9% overall accuracy versus 84.7% for the ML classifier. They also stated that the two-layer ANN is able to smooth the resulting classified image. Crane (1992) performed post processing editing on a classified image using an ANN. He utilized spatial information to correct misclassified pixels in a large classified LANDSAT TM scene with an abundance of lakes, marshes, small wetlands and rich soils. He concluded that the ANN is able to learn the characteristics of the spatial data and thus the overall accuracy is improved.

Atkinson and Tatnall (1997) also mentioned that the number of applications of ANN classification increases rapidly because of its capabilities to perform more accurately and more rapidly than other techniques such as statistical classifiers, particularly when the feature space is complex and the source data has different statistical distributions. ANNs incorporate *a priori* knowledge, realistic physical constraints, and different types of data (including those from different sensors) into the analysis, thus facilitating synergistic studies.

Gopal et al. (1999) also indicated that the accurate classification results of the ANN-based technique is due to multiple factors: 1) neural network classifiers are distribution-free and can detect and exploit nonlinear data patterns, 2) neural network classification algorithms can easily accommodate ancillary data, 3) neural network architectures are quite flexible and can be easily modified to optimize the performance, and 4) neural networks are able to handle multiple subcategories per class.

Multispectral image information can sometimes be insufficient for differentiating species-level land cover classes because of the effects of local topography, background

reflectance from soils or understory land cover, and high within-class variance due to the structure and patchiness of vegetation canopies. Therefore, ancillary data have often been used to help differentiate vegetation types in land cover mapping. Phenological information, which provides seasonal characteristics of different land cover types, is a key ancillary data source for land cover classification. ANNs be adapted to use these spectral and ancillary data in order to improve classification performance.

Additionally, ANNs are utilized for soft classification and being developed into different techniques. Multi-layer perceptron (MLP) neural networks have been widely used in many remote sensing studies (Watanachaturaporn, 2005). This technique is different from LMM and MLC, which are statistical methods, because MLP does not assume that the data follow a probability distribution. Fuzzy ARTMAP, mixture discriminant analysis (MDA), counter propagation network (CPN), and regression tree algorithm (RTA) are other classifiers for soft classification. In remote sensing studies, it has been found that these techniques were limited to some applications (Watanachaturaporn, 2005). However, the self-organizing map (SOM) neural network demonstrates the great potential for soft classification and overcomes the weaknesses and limitations of ANNs.

2.4 Self-organizing map (SOM) neural network

There has been considerable interest in applications of ANNs for remotely sensed image classification; therefore, several techniques have been developed including the self-organizing map (SOM). SOM developed by Kohonen (1989, 1990) is a prominent unsupervised and nonparametric neural network approach. The original concept of SOM is based on competitive learning in which lateral interaction in the output layer self-adaptively leads to regional

organizations of neurons (a topology) that become special detectors for different signals such as land cover classes (Bagan *et al*., 2005). The output layer links with input vectors by random weights and the adjustment of weights is spread spatially to neighboring neurons using a distance decay function (Li & Eastman, 2006). In the last step, neurons are organized into clusters of association with input vectors.

The basic architecture of SOM is shown in Figure 2.1. According to Kohonen (1989, 1990), Hagan *et al.* (1996), Bagan *et al*. (2005) and Li & Eastman (2010), the input layer represents the input feature vector and contains neurons for each measurement dimension. For example, a separate neuron in remotely sensed data refers to each reflectance band. The output layer (or competitive layer) of SOM is typically organized as a two-dimensional array of neurons. A set of neurons has the neighborhood relationships among the neurons. Synaptic weights function as the connection between all neurons in the input layer and each output layer. The synaptic weights are initialized to random values from 0 to 1. The weights are adjusted in the learning procedure according to normalized input feature vectors and lateral interaction between neurons in the output layer. The lateral interaction varies in the manner of a distance decay function. Therefore, each input vector is assigned to the neuron with the nearest weight vector. When the process finds the winning neuron, the weight of the winning neuron and all neurons in the neighborhood of the winning neuron are updated. The SOM is able to divide the input space into regions with common nearest weight vectors. Finally, input patterns with similar attributes will be clustered spatially in the output layer.

SOM has been applied in supervised classification research and showed effective results in image classification. Although MLP neural networks are widely applied to image classification, there are limitations with this technique. MLPs use a multilayer feedforward

approach with a sufficient number of hidden nodes between the input and output units. A variable number of hidden layers in MLPs are organized to accommodate the complexity of hypersurfaces needed to separate the input classes. On the other hand, SOM has only two layers (an input and an output layer) with an emphasis on lateral organization. Therefore, the capability of SOM is to map high dimensional input vectors onto an array of low dimensional units and to preserve the topology of the input pattern in the low dimension after dimension reduction (Bagan *et al*., 2005). With this capability, SOM is very useful for analyzing large datasets because this technique can produce both a reduced amount of data by clustering and a projection of dominant data patterns on a lower-dimensional display.

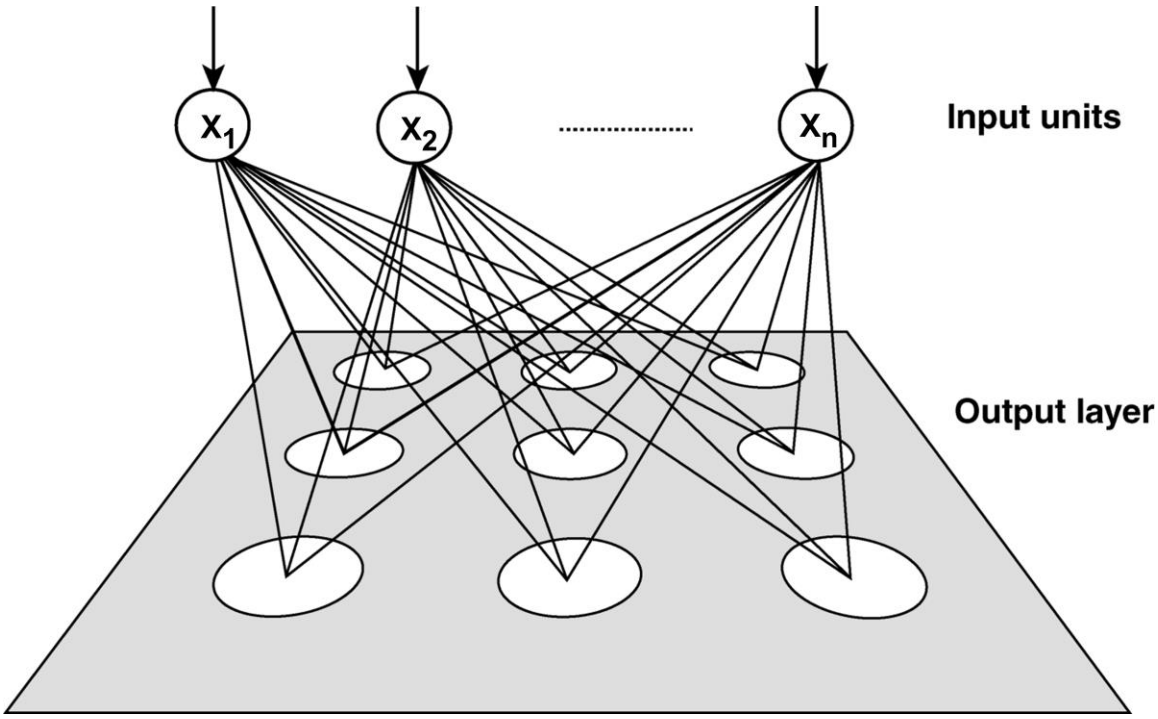


Figure 2.1. An example of the architecture of a SOM.

SOM is similar to Fuzzy ARTMAP in that it uses multivariate automated procedures for cluster analysis. However, each neuron in the output layer of SOM relates to a fixed position in a two-dimensional plane (n x m neurons), whereas the output layer of Fuzzy ARTMAP is one-dimensional and has no topology among them (Li, 2007).

Bagan *et al*. (2005), Li & Eastman (2006) and Hu (2009) developed a supervised SOM algorithm for effective image classification based on *coarse-tune and fine-tune* processes. There are four steps in this supervised classification. The first step is the unsupervised clustering process where the network training in SOM implements the coarse tuning. In this coarse tuning stage, the weights are adjusted based on the normalized input feature vector and the lateral interaction between neurons in the output layer. The radius of the zone of lateral interaction will decrease. While neuron weights, which represent the underlying clusters and sub-clusters in the input data, are generated, input patterns with similar attributes also are clustered in the output layer. The second step is "code book labeling". During this step, majority voting takes place to identify a class for each output neuron and to establish identities of regional groups by comparison with training data. When a group of neurons is labeled with a single information class, a group of neurons will commonly cover the range of variability in the reflectance associated with the information class. The third step is the fine tuning. In this stage, the weight vectors are adjusted by a Learning Vector Quantization (LVQ) algorithm to improve the discriminability of decision boundaries. The specific boundaries between neurons based on specific information classes are defined by using training site data. Finally, in the classification step, each image pixel is assigned the class label of the neuron that is most similar in its weight structure to the pixel vector of reflectance.

With the coarse-tune and fine-tune processes, Bagan *et al*. (2005) employed the supervised SOM technique to classify land cover types using a 17-dimensional dataset that was generated from 16-day interval MODIS-EVI data with a spatial resolution of 500 m in eastern China during the growing season. The accuracy of SOM is higher than the conventional method (MLC) that uses high-dimensional MODIS time-series data. The research of Li & Eastman (2010) showed successful modification of SOM for image supervised classification by using SPOT and AVIRIS (hyperspectral image). SOM Commitment (SOM-C) and SOM Typicality (SOM-T) in this research outperformed a parametric Bayesian posterior probability classifier and Mahalanobis typicality classifiers.

Alternatively, Liu *et al*. (2010) introduced a new method for supervised SOM based on a *tagging* technique by using synthetic data experiments and hyperspectral remote sensing imagery. The results demonstrated that SOM is suitable for the decomposition of mixed pixels in hyperspectral images, particularly for nonlinear spectral mixing. In addition, the learning process in supervised SOM can avoid the issue of the local optimum, which is the main problem of other techniques, such as Fuzzy C-means and MLP.

However, these improved supervised SOM techniques have some limitations such as unlabelled neuron unclassified pixels, and they can be very time consuming.

2.5 Significance of the study

Several approaches based on ANN classification have been developed to improve classification accuracy, particularly SOM. SOM with supervised classification has been proven to be an effective technique to enhance the accuracy of image classification. However, few studies have utilized SOM for supervised classification. Additionally, when the SOM is

associated with a supervised classification, a majority voting technique is usually used to associate these neurons with training data classes. However, this technique cannot guarantee that every neuron in the output layer will be labeled, and thus causes unclassified pixels in the final map (Li & Eastman, 2006).

Bagan *et al*. (2005) also implemented supervised SOM based on the coarse-tune and fine-tune technique with a majority voting principle and they solved the problem of unlabelled neurons by selecting a suitable threshold value to label all neurons under the condition of unclassified neurons. Li & Eastman (2006) proposed the auxiliary labeling algorithm to assign unlabelled neurons to the clusters already formed from the supervised stage. Although these studies illustrated how to solve the problem of unlabelled neurons, whole input data fed into the coarse-tuning stage for unsupervised classification can be considerably time consuming.

In addition to the coarse-tune and fine-tune technique, Liu *et al.* (2010) demonstrated the capabilities of a tagging technique with fuzzy membership for supervised SOM classification. Although this method showed successful results for the decomposition of mixed pixels in hyperspectral images, using a Fuzzy C-means function in the classification stage may lead to the problem of information loss due to a hard neural network topology, and this technique is considered as partially-soft classification.

This research attempted to develop a technique to improve classification accuracy. The objective of this study is to improve SOM to soft supervised classification using phenological information from satellite time-series data in order to overcome the mixed pixel problem. The new technique, *soft supervised self-organizing map (3SOM)*, is proposed to improve image classification. The innovative method applies a *"Class information attaching"* technique to solve the problem of unlabelled neurons. This approach directly obtains class information membership

from the output layers; therefore, the classification process is faster with no information loss. Additionally, it has a soft neural network topology to support a fully soft classification. The new approach based on the phenological information will be able to extract effective texture information from satellite imagery, which benefits regional scale land use/land cover classification and change detection as well as contributes to various disciplines such as map updating, agricultural area estimation, cartography, and urban planning.

Chapter 3

Research Methodology

3.1 Research dataset and study area

3.1.1 Synthetic remotely sensed data

Synthetic data are generally preferred of testing new classification algorithms because the actual class proportions of each pixel in these images are known beforehand for validation and accuracy assessment purposes. Therefore, a synthetic image is generated to facilitate the design and development of new classification algorithms and to conduct experiments on soft supervised self-organizing map (3SOM) classification. By breaking down the elements of real world imagery into simplified representations, understanding and improving such an image processing technique becomes easier.

The synthetic data is generated based on a time-series of MODIS-EVI image (23 dates per year). The process is shown in Figure 3.1. To reduce computational time, the synthetic image is relatively small in size, corresponding to 50 x 50 pixels in 23 layers. Each layer consists of four assumed land cover types derived from the MODIS-EVI values of pure pixels located within large homogeneous areas. The four land cover classes identified from the MODIS-EVI data are verified through land cover reference images.

The mean and standard deviation are extracted from each class of pure pixels as shown the details in Appendix A. Figure 3.2 shows the mean standard EVI temporal profiles of four land cover classes. Class A tends to have the highest EVI profile, while the lowest EVI profile is obtained from Class D. The similarity of class B and C pose challenges when attempting to distinguish between these two classes.

Four images, one corresponding to each land cover class, are created through a random number generator based on a normal distribution using the extracted mean and standard deviation of each class. Once the images are generated for the four classes, the synthetic EVI values for each pixel are derived using a weighting scheme according to the Linear Mixture Model (LMM) given by Equation (1). with known class proportions distributed as shown in Figure 3.3. and Table 3.1.

$$X = \sum_{i=1}^{c} \rho_i S_i \qquad (1)$$

In the LMM equation, X is a mixed value for an individual pixel, ρ_i is the class proportion value of class i (shown in Figure 3.3.), S_i is the pure value of class i (obtained from the randomly generated land cover class datasets), and c is the number of classes.

Table 3.1. The set of proportions corresponding to each index zone from Figure 3.1.

	Class Proportion					**Class Proportion**					**Class Proportion**			
ID	**A**	**B**	**C**	**D**	**ID**	**A**	**B**	**C**	**D**	**ID**	**A**	**B**	**C**	**D**
1	1.0	0.0	0.0	0.0	13	0.0	0.0	0.8	0.2	25	0.5	0.1	0.3	0.1
2	0.0	1.0	0.0	0.0	14	0.0	0.0	0.6	0.4	26	0.4	0.2	0.2	0.2
3	0.0	0.0	1.0	0.0	15	0.0	0.0	0.4	0.6	27	0.2	0.4	0.2	0.2
4	0.0	0.0	0.0	1.0	16	0.0	0.0	0.2	0.8	28	0.1	0.5	0.1	0.3
5	0.8	0.2	0.0	0.0	17	0.8	0.0	0.2	0.0	29	0.3	0.1	0.5	0.1
6	0.6	0.4	0.0	0.0	18	0.6	0.0	0.4	0.0	30	0.2	0.2	0.4	0.2
7	0.4	0.6	0.0	0.0	19	0.4	0.0	0.6	0.0	31	0.2	0.2	0.2	0.4
8	0.2	0.8	0.0	0.0	20	0.2	0.0	0.8	0.0	32	0.1	0.3	0.1	0.5
9	0.0	0.8	0.0	0.2	21	0.7	0.1	0.1	0.1	33	0.1	0.1	0.7	0.1
10	0.0	0.6	0.0	0.4	22	0.5	0.3	0.1	0.1	34	0.1	0.1	0.5	0.3
11	0.0	0.4	0.0	0.6	23	0.3	0.5	0.1	0.1	35	0.1	0.1	0.3	0.5
12	0.0	0.2	0.0	0.8	24	0.1	0.7	0.1	0.1	36	0.1	0.1	0.1	0.7

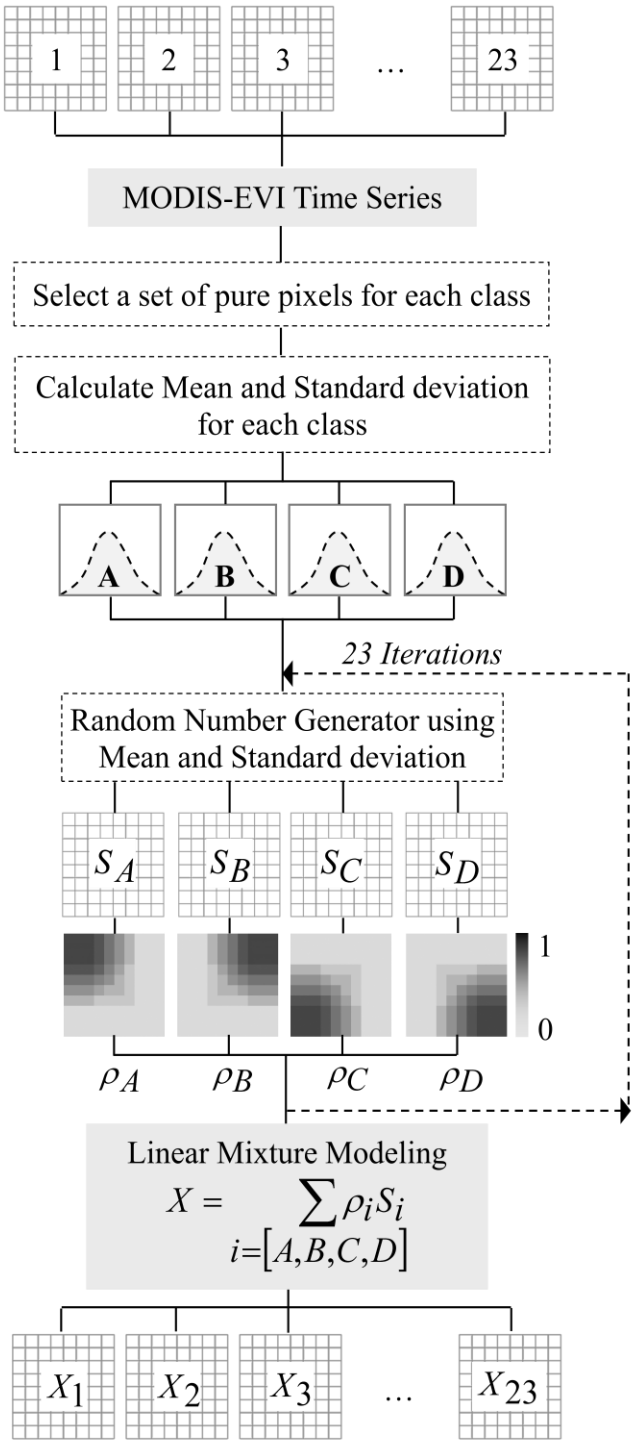


Figure 3.1. The process to simulate the remotely sensed synthetic data.

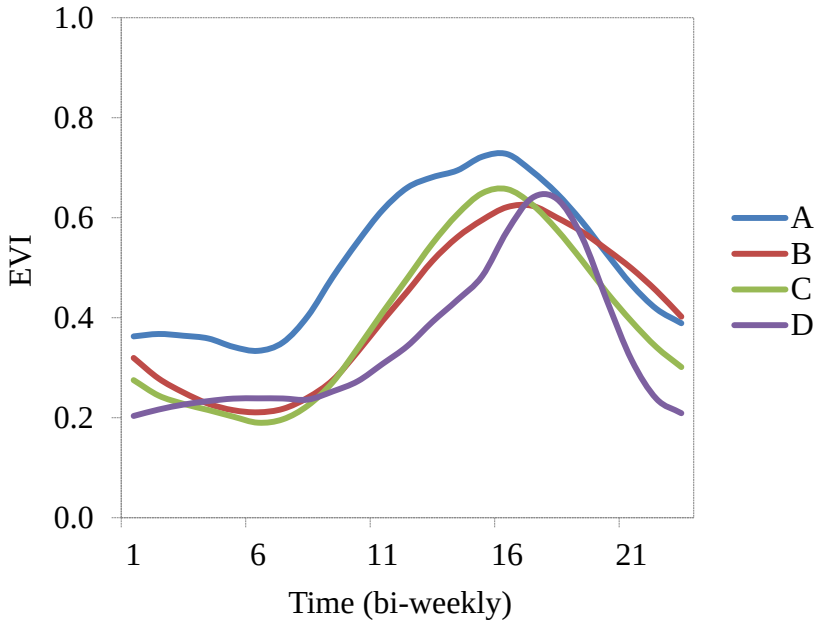
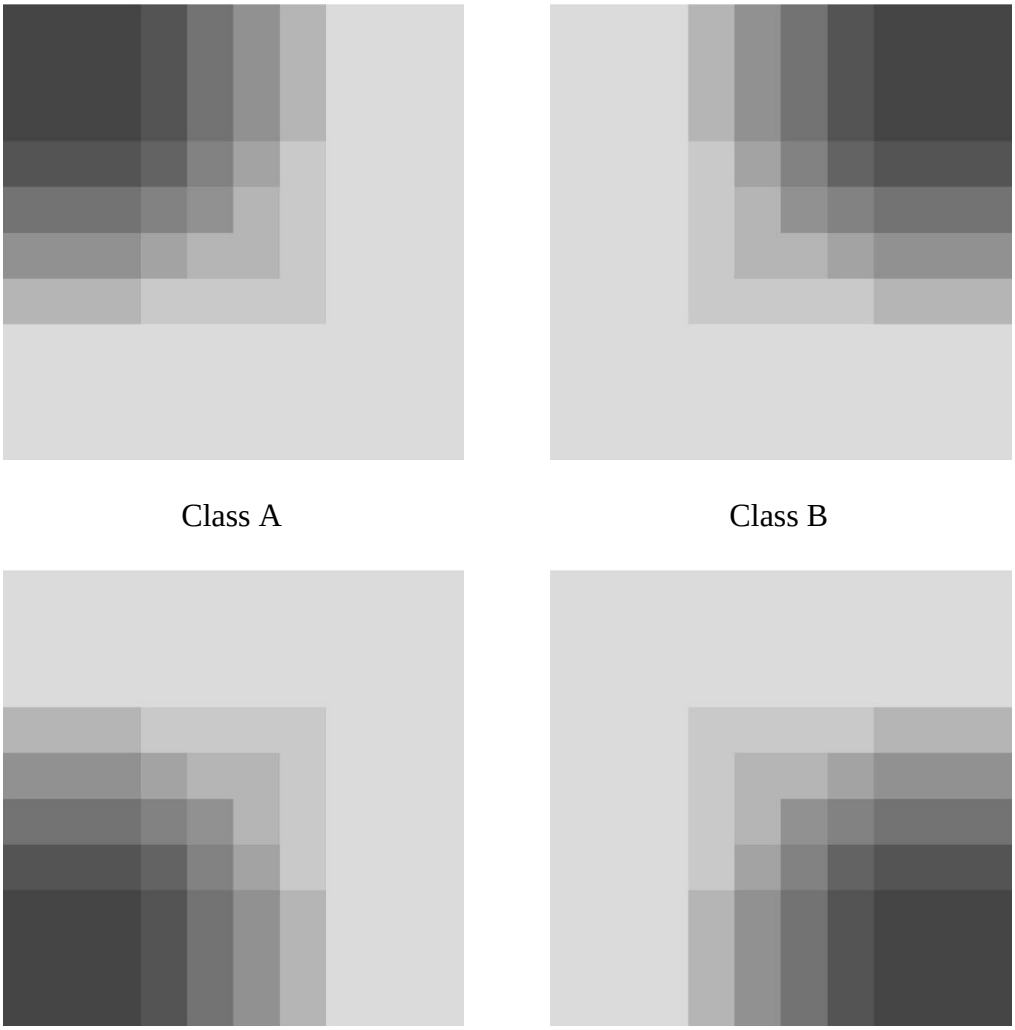


Figure 3.2. The standard EVI temporal profiles for each land cover class.
"For interpretation of the references to color in this and all other figures, the reader is referred to the electronic version of this dissertation."

1	1	1	5	6	7	8	2	2	2
1	1	1	5	6	7	8	2	2	2
1	1	1	5	6	7	8	2	2	2
17	17	17	21	22	23	24	9	9	9
18	18	18	25	26	27	28	10	10	10
19	19	19	29	30	31	32	11	11	11
20	20	20	33	34	35	36	12	12	12
3	3	3	13	14	15	16	4	4	4
3	3	3	13	14	15	16	4	4	4
3	3	3	13	14	15	16	4	4	4

Figure 3.3. Index zones representing the class proportions of 5 x 5 blocks of pixels in a 50 x 50 pixel synthetic image.

The synthetic image is obtained by weighting the known class proportions. The darker pixel refers to the higher proportion of a class. Thus, the black color represents a pure pixel of a class (i.e. 100% class proportion). The individual class proportion images are also known as fraction images shown in Figure 3.4. These images represent actual class proportions in the synthetic data and, thus, provide the soft reference data.



Class A Class B

Class C Class D

Figure 3.4. The individual class proportion images of reference image.

3.1.2 Real remotely sensed data

The remotely sensed images used in this research are MODIS 16-day composites of EVI with a spatial resolution of 250 m (MOD13Q1) in 2010, acquired through the NASA Distributed Active Archive Center (DAAC), EROS Data Center. These data are made available to the public free of charge and distributed by the USGS Land Processes Distributed Active Archive Center (LP DAAC, 2008). MOD13Q1 offers consistent spatial and temporal comparisons of vegetation conditions where the MODIS daily vegetation indices are determined by the blue, red, and near-infrared reflectance. The canopy background variations are minimized while the sensitivities over dense vegetation conditions are maintained in the MODIS-EVI images.

In this research, two study areas are used to confirm the classification performance: the Midwestern US and Thailand. These two study areas provide significantly different land cover characteristics. The vegetation development in Thailand is quite heterogeneous – a consequence of the mixture of land cover types. The agricultural areas in Thailand are small in size and exhibit diverse topological features, which often times cause each single pixel to contain more than one land cover classes. On the contrary, the agricultural areas in the Midwestern U.S. are large, quite homogeneous and demonstrate fewer topological differences. Therefore, in this study area a MODIS pixel generally contains only one land cover type, such as large planting corn or soybean fields. Nevertheless, there are still some challenges associated with the data corresponding to the agricultural areas in the Midwest. This is because the EVI temporal profiles of both corn and soybean are similar which results in spectral and temporal confusions.

3.1.3 Reference land cover data

The reference data utilized to validate the classification accuracy include:

1) The dataset of Thailand's National Land Use from the Land Development Department, Ministry of Agriculture, Thailand. This data is generated from the digital color aerial images at a scale of 1:4,000 from 2004, SPOT-5 images with a spatial resolution of 5 m from 2007, and THEOS images with a spatial resolution of 2 m from 2010. The digital color aerial images were geo-referenced and ortho-rectified using field-collected ground control points (GCPs) and fine-resolution digital elevation models (DEM). The land cover data was generated in a vector format at a scale of 1:50,000 by conducting a visual interpretation of digital aerial images, which are updated by SPOT-5 and THEOS images in association with field support data.

2) The U.S. Cropland Data Layer (CDL) from 2010 derived from the LANDSAT 5-TM. This data is published by National Agricultural Statistics Service (NASS), which is part of The United States Department of Agriculture (USDA). The CDL is available in raster format with a spatial resolution 30 m.

3.2 Data filtering and phenological parameters extraction

3.2.1 Data filtering

There are always some errors in standard MODIS reflectance data products associated with georeferencing, cloud contamination, atmospheric conditions, and bidirectional effects (Doraiswamy & Stern, 2007; Jin & Sader, 2005). To eliminate these unfavorable factors, a

Savitzky–Golay filtering technique (Jonsson & Eklundh, 2004) is applied to the EVI time series data to remove the spikes and irregular values of the original image.

Based on the moving window using a simple least-squares filter described by Savitzky and Golay (1964), an adjusted Savitzky-Golay filter was proposed by Chen *et al*., (2004), which applies a weighted moving average filter to an NDVI time series, with the weighting given as a polynomial of a particular degree. A polynomial least-squares fit is applied within the filter window by the weight coefficients. The width of the moving window determines the degree of smoothing, but it also affects the ability to follow a rapid change. This research applies an adjusted Savtizky-Golay filter to the MODIS-EVI time series within the TIMESAT 2.3 program (Chen *et al*., 2004, Jönsson & Eklundh, 2006). The filter can be generally described by Equation (2):

$$EVI_j^* = \sum_{i=-n}^{n} c_i EVI_{j+i} \qquad (2)$$

where EVI_j^* is the j^{th} new filtered EVI value of the window, EVI_j is the original EVI value, the smoothing window size (filter size) is N, which consists of $2n+1$ points, and c_i is the coefficient for the i^{th} EVI value of the filter (Chen *et al*., 2004).

Also within the TIMESAT program, a quadratic polynomial is fit to all points in the moving window, replacing the EVI value at each data point with that of the polynomial. The resulting fitted curve is referred to as the EVI profile (Jönsson & Eklundh, 2006).

3.2.2 Phenological parameters extraction

To extract meaningful phenological information about the vegetation growing season, it is necessary to generate a smooth time-series from noisy satellite sensor data as described above. TIMESAT is a program for analyzing such satellite time-series data. It implements a processing

method to estimate various phenological parameters from the EVI profile. The TIMESAT program consists of eleven general phenological parameters extracted in the following series: (1) start of season, (2) end of season, (3) length of season, (4) base value, (5) position of middle of season, (6) maximum of fitted data, (7) amplitude, (8) left derivative, (9) right derivative, (10) large integral, and (11) small integral (Jonsson & Eklundh, 2006).

Figure 3.5 shows the growing season of an agricultural crop. The beginning of the season, marked by (a) in the figure, is defined from the fitted function as the point in time for which the value has increased by a certain proportion. This value is currently set to 10% of the distance between the left minimum level and the maximum level. The end of the season (b) is defined in a similar way. The middle of the season is difficult to define, but a reasonable estimate is obtained as the position (e) between the positions (c) and (d) for which the value of the fitted function has increased to 90% of the distance between the left and right minimum levels and the maximum.

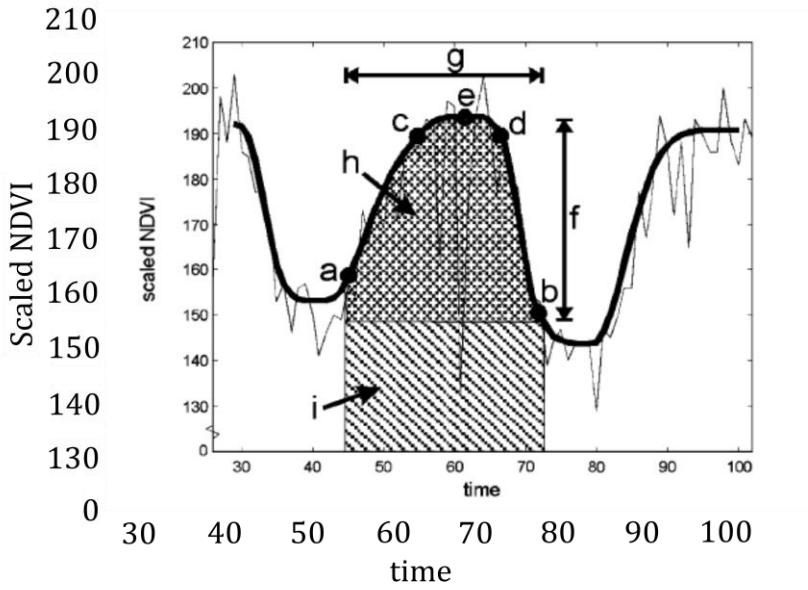


Figure 3.5. A simple NDVI profile for a typical patch of vegetation (Jonsson & Eklundh, 2004).

The amplitude (f) of the season is defined as the difference between the peak value and the average of the left and right minimum values. The first (i.e., small) integral (h), given by the area of the region between the fitted function and the average level of the left and right minima, represents the seasonally active vegetation, which may be fairly small for evergreen areas. The second (i.e., large) integral (i), given by the area of the region between the fitted function and the zero level, represents the total vegetation production. In evergreen areas the first integral may be small even if the total vegetation production is large (Jonsson & Eklundh, 2004).

3.3 Selection of training and testing data

In this study, two sets of randomly sampled pixels are selected from the reference data for training and testing purposes. They include hard training data (only pure pixels) and soft training data (mixed and pure pixels). For comparative purposes, both sets have the same sample size. Generally, soft classifications deal with class mixing in only the classification stage, but do not accommodate class mixing in the ground reference data used in training and testing stages (Foody 1995, 1996a, 1997). Such a classification may be termed partially-soft classification (Zhang & Foody, 1998; 2001) because class mixing is not fully accommodated throughout the classification. Fully-soft classification, which employs both pure and mixed pixels for training and testing stages, is also investigated in this study to test the performance of classifiers.

Only hard training and testing data are used to train the classifiers and validate the outputs of the hard classification, while both hard and soft training and testing data are used to compare fully-soft and partially-soft classifications.

Li and Eastman (2010) mentioned that the training sample size of the SOM is not as sensitive as other neural network models, and that SOM does not need a very large number of training samples to compensate for the high spectral dimensionality.

In this research, both the hard and soft training data have a sample size of 240. Hard and partially-soft classifications are trained by all samples (100% training data) for only pure pixels, while fully-soft classification is trained by 96 samples (40% training data) for pure pixels and 144 samples (60% training data) for mixed pixels.

Moreover, it is important to investigate the performance of a classifier by using different training data due to a significant impact of training samples on the performance of a classifier (Kavzoglu & Mather, 2003), particularly neural network classifiers. Therefore, 500 different random training datasets were generated by a random selection of samples to train the classifier in this study.

3.4 Analysis of the neural network architecture and internal parameter values

The performance of neural network-based classifiers significantly differs depending on the setting of the network structure and internal parameter values. This is because the speed and effectiveness of the learning process are determined by the network architecture and internal parameter values. To achieve high classification accuracy, some adjustments to the network structure and the parameter values should be implemented.

Since there is no standard procedure for choosing the suitable configuration, both series of configurations are set and run by trial-and-error on a case-by-case basis. All trials are carried out on the same training and testing data. The accuracy of a backpropagation neural network

(BPNN) and the supervised self-organizing map (SSOM) in different configurations are evaluated. The configuration that provides the highest accuracy is selected as the suitable one.

In this study, the performance of the (BPNN) and the (SSOM) are examined by setting different values for each parameter. In order to examine the performance of the BPNN, 455 neural network configurations are formed based on primary parameters which have been identified by Kavzoglu & Mather (2003). These parameters include seven different numbers of hidden layer neurons (HN), thirteen learning rate (LR) and momentum factor (MF) values, and five numbers of iterations (ITER) as shown in detail in Table 3.2, 3.3 and 3.4. Furthermore, the performance of SSOM is studied by conducting 300 different neural network configurations. The conduction is done by six different numbers of competitive layer neurons (NET). For each NET, 10 trials with different initial LRs are performed and five different ITER variants of (ITER) were examined (Table 3.5).

Table 3.2. Parameters and values used to investigate the suitable configuration of BPNN.

Parameters	**Values**
Number of input layer neuron	23
Number of output layer neuron	4
Number of hidden layer neuron	4, 14, 15, 46, 47, 54, 69
Initial weight	0.0 – 1.0
Learning rate & momentum factor	(0.01,0.00005), (0.05, 0.0), (0.05, 0.5), (0.1, 0.0), (0.1, 0.3), (0.1, 0.9), (0.15, 0.075), (0.2, 0.0), (0.2, 0.6), (0.25, 0.9), (0.3, 0.6), (0.5, 0.9), (0.8, 0.0)
Iterations	50, 100, 200, 500, 1000

Table 3.3. Heuristics proposed to compute the optimum number of hidden layer nodes (Kavzoglu and Mather, 2003).

Heuristic	**Hidden nodes**	**Reference**
$2N_i$, or $3N_i$	46 or 69	Kanellopoulos and Wilkinson (1997)
$3N_i$	69	Hush (1989)
$2N_i+1$	47	Hecht (1987)
$2N_i/3$	15	Wang (1994)
$(N_i+N_o)/2$	14	Ripley (1993)
$N_i/[r(N_i+N_o)]$	4	Garson (1998)
$(2+N_o \cdot N_i+0.5N_o(N^2_i+N_i)-3)/N_i+N_o$	54	Paola (1994)

Table 3.4. The configurations of learning rate and momentum factor (Kavzoglu and Mather, 2003).

Learning rate	**Momentum factor**	**Reference**
0.01	0.00005	Paola and Schowengerdt (1997)
0.05	-	Lawrence et al. (1996)
0.05	0.5	Partridge and Yates (1996)
0.1	-	Haykin (1999), Gallagher and Downs (1997)
0.1	0.3	Ardo et al. (1997)
0.1	0.9	Foody et al. (1996), Pierce et al. (1994)
0.15	0.075	Ederhart and Dobbins (1990)
0.2	-	Bisshof et al. (1992)
0.2	0.6	Gong et al (1996)
0.25	0.9	Swingler (1996)
0.3	0.6	Gopal and Woodcock (1996)
0.5	0.9	Hara et al. (1994)
0.8	-	Staufer and Fischer (1997)

Table 3.5. Parameters and values used to investigate the suitable configuration of SSOM.

Parameters	**Values**
Number of input layer neuron	23
Number of output layer neuron	4
Number of competitive layer neuron	2 x 2, 4 x 4, 6 x 6, 10 x 10, 15 x 15, 20 x 20
Initial neighborhood radius	Automatic
Initial weight	0.0
Initial learning rate	0.001, 0.005, 0.01, 0.025, 0.05, 0.075, 0.1, 0.25, 0.5, 0.9
Iterations	50, 100, 200, 500, 1000

3.5 Classification

For image classification, three hard classifiers are tested in this study, which include a Gaussian maximum likelihood classifier (GMLC), a backpropagation neural network (BPNN), and a supervised SOM (SSOM) neural network. Moreover, a soft supervised SOM (3SOM) neural network is used to classify the data to improve image classification accuracy according to the soft classification.

The classification scenarios are shown in Table 3.6. The training and testing data consisting of pure pixel values are used for the hard classifiers (i.e., GMLC, BPNN, and SSOM). The output from the GMLC, BPNN, and SSOM are thematic images (i.e., hardened images), which are evaluated by using measures of accuracy assessment for hard classification.

For the soft classification, two approaches of 3SOM are employed to derive the classification accuracy, which are the fully-soft (trains with pure and mixed pixels) and partially-soft (trains with only pure pixels). Appropriate measures of a soft classifier are used to assess classification accuracy for the soft classification.

Table 3.6. Classification scenarios

	Classification				
	Hard classification			Soft classification	
	GMLC	BPNN	SSOM	3SOM	
				Partially	Fully
Training data	H	H	H	H	S
Testing data	H	H	H	S	S

H represents hard training and testing data (only pure pixels) and S represents soft training and testing data (both pure and mixed pixels).

The algorithms of the GMLC, BPNN, SSOM, and 3SOM classifiers used in this research are described below.

3.5.1 Gaussian maximum likelihood classifier (GMLC)

The Gaussian maximum likelihood classifier (GMLC) is a probabilistic classifier and the most widely available and used classification algorithm in remote sensing. This classifier has often been treated as a benchmark to evaluate the performance of new classifiers. In most studies, the GMLC has generally been used as a hard classifier. The classification is based on the probability density function from which the posterior probability of class membership is given by Equation (3) (Foody, 1992; Foody, 1996a).

$$P(x \mid i) = \frac{P(i)p(x \mid i)}{\sum_{i=1}^{c} P(i)p(x \mid i)} \qquad (3)$$

In Equation (3), $P(x \mid i)$ is the posterior probability of pixel x belonging to class i, $p(x \mid i)$ is a probability density function derived from Equation (4), $P(i)$ is the a priori

probability for class *i*, and *c* is the total number of classes. Each pixel is then allocated to the class with which it has the highest a posteriori probability of membership. The actual magnitude of the class membership probabilities is ignored yet can provide useful information on the quality of the class allocation. In Equation (4) below:

$$p(x \mid i) = \frac{1}{(2\pi)^{k/2} \left| \sum i \right|^{1/2}} \exp\left(-1/2 \left[\left(X - \mu_i^{\bullet} \right)^T \sum_i^{-1} \left(X - \mu_i^{\bullet} \right) \right] \right) \qquad (4)$$

$p(x \mid i)$ is a probability density function of a pixel *x* as a member of class *i*, *k* is the number of bands, and *X* is the vector denoting the spectral response of the pixels.

3.5.2 Backpropagation neural network (BPNN)

With the existence of distribution-free classifiers, one particularly attractive alternative for the supervised classification of remotely sensed data is the use of artificial neural networks.

The backpropagation neural network (BPNN) is one of the most frequently used supervised classifiers for remotely sensed imagery. Therefore, the performance of this approach is evaluated in this research.

A typical BPNN consists of three layers of neurons: an input layer that receives external inputs, one hidden layer, and an output layer which generates the classification results (Figure 3.6).

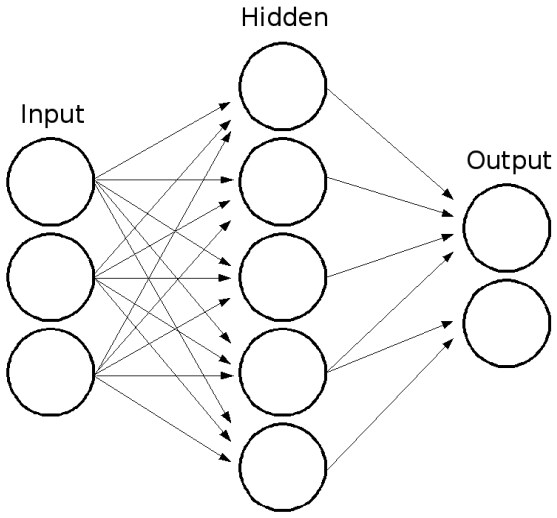


Figure 3.6. A typical backpropagation neural network.

Each neuron in a layer is connected to every neuron in the next layer. The data are entered at the input layer and pass through a hidden layer to the output layer. Each neuron calculates a weighted sum of the outputs from neurons in the preceding layer to which it is connected, passes this through a transfer function to derive its own output, which is then fed on to neurons in the next layer. Therefore, the input to a given neuron may be determined from Equations (5) and (6).

$$net_j = \sum_{i=1}^{n} o_i \omega_{ji} \tag{5}$$

and

$$o_j = f(net_j) = \frac{1}{\left(1 + e^{-net_j}\right)} \tag{6}$$

where o_i is the input to neuron j from neuron i, ω_{ji} is the weight for the connection linking neuron i and neuron j, n is the total number of neurons having links with neuron j in the

same layer as neuron *i*, o_j is the output from neuron *j*, known as the activation level, and *f* stands for an activation function as a sigmoid function (Schalkoff, 1992)

Generally, a back-propagation learning algorithm (Schalkoff, 1992) is used to train the network. With this algorithm, training pixels are presented to the network via the input layer and fed forward through the network using Equation (5) and (6). In the output layer, network outputs are compared with the target outputs, which are known for training pixels. The error, if any, is then propagated backward through the network to the input layer with the weights for relevant connections corrected via a relation given by Equation (7).

$$\Delta \omega_{ji}^{t+1} = \eta \delta_j o_i + \alpha \Delta \omega_{ji}^{t} \qquad (7)$$

where *t* is an iteration, η is a learning rate, δ_j is a computed error, and α is the momentum factor.

The training data are then entered again and the process repeated until the overall error is minimized or declines to an acceptable level.

3.5.3 Supervised self-organizing map (SSOM)

Similar to an original SOM, the SSOM architecture is composed of the input layer and the output or competitive layer (typically organized as a two-dimensional array of neurons). All neurons in the input layer and each output layer are connected by synaptic weights (Bagan *et al.*, 2005; Li & Eastman, 2010).

In the competitive layer, the neurons are constructed during the training phase and each neuron represents a reference pattern. In general, an initial SSOM neural network represented by

a two dimensional map space containing M x N = U neurons is generated. The number of neurons can be chosen according to the complexity of the problem, with more neurons being required if several complex trends or groups are to be represented.

1) Architecture of SSOM

For SSOM neural networks, each neuron (u) in the competitive layer of a network is characterized by a weight vector (w), which has dimensions of $J + K$, where J is the number of features in the data and K is the number of classes in the data. The weight vector of each neuron for a SSOM consists of a feature weight vector $w_f = [\omega_{uf1}, \omega_{uf2}, \omega_{uf3,\ldots}, \omega_{ufJ}]$ and a class weight vector $w_c = [\omega_{uc1}, \omega_{uc2}, \omega_{uc3,\ldots}, \omega_{ucK}]$, where ω_{uf} and ω_{uc} are a feature weight value and a class weight value, respectively, of neuron u. To initialize the network, a weight value (ω) is randomly generated by a normal distribution between 0 and 1. The topology of the competitive layer and weight vector structure are shown in Figure 3.7.

In the input layer of the SSOM neural network, the input neuron is structured by a input vector (x), which contains two parts including a feature input vector $x_f = [x_{f1}, x_{f2}, x_{f3,\ldots}, x_{fJ}]$ and a class input vector $x_c = [x_{c1}, x_{c2}, x_{c3,\ldots}, x_{cK}]$, where x_f is an input value of a data feature and x_c is the class information. The dimensions of the x_f and x_c vectors are dependent on the number of features (J) and number of classes (K) in the data, respectively. The structure of input data is shown in Figure 3.8.

In the training stage, the SSOM is a simple modification of the SOM-algorithm. The training process of the SSOM neural network is a kind of competitive learning without any objective function, thus, local optimum problems do not exist.

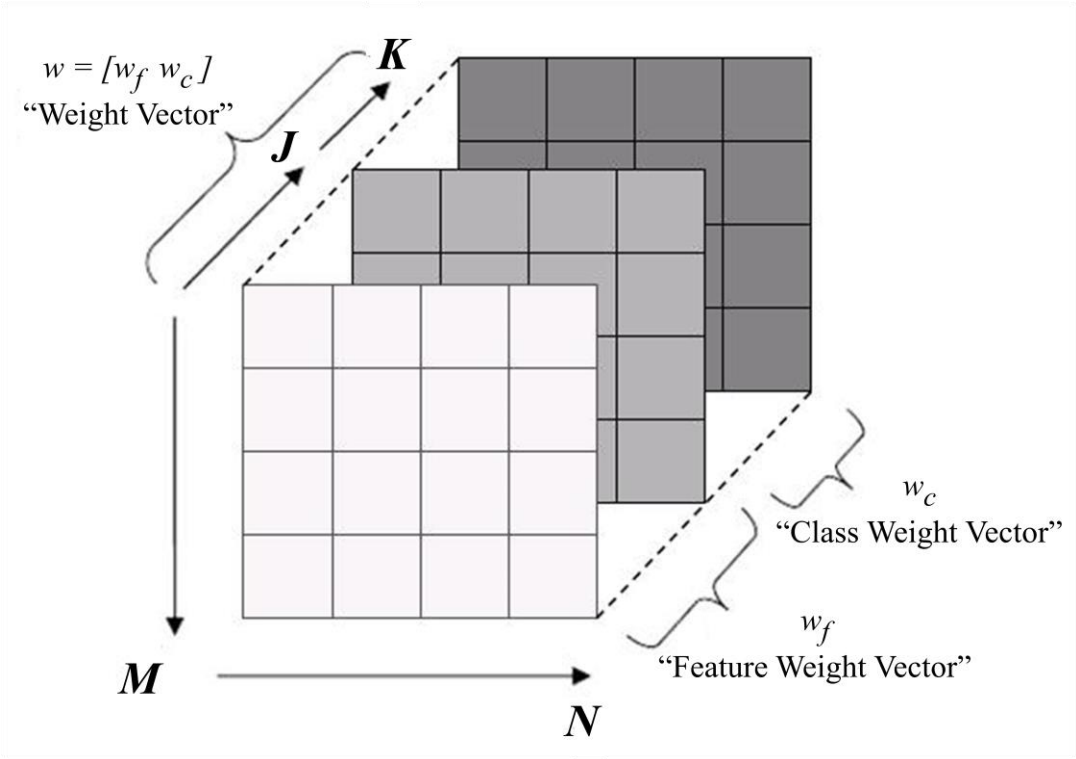


Figure 3.7. The topology of the competitive layer and weight vector structure for a SSOM neural network.

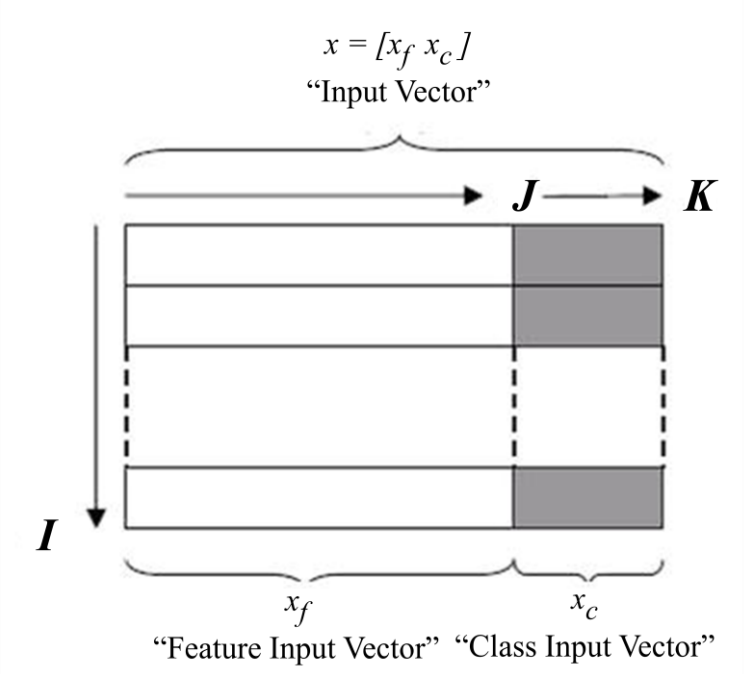


Figure 3.8. The structure of input data for a SSOM neural network.

2) Learning algorithm of SSOM

Prior to training, each neuron's weights must be initialized. Typically these will be randomly generated from a normal distribution between 0 and 1.

The identification of winners is a key procedure for SOM because it is based upon competitive learning. The only difference between SOM and SSOM is how the network is used. In order to implement the SOM in supervised mode, the training of SSOM is done using both feature and class input vectors (x_f and x_c), but while finding the winner neuron only the x_f is considered. To determine the winner neuron, the Euclidean distances between a feature input vector (x_f) and a feature weight vector (w_f) are calculated for each neuron in the competitive layer, and then the neuron with the minimum distance (its weight vector is closest to the input vector) is assigned as the winner neuron or the best matching unit (BMU). The BMU is given as:

$$BMU = \arg\min_{u} \left\{ \sqrt{\sum_{j=1}^{J} \left(x_{fj}^{t} - \omega_{ufj}^{t} \right)^2} \right\} \tag{8}$$

where x_{fj}^{t} is the value of input feature j at iteration t, ω_{ufj}^{t} is weight value of the feature j of neuron $u \in \{1 \cdots\cdots U\}$ in competitive layer at iteration t, and J is number of data features.

In the next step, both the feature weight values and class weight values of the winner neuron and its neighborhood are updated according to the following equation:

$$\begin{aligned} \omega_{ufj}^{t+1} &= \omega_{ufj}^{t} + \eta^{t} \alpha^{t} \cdot \left(x_{fj}^{t} - \omega_{ufj}^{t} \right) \\ & \qquad \text{if } D \le \sigma^{t} \\ \omega_{uck}^{t+1} &= \omega_{uck}^{t} + \eta^{t} \alpha^{t} \cdot \left(x_{ck}^{t} - \omega_{uck}^{t} \right) \end{aligned} \tag{9}$$

Or

$$\omega_{ufj}^{t+1} = \omega_{ufj}^{t}$$

$$\text{if } D > \sigma^{t} \qquad (10)$$

$$\omega_{udk}^{t+1} = \omega_{udk}^{t}$$

where ω_{uck}^{t} is the weight value of class $k \in \{1 \cdots\cdots K\}$, η^{t} is the amount of influence of a neuron's distance from the BMU at time *t*. η^{t} is given by Equation (11), α^{t} is the learning rate at iteration *t* and is given by Equation (12). In Equation (11), *D* is the distance between the BMU and other neurons in the competitive layer, and σ^{t} is the radius of the neighborhood at time *t*, which can be calculated by Equation (13).

$$\eta^{t} = \exp\left(-\frac{D^{2}}{2 \cdot \left(\sigma^{t}\right)^{2}}\right) \qquad (11)$$

$$\alpha^{t} = \alpha_{0} \exp\left(-\frac{t}{\lambda}\right) \qquad (12)$$

In Equation (12), α_0 is the learning rate constant at time t_0, *t* is the current time-step, and λ is the time constant that can be defined by Equation (14).

$$\sigma^{t} = \sigma_{0} \exp\left(-\frac{t}{\lambda}\right) \qquad (13)$$

Equation (13) demonstrates the calculation of the radius constant of the neighborhood (i.e., σ_0) at time t_0, where *t* is the current time-step, and λ is the time constant that can be defined by Equation (14):

$$\lambda = \frac{T}{\log(\sigma_0)} \qquad (14)$$

where *T* is the number of iterations.

With these training samples as input, the SSOM neural network is trained until it converges.

3) Classification

In the classification stage, the SSOM is used as a classifier to determine the class of an unknown sample by locating the BMU of the unknown input. This is done using only the feature weight vector (w_f) for each neuron, and assigning the input to the class in the class weight vector (w_c) of the BMU that has the largest value. The specified endmember (hardened-class) is indicated by the class index that has the largest class weight value.

3.5.4 Soft-supervised self-organizing map (3SOM)

The 3SOM uses the same method as SSOM for input and competitive layers as well as the learning algorithm and the weight based on the distance decay function. However, the 3SOM differs during the classification stage in determining the output. The 3SOM is used as a classifier to determine the fuzzy membership of an unknown input by locating its BMU using only the feature weight vector (w_f) for each neuron, and assigning the input to the fuzzy membership in the class weight vector (w_c) of the BMU. The set of class weight values obtained from the class weight vector $w_c = [\omega_{uc1}, \omega_{uc2}, \omega_{uc3}, \ldots, \omega_{ucK}]$ is the mixture proportion of each class, and may be treated as fuzzy membership values.

3.6 Evaluation of classification accuracy

Accuracy assessments in this research are performed) to calculate the accuracy of both the hard classification and soft classification. All measured accuracy approaches used in this research are described as follows:

3.6.1 Accuracy assessment of hard classification

The most widely used measures are derived from a classification confusion or error matrix that shows a cross-tabulation of the class labels in the output of a classification against those in the ground truth data. The overall accuracy and the Kappa coefficient, which are efficient and reliable measures, are used in this research (Zhan *et al.*, 2002).

1) Overall accuracy (OA)

$$OA = \frac{\sum_{i=1}^{c} n_{ij}}{N} \tag{15}$$

Where n_{ij} is the number of pixels classified correctly, N is the total number of pixels and c is the number of classes.

2) Kappa coefficient (KAP)

Kappa analysis is a discrete multivariate technique used in accuracy assessment for determining statistically if one error matrix is significantly different from another (Bishop *et al.*, 1975 as cited in Zhan *et al.*, 2002). and has become a popular component of accuracy assessment (Hudson & Ramm, 1987; Congalton, 1991; Richards, 1993; Foody, 1995; Congalton & Green, 1999 as cited in Zhan *et al.*, 2002). The Kappa statistic is given by:

$$KAP = \frac{N\sum_{i=1}^{c} n_{ii} - \sum_{i=1}^{c} n_{i+}n_{+i}}{N^2 - \sum_{i=1}^{c} n_{i+}n_{+i}} \tag{16}$$

where n_{ij} is the number of pixels classified correctly, n_{i+} is the number of pixels classified into class *i*, n_{+i} is the number of pixels classified into class *i* in the testing data set, *N* is the total number of pixels and *c* is the number of classes.

3.6.2 Accuracy assessment of soft classification

For the soft classification output, four measures of accuracy are estimated to assess the difference between each classified image and the reference images (Tatem *et al.*, 2002). The four measures are described as follows:

1) Area error proportion (AEP)

One of the simplest measures of agreement between a set of known proportions in matrix *y*, and *a* set of predicted proportions in matrix *a*, is the area error proportion (*AEP*) per class,

$$AEP_j = \frac{\sum_{i=1}^{n} \left(y_{ij} - a_{ij} \right)}{\sum_{i=1}^{n} a_{ij}} \qquad (17)$$

where, *j* is the class and *n* is the total number of pixels. This statistic informs about bias in the prediction image.

2) Correlation coefficient (CC)

The correlation coefficient, *r*, measures the amount of association between a target, *y*, and a predicted set of proportions, *a*,

$$r_j = \frac{c_{yj \cdot aj}}{s_{yj} \cdot s_{aj}}, c_{yj \cdot aj} = \frac{\sum_{i=1}^{n} \left(\bar{y}_{ij} - y_{ij}\right) \cdot \left(\bar{a}_{ij} - a_{ij}\right)}{n-1} \quad (18)$$

where, $c_{yj \cdot aj}$ is the covariance between y and *a* for class *j*, and s_{yj} and s_{aj} are the standard deviations of *y* and *a* for class *j*. This statistic provides information about the prediction variance.

3) Closeness (S)

Foody (1996a) suggests a measure related to the Euclidean distance between the land cover proportions predicted by the classification and those of the reference data. It measures the separation of the two data sets, per pixel, based on the relative proportion of each class in the pixel. It is calculated as:

$$S_i = \sum_{j=1}^{c} \left(y_{ij} - a_{ij}\right)^2 / c \quad (19)$$

where y_{ij} is the proportion of class *i* in a pixel from the reference data, a_{ij} is the measure of the strength of membership to class *j* taken to represent the proportion of the class in the pixel from the soft classification, and *c* is the total number of classes.

4) Root mean square error (RMSE)

The root mean square error (RMSE) is used to estimate the overall accuracy for each class of the soft classification. It is calculated by:

$$RMSE_j = \sqrt{\frac{\sum_{i=1}^{n} \left(y_{ij} - a_{ij}\right)^2}{n}} \quad (20)$$

The RMSE provides a measure of the inaccuracy of the prediction (bias and variance).

3.7 Evaluation of uncertainty in classification accuracy

Uncertainty has been receiving increased attention in geographical information science research for over a decade (Goodchild & Gopal, 1989; Heuvelink, 1998; Zhang & Goodchild, 2002). The uncertainty in the spatial output of geographical information system (GIS) and remote sensing analyses needs to be assessed, particularly in classification accuracy (Atkinson & Foody, 2002; Fisher, 1994; Foody, 2002).

Although most studies try to improve classification methods to increase accuracy, there is always an element of uncertainty in the classification results. Failure to recognize uncertainty may lead to erroneous and misleading interpretations.

Dungan (2002) defined uncertainty as a "quantitative statement about the probability of error". With accurate measurements, the estimated or predicted values will have small uncertainty, whereas with inaccurate measurements, estimates or predictions should be associated with large uncertainty.

The sources of uncertainty in remote sensing analyses are considered in five aspects (Dungan, 2002): parameter uncertainty (parameters in the models or equations), uncertainty about the model (the form or structure of the model), uncertainty about the support (the area over which a variable is measured or predicted, e.g., the instantaneous field of view, flight variables), position uncertainty (the location of data values, e.g., GCP), and variable uncertainty (input variables).

Monte Carlo simulation is a well-established technique which involves the computation of uncertainty in the output induced by the quantified uncertainty in the input and model (Canters, 1997; Heuvelink, 1999). Although this technique is computationally intensive, it has

the advantage of being universally applicable to analyze the propagation of error (Heuvelink & Burrough, 1993; Canters, 1997).

Therefore, the aim of this research is to evaluate the uncertainty in the classification accuracy by considering the impact of possible factors on the spatial variation in the classifier. Only the synthetic data is used in this evaluation. Furthermore, the Monte Carlo simulation technique is applied to assess the reliability of the classification output by focusing on the uncertainty associated with the input data, training data, and the classifier itself.

3.7.1 Uncertainty associated with input data

The variations of environmental conditions (e.g., land management practices, climate change, atmosphere interactions, and soil fertility) and data preprocessing can affect the accuracy and reliability of classification results. Those variations have an influence on the input data resulting in classification uncertainty. In order to evaluate the uncertainty in the classification accuracy associated with input data, "noise" is added to the input images. Noise represents the variations of environmental conditions that lead to uncertainty in input image. Noise is derived from a random number generator based on a normal distribution using the extracted mean and standard deviation of each class. The uncertainty in the classification accuracy associated with the input image is analyzed by running the same classifier and training data and varying the input data.

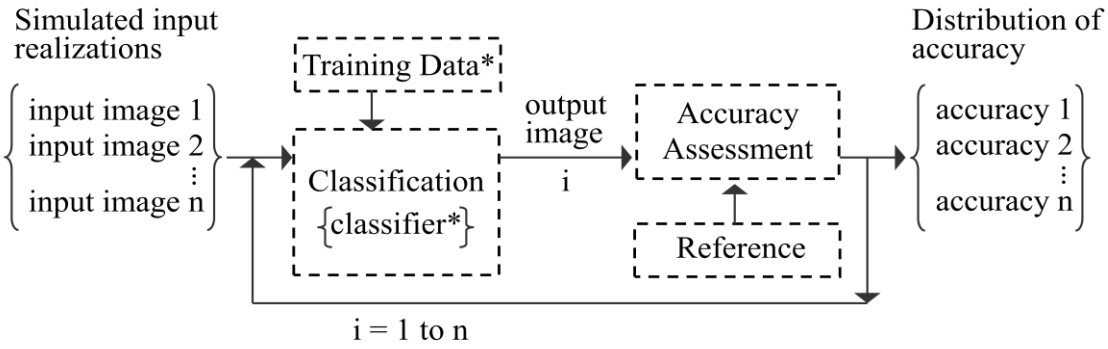


Figure 3.9. Evaluating the uncertainty in classification accuracy associated with input data.

The process shown in Figure 3.9 is repeated many times with a new realization of the input image. The output images of each run are evaluated to assess the classification accuracy. The classification accuracy derived from the output images and reference data of each run are used to generate the classification accuracy distribution. Moreover, they are accumulated and calculated to represent the accuracy possibility.

3.7.2 Uncertainty associated with training data

Training data is not only an important component of the classification process but also to the accuracy of classification, therefore, two criteria for generating a set of training data, which are random selecting and shuffling sequence, are established to study the impact of training data on the classification accuracy. With the same image input and classifier setting, simulations are run with different random selections of training data for the first test, then simulations are run with different random sequences of the training sample in the same training data for the second test. The output images are evaluated for accuracy by generating and the accuracy possibility and distribution to show the classification reliability (Figure 3.10).

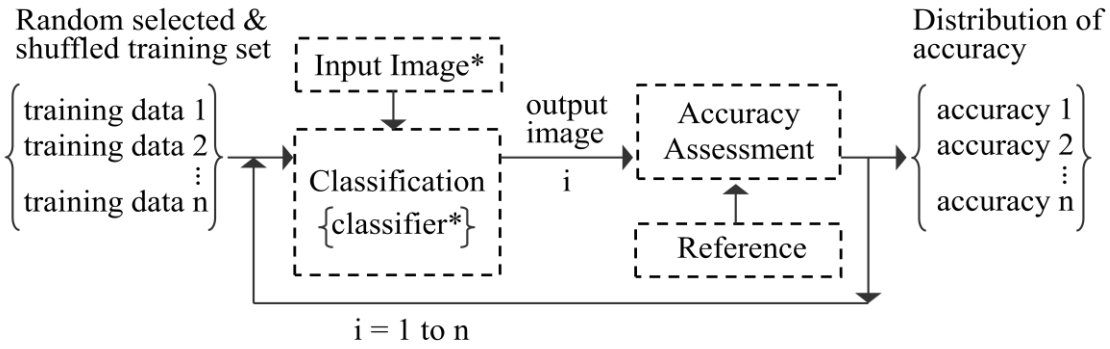


Figure 3.10. Evaluating the uncertainty in classification accuracy associated with training data.

3.7.3 Uncertainty associated with classifier

The accuracy of neural network-based classification is determined by the network architecture and internal parameters. Different neural network design and settings can lead to uncertainty in classification accuracy. To examine how the classifier itself affects the classification accuracy, the classifications are run using varied parameter values while the input image and training data are kept constant. The classification accuracy is assessed by comparing the output images of each classification with the reference data. Then, the accuracy distribution and possibility are analyzed to illustrate the sensitivity of classification results (Figure 3.11).

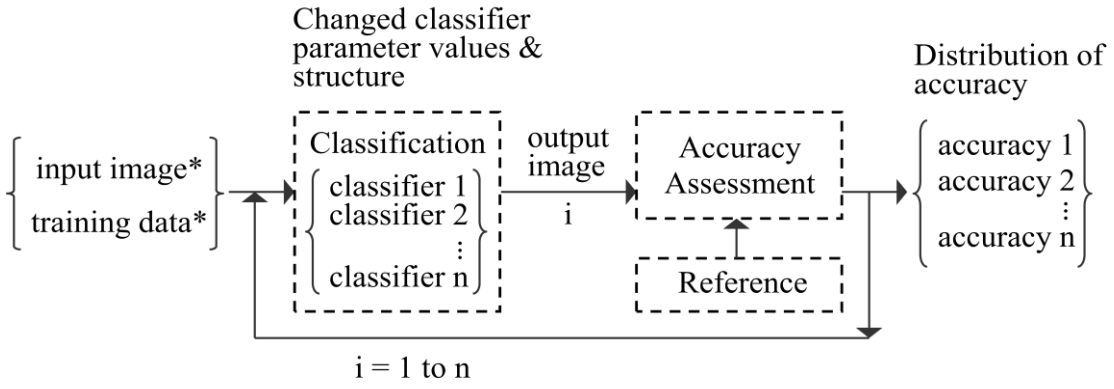


Figure 3.11. Evaluating the uncertainty in classification accuracy associated with classifier.

Chapter 4

Testing and Developing Suitable Method Using Synthetic Data

To test the effectiveness of the proposed method, a synthetic dataset is generated based on a remotely sensed time-series image of MODIS-EVI. This data is simulated using a linear mixture model as described in Chapter 3.1 for realization. Accordingly, the properties of this data can be clearly defined and used for reliable verification of the testing results.

In this chapter, a suitable classification method is developed and investigated by using synthetic data. The input data are examined to determine the appropriate input for land cover classification. Then, the suitable neural network configuration is configured to tune classifiers for the best performance. In addition, this chapter investigates the performance of hard and soft classifications as well as assesses the sensitivity and reliability of the classification output with an emphasis on the uncertainty in classification associated with the input data, training data, and the classifier itself. The following six experiments were performed and are described in this chapter: 1) comparison of input images between time-series and phenology images by using the self-organizing map classifier (SSOM), 2) selection of the suitable neural network configuration of the back-propagation neural network classifier (BPNN) and the SSOM, 3) comparative evaluation of the SSOM with the Gaussian maximum likelihood classifier (GMLC) and the BPNN, 4) comparative evaluation of the SSOM with the soft-supervised self-organizing map classifier (3SOM), 5) comparative evaluation of the fully-3SOM with the partially-3SOM, and 6) assessing the uncertainty in classification accuracy of SSOM.

4.1 SSOM approach to land cover classification using time-series and phenology images

The aim of this experiment is to determine the appropriate input data for land cover classification. The experimental procedures are shown in Figure 4.1. In this experiment, the time-series and phenology images are classified using the same method to investigate which input provides the highest classification accuracy.

The first input dataset, the time-series image (TIME), is smoothed by the Savitzky-Golay filtering technique (Jonsson & Eklundh, 2004) to remove the atmospheric and cloud conditions as described in Chapter 3.2.

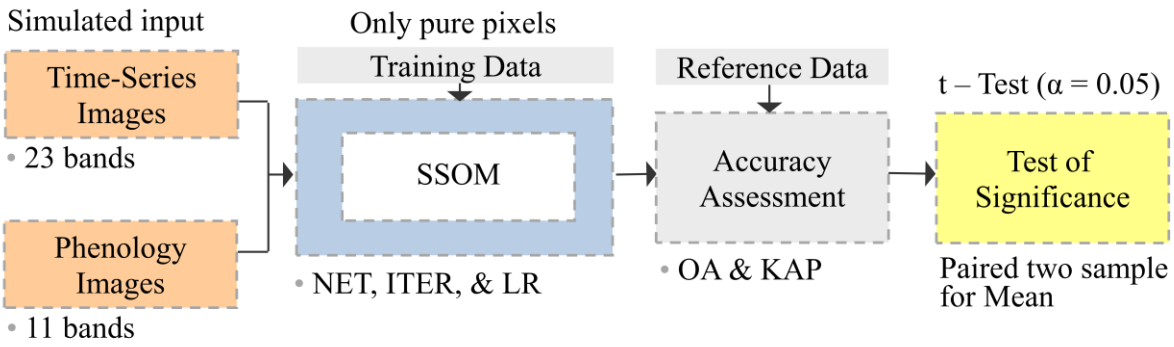


Figure 4.1. Experimental procedure of comparative evaluation of SSOM using TIME and PHEN.

Using the feature extraction approach, the phenology image (PHEN), which is the second input dataset, is extracted from the time-series image. This data represents important phenological information about the vegetation growing season. Eleven meaningful phenological parameters consist of: 1) start of season, 2) end of season, 3) length of season, 4) base value, 5) position of middle of season, 6) maximum of fitted data, 7) amplitude, 8) left derivative, 9) right derivative, 10) large integral, and 11) small integral (Jonsson & Eklundh, 2006).

In order to investigate which input data is most suitable, the supervised self-organizing map (SSOM) classification with 300 different neural network configurations is applied to both the TIME and PHEN datasets. These configurations are constructed from three significant parameters consisting of six numbers of competitive layer neurons (NET), ten values of initial learning rates (LR), and five numbers of iterations (ITER) as show in Table 3.5.

In this experiment, the SSOM is trained by randomly selecting 240 pure pixels (60 pixels for each class) to classify each input data. The classification accuracy is assessed by calculating the overall accuracy (OA) and the Kappa coefficient (KAP). Additionally, the distributions of classification accuracies are also generated to quantify the accuracy and robustness for each input dataset. Then, a t-test is performed to test the statistical significance difference between the mean accuracy derived from TIME and PHEN at $\alpha = 0.05$.

The classification results revealed that SSOM produces considerably higher classification accuracy when applied to the TIME dataset as compared to the classification results when applied to PHEN. Moreover, SSOM with TIME achieves higher accuracy for most configurations (Figure 4.2). Table 4.1 also shows that the mean OA and KAP that obtained by TIME (OA=81%, KAP=0.75) is higher than that obtained by PHEN (OA=74%, KAP=0.66).

Figure 4.3 shows the classified images that provided the highest accuracy using SSOM with TIME and PHEN. Visual interpretation indicates that the classified map of SSOM with TIME provides a more accurate classification with sharper boundaries of each class. In addition, OA and KAP obtained by TIME (OA=88%, KAP=0.85) is higher than that obtained by PHEN (OA=83%, KAP=0.77).

For statistical significance testing, a t-test is performed for these two input datasets to determine whether the accuracy derived from TIME is statistically different from that derived

from PHEN. The results of the paired difference t-statistic are listed in Table 4.2. The results reveal that the differences in OA and KAP are statistically significant at an alpha level of 0.05. The paired-difference means also indicate that SSOM with TIME can achieve significantly higher OA and KAP than that with PHEN.

In this comparison, TIME has achieved higher accuracy than PHEN. Although dimensional space is reduced by extracting PHEN, this extraction may cause information loss for image classification. On the other hand, TIME has a high-dimensional space and may lead to the "curse of dimensionality". The SSOM is capable of mapping high dimensional inputs onto low dimensional units and preserving the topology of input patterns in the low dimension after dimension reduction (Bagan *et al*., 2005). SSOM is consequently very useful for analyzing large datasets because this technique can produce both a reduced amount of data by clustering and a projection of dominant data patterns on a lower-dimensional display.

Therefore, the time-series image is a potentially useful input dataset for land cover classification. This data will be used to investigate the suitable neural network configuration and to compare the classifiers in the following experiments.

Table 4.1. Statistics of classification accuracy derived from TIME and PHEN

	Overall Accuracy (%)		Kappa Coefficient	
	TIME	PHEN	TIME	PHEN
Minimum	63.24	57.96	0.5099	0.4395
Maximum	88.36	82.88	0.8448	0.7717
Mean	81.39	74.47	0.7519	0.6596
Standard Deviation	4.64	5.30	0.0619	0.0707

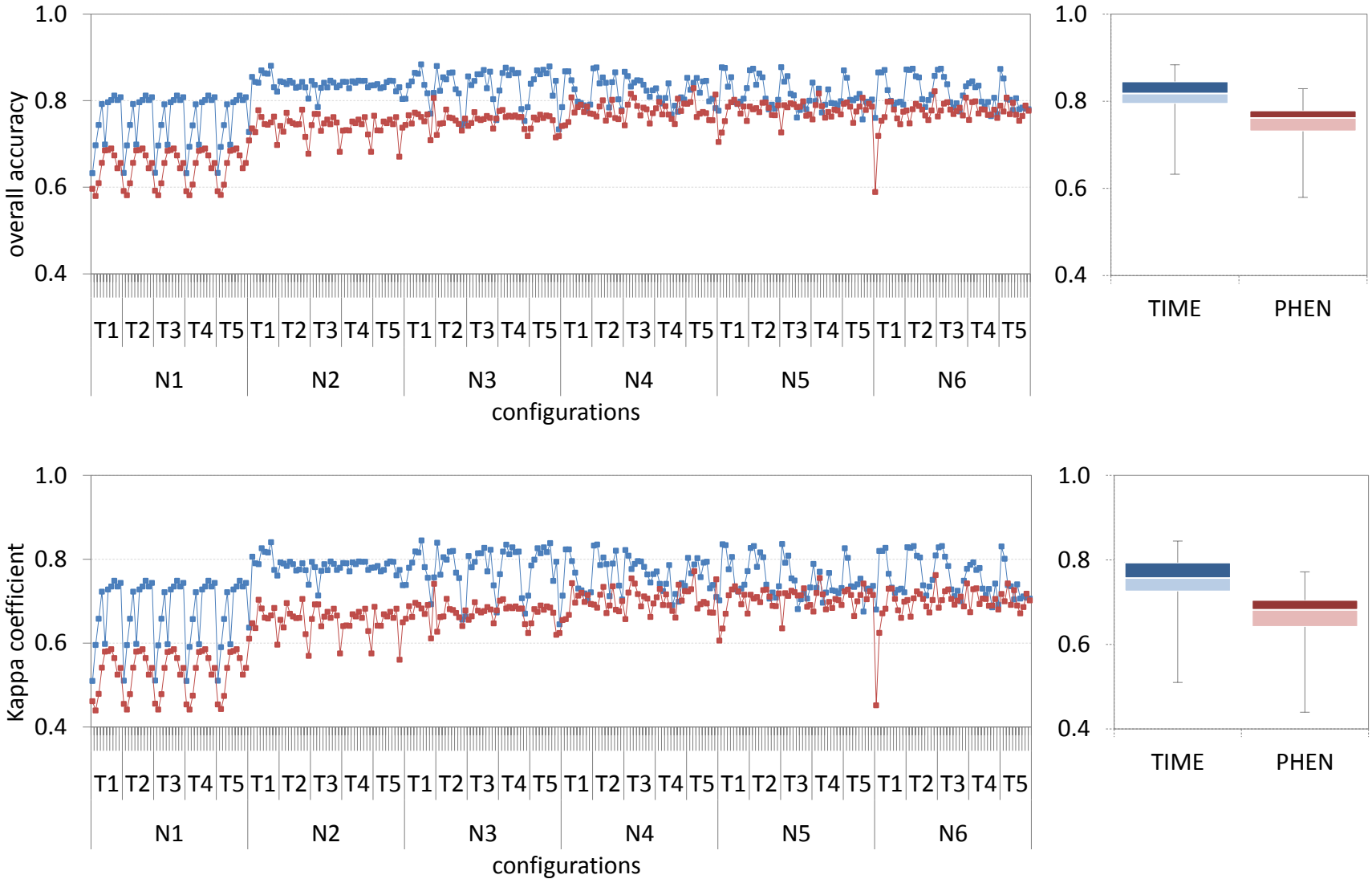
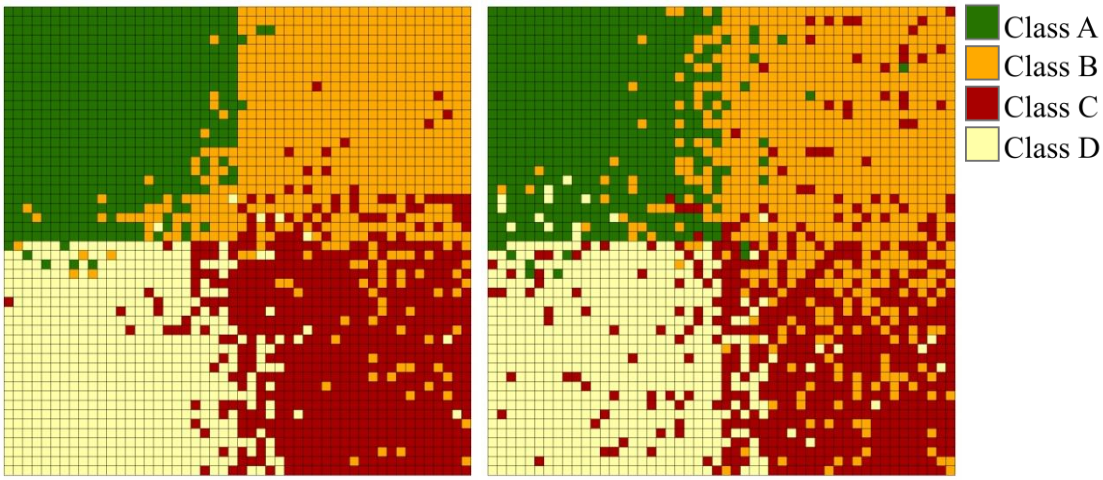


Figure 4.2. Distribution of classification accuracy derived from TIME and PHEN in different neural network configurations.

Table 4.2. Test of significance difference in accuracy between TIME and PHEN

	Paired Difference				
	Mean	Std. Deviation	Std. Error Mean	t	Sig. (2-tailed)
Overall Accuracy (%)					
TIME – PHEN	6.92	4.76	0.28	25.198	< .001
Kappa Coefficient					
TIME – PHEN	0.0923	0.0645	0.0036	25.197	< .001



(a) (b)

Figure 4.3. The classified images derived from (a) TIME and (b) PHEN providing the highest accuracy.

4.2 Selecting the suitable neural network configuration of BPNN and SSOM

The speed and effectiveness of the learning process are vital components of successful neural network-based classifications, and are determined by the network architecture and internal parameter values. A careful designing phase has been carried out to tune each classifier for optimal performance using the already-prepared training and testing data. In this study, the

performance of the backpropagation neural network (BPNN) and the supervised self-organizing map (SSOM) are examined by establishing different values for each parameter. The experimental procedures are shown in Figure 4.4. Then, the configuration that provides the highest classification accuracy is selected as the suitable configuration and is used for subsequent experiments.

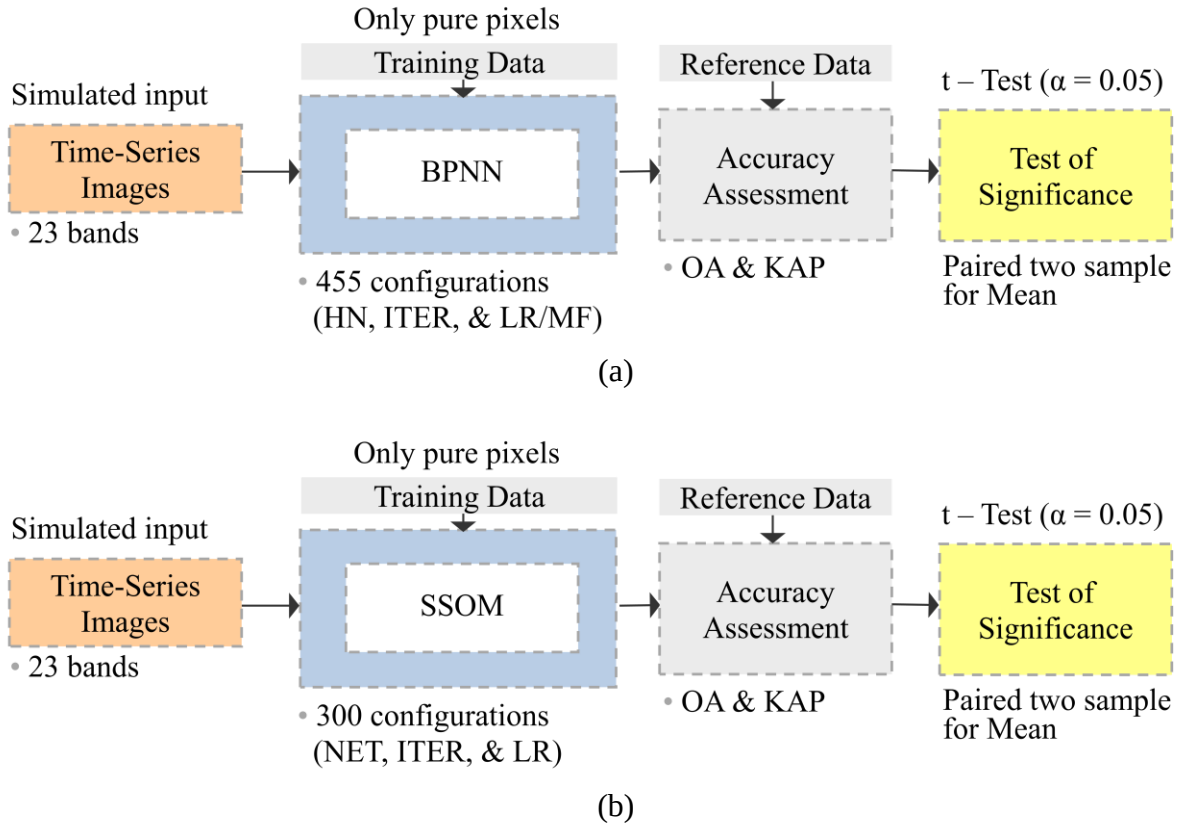


Figure 4.4. Experimental procedures to investigate the suitable neural network configuration of (a) BPNN and (b) SSOM.

To investigate the performance of the BPNN, 455 neural network configurations are formed based on three primary parameters. Each parameter value is identified by Kavzoglu & Mather (2003), which consists of seven different numbers of hidden layer neurons (HN), thirteen learning rate (LR) and momentum factor (MF) values, and five numbers of iterations (ITER) as shown in Tables 3.2 to 3.4. To investigate the performance of SSOM, 300 different neural network configurations are conducted by six different numbers of competitive layer neurons (NET). For each NET, ten trials with different initial learning rates (LR) are performed and five different variants of iterations (ITER) are examined. These parameters and values are shown in Table 3.5.

Both series of configurations are set and run on a case-by-case basis by using trial-and-error analysis because there is no standard procedure for choosing the suitable configuration. All trials are performed on the same training and testing data. The classification accuracy results in different configurations of the BPNN and SSOM are measured in terms of overall accuracy (OA) and the Kappa coefficient (KAP).

Figure 4.5 shows the classification accuracy of the BPNN with different neural network configurations. The highest classification accuracy is achieved by the combination of a HN of 46, LR and MF of 0.25 and 0.9, and ITER of 50 (Table 4.3). With the suitable configuration, the BPNN produces an OA of 87% and a KAP of 0.83.

Figure 4.6 illustrates the classification accuracy of the SSOM with different neural network configurations. Table 4.4 represents the suitable values of each parameter including a NET of 6 × 6, LR of 0.075, and ITER of 50. This combination is the suitable neural network configuration because with this configuration the SSOM yields the highest OA and KAP, which are 88% and 0.85, respectively.

Table 4.3. The suitable configuration of BPNN

Parameters	**Suitable values**
Number of input layer neuron	23
Number of output layer neuron	4
Number of hidden layer neuron (HN)	46
Learning rate & momentum factor (LR&MF)	(0.25, 0.9)
Iterations (ITER)	50

Table 4.4. The suitable configuration of SSOM

Parameters	**Suitable values**
Number of input layer neuron	23
Number of output layer neuron	4
Number of competitive layer neuron (NET)	6 x 6
Initial learning rate (LR)	0.075
Iterations (ITER)	50

The results of this experiment indicate that neural network configuration has an influence on classification results. A careful architecture design and internal parameter configuration should be prepared for the best performance of the classifier. This experiment demonstrates that the classification accuracy varies across different neural network models associated with different internal parameter settings. Increasing the network size will not generate much accuracy improvement, but it will result in more computational time. For example, Figure 4.5 and Figure 4.6 show that the increase of HN in the BPNN and NET in the SSOM does not improve the classification accuracy, but it can be a cause of intensive computational time.

These guidelines will facilitate the process of design and use of the BPNN and SSOM in remote sensing classification. It should be noted that they are only valid for similar datasets and

classification problems to those used in this study. In addition, these configurations are suitable for few output classes. More complex topology may be required when a large number of input data are used or when many different classes are to be generated

4.3 Comparative evaluation of SSOM with GMLC and BPNN

This experiment is aimed to identify the appropriate classifier by comparing the classification accuracy of the SSOM to the GMLC and the BPNN. In this experiment, these three classifiers are employed in a hard classification mode and applied to the same input and training data. The BPNN and SSOM are applied using the suitable neural network configurations, which are described in the previous section.

Due to a significant impact of training samples on the performance of a classifier, particularly neural network classifiers, it is important to investigate the performance of a classifier by using different training data (Kavzoglu & Mather, 2003). In this experiment, there are two comparative evaluation tests of the GMLC, BPNN, and SSOM in different scenarios: one test with different simulated input data and another with different random training data (Figure 4.7). For the first test, 500 simulations are run with different simulated input data with the same training data. Different simulated input data are generated through a random number generator based on a normal distribution using the extracted mean and standard deviation of each class (see details in Chapter 3.1).

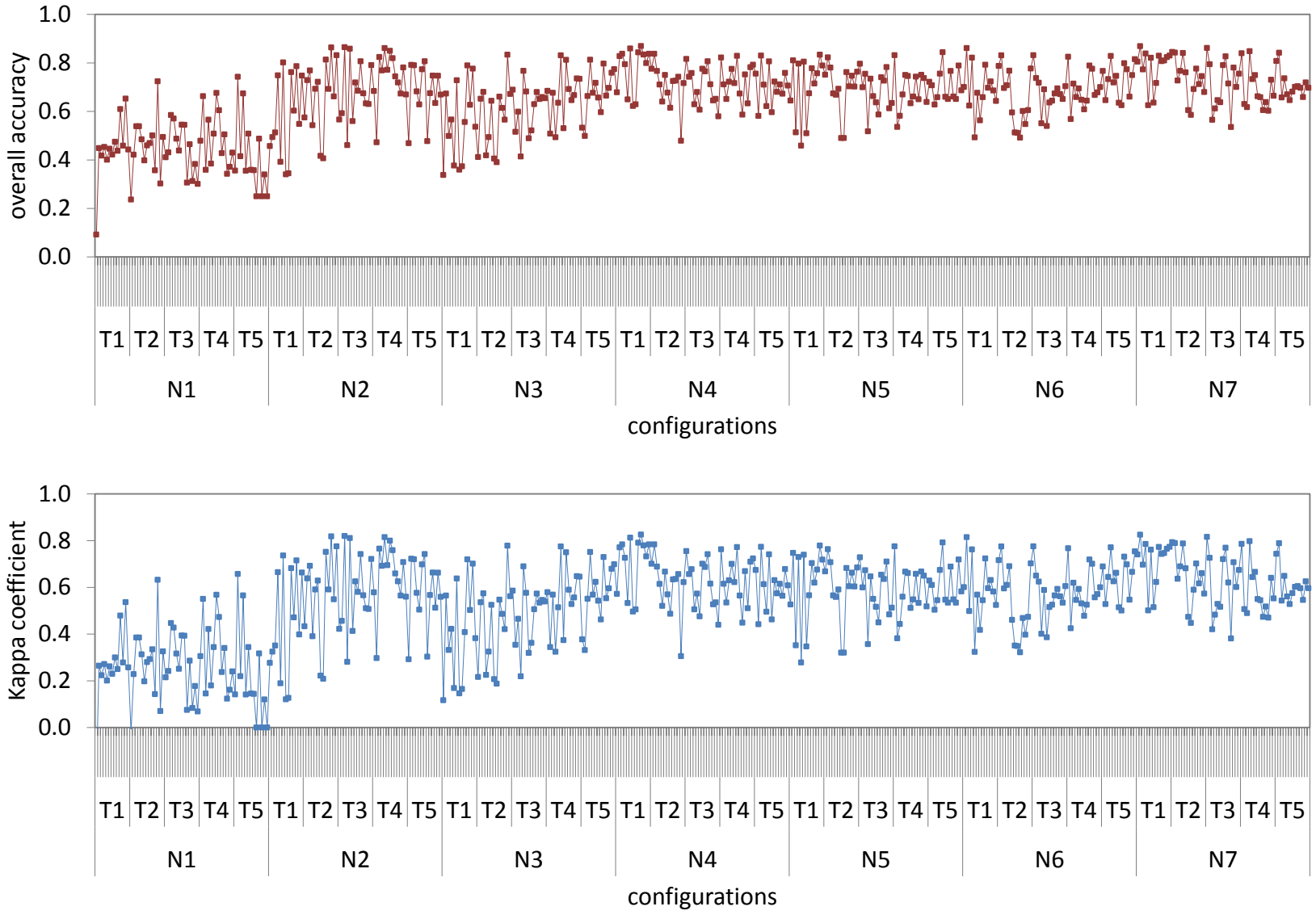


Figure 4.5. Classification accuracy of BPNN in different neural network configurations

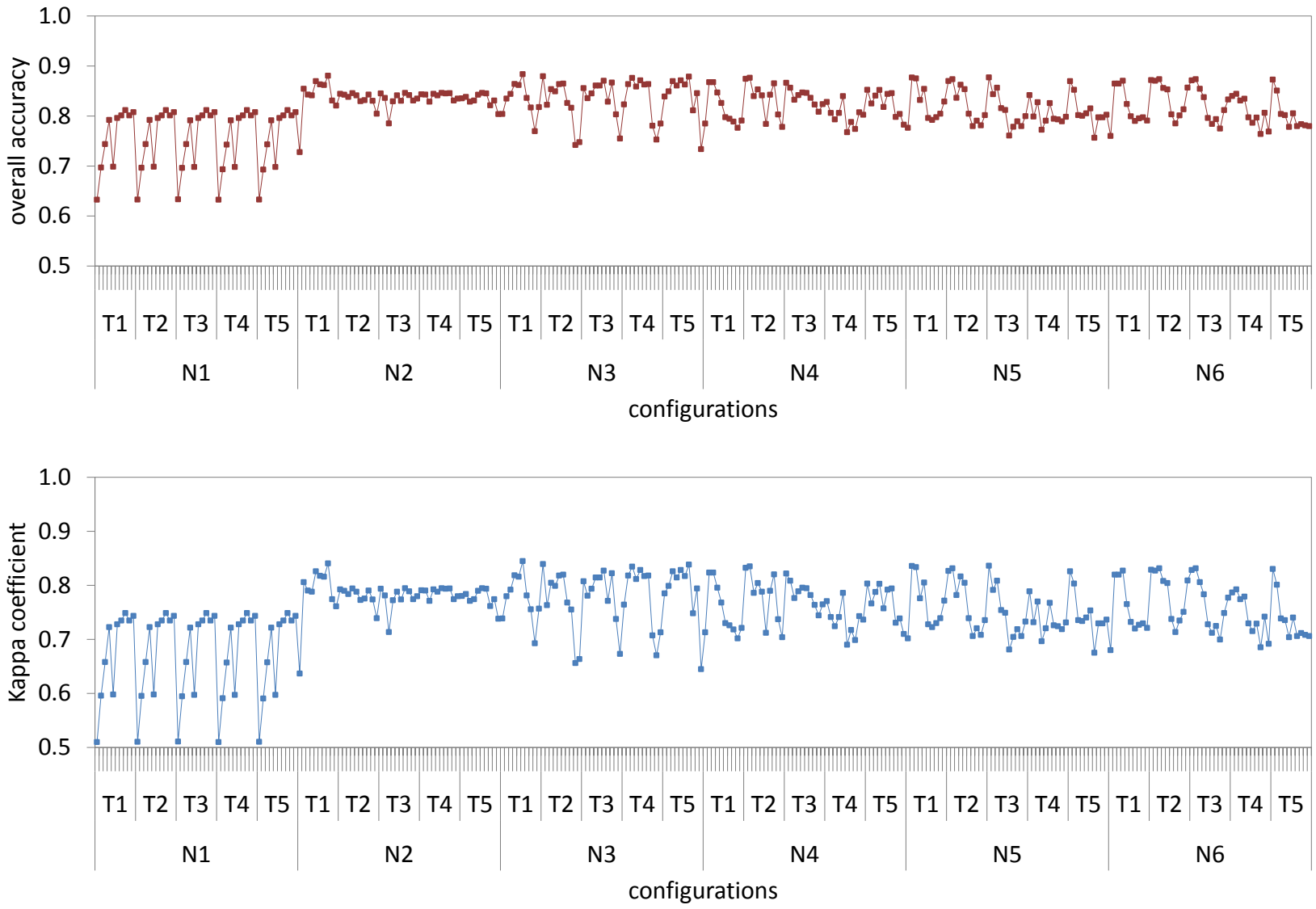


Figure 4.6. Classification accuracy of SSOM in different neural network configurations.

The second test utilizes 500 different training datasets. Each dataset is generated by randomly selecting 240 pure pixels (60 pixels per class) to train all classifiers. In this test, 500 simulations with different training data are applied to the same synthetic time-series image.

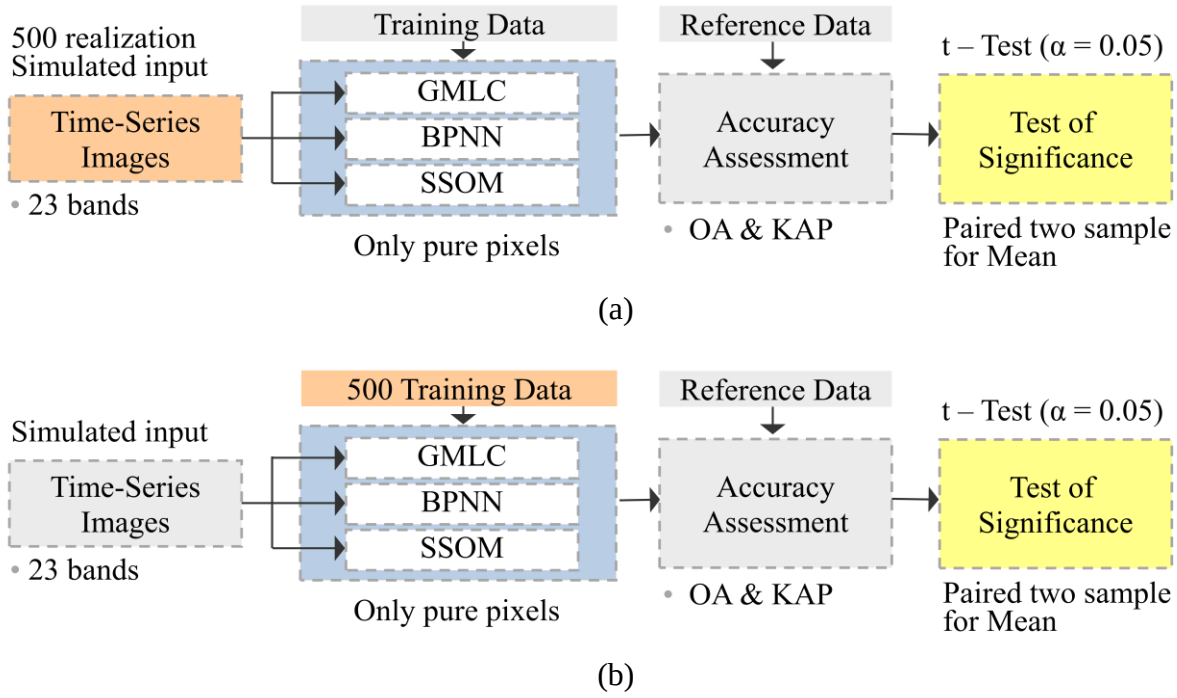


Figure 4.7. Experimental procedures of comparative evaluation of SSOM with GMLC and BPNN (a) in different simulated input data and (b) in different random training data.

The performance of the three classifiers is assessed by evaluating the classification accuracy. Distribution plots of overall accuracy (OA) and the Kappa coefficient (KAP) are also generated to compare the accuracy and robustness of each classifier. Then, the t-test is performed to determine if there is a statistically significant difference among three classifiers at $\alpha = 0.05$.

Figure 4.8 and 4.9 show the classification accuracy distributions of the GMLC, BPNN, and SSOM derived from 500 different simulated input datasets and 500 different random training datasets, respectively. The results reveal that the highest OA and KAP are obtained by the SSOM in all 500 simulations of both tests. Several accuracy distribution statistics are also reported in

Table 4.5. For the first test, the mean OA of the GMLC, BPNN, and SSOM is 81%, 78%, and 88% respectively. Similar findings are also observed in the second test, which show OAs of 80%, 78%, and 86% for the GMLC, BPNN, and SSOM respectively. Moreover, the SSOM also has the lowest standard deviation of OA (1.21% and 1.14%) for both tests, whereas the highest standard deviation of OA (6.15% and 4.16%) for both tests is obtained by BPNN.

For spatial comparison, Figure 4.10 shows the GMLC, BPNN, and SSOM classified images with the highest accuracy. In both tests, the results demonstrate that SSOM achieves more meaningful classification results than those obtained from GMLC and BPNN. The visual depiction of the results also demonstrates that classes are uniformly defined by SSOM, whereas GMLC and BPNN produce misclassification results along the boundaries of classes or in the areas of mixed pixels.

In statistical comparisons, the results of the t-statistic of each paired difference are listed in Table 4.6. In both tests, the results show that the differences in the mean accuracy of these three classifiers are statistically significant. The mean differences also indicate that the mean OA and KAP of SSOM is significantly higher than those of GMLC and BPNN. Moreover, comparing to other classifiers, BPNN performs less satisfactorily as illustrated by considerably low mean OA and KAP.

Results from both tests illustrate that BPNN performs less satisfactorily as indicated by lower accuracy comparing to other classifiers. Although GMLC shows ability to control uncertainty in the classification accuracy, the multivariate normal model of GMLC is not as effective as the SSOM in the classification of time-series images. This is because the GMLC highly depends on an assumption of the distribution of data. In reality, classes often display non-normal distributions, which can be difficult to correct.

In addition, the results show the unstable nature of the BPNN, which produces a large variation in the accuracy distribution. It can be assumed that the BPNN is unable to maintain variation in input and training data, whereas the SSOM is more stable and robust. This classifier provides high accuracy with very small variation. Uncertainty in input and training data has only a slight effect on the classification accuracy of the SSOM indicating that it outperforms the GMLC and the BPNN.

Table 4.5. Statistics of classification accuracy of GMLC, BPNN, and SSOM.

(a) Derived from 500 different simulated input data

	Overall Accuracy (%)			Kappa Coefficient		
	GMLC	BPNN	SSOM	GMLC	BPNN	SSOM
Minimum	74.20	50.56	85.88	0.6560	0.3408	0.8117
Maximum	85.12	89.84	91.68	0.8016	0.8645	0.8891
Mean	81.01	78.09	87.54	0.7468	0.7079	0.8339
Standard Deviation	1.70	6.15	1.21	0.0227	0.0820	0.0161

(b) Derived from 500 different randomly training data

	Overall Accuracy (%)			Kappa Coefficient		
	GMLC	BPNN	SSOM	GMLC	BPNN	SSOM
Minimum	74.04	70.04	84.08	0.6539	0.6005	0.7877
Maximum	84.72	86.92	89.64	0.7963	0.8256	0.8619
Mean	79.58	78.25	86.01	0.7277	0.7100	0.8135
Standard Deviation	1.75	4.16	1.14	0.0234	0.0554	0.0152

Table 4.6. Test of significance difference in accuracy of GMLC, BPNN, and SSOM.

(a) Derived from 500 different simulated input data

	Paired Difference				
	Mean	Std. Deviation	Std. Error Mean	t	Sig. (2-tailed)
Overall Accuracy (%)					
BPNN – GMLC	-2.91	62.69	0.28	-10.40	< .001
SSOM – GMLC	6.53	14.42	0.06	101.28	< .001
SSOM – BPNN	9.45	61.63	0.28	34.27	< .001
Kappa Coefficient					
BPNN – GMLC	-0.0389	0.0836	0.0038	-10.40	< .001
SSOM – GMLC	0.0871	0.0192	0.0009	101.28	< .001
SSOM – BPNN	0.1259	0.0822	0.0037	34.27	< .001

(b) Derived from 500 different random training data

	Paired Difference				
	Mean	Std. Deviation	Std. Error Mean	t	Sig. (2-tailed)
Overall Accuracy (%)					
BPNN – GMLC	-1.33	4.49	0.20	-6.59	< .001
SSOM – GMLC	6.44	2.08	0.09	69.14	< .001
SSOM – BPNN	7.76	4.34	0.19	40.02	< .001
Kappa Coefficient					
BPNN – GMLC	-0.0177	0.0599	0.0027	-6.59	< .001
SSOM – GMLC	0.0858	0.0278	0.0012	69.14	< .001
SSOM – BPNN	0.1035	0.0578	0.0026	40.02	< .001

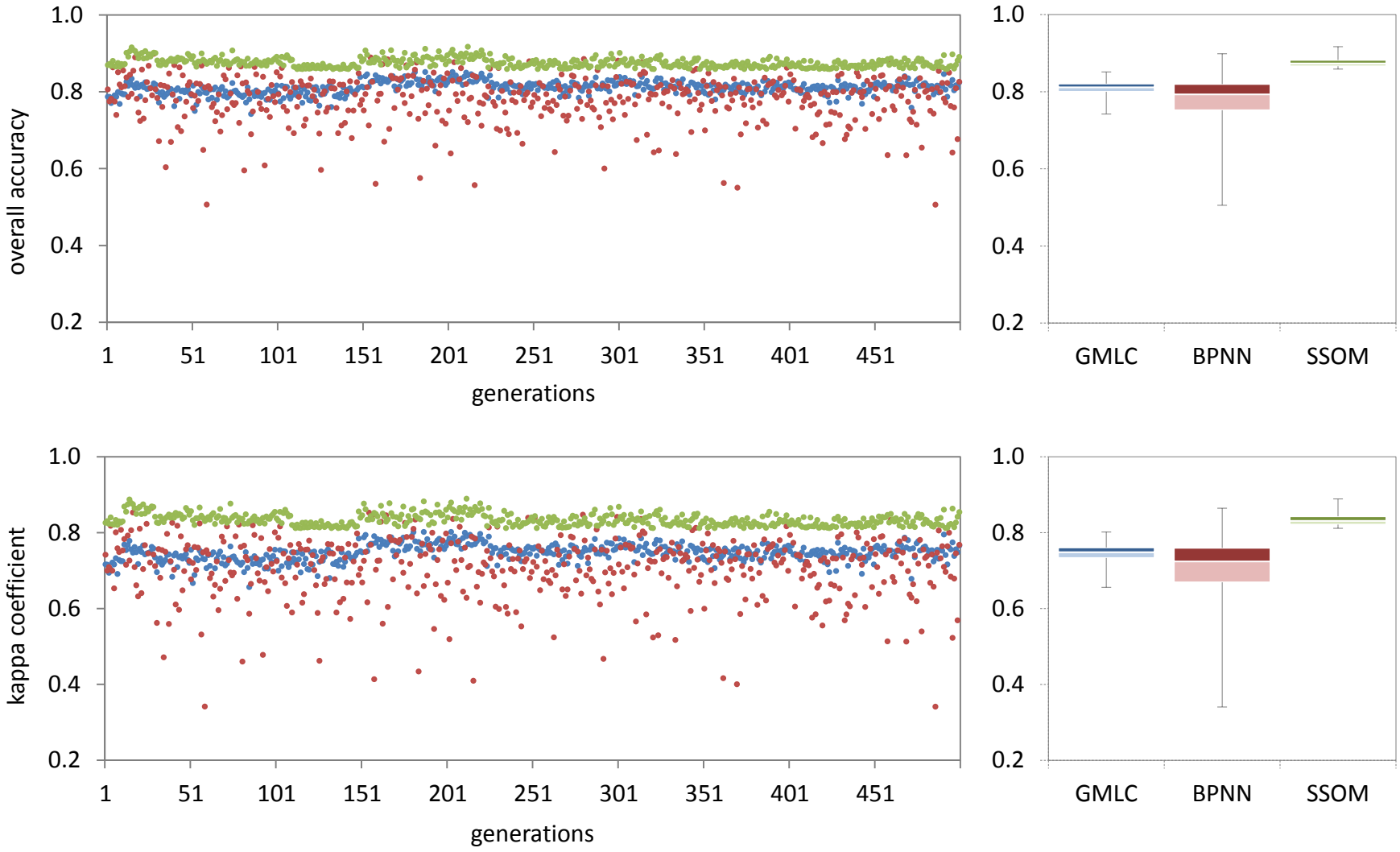


Figure 4.8. Distribution of classification accuracy of GMLC, BPNN, and SSOM in different simulated input data

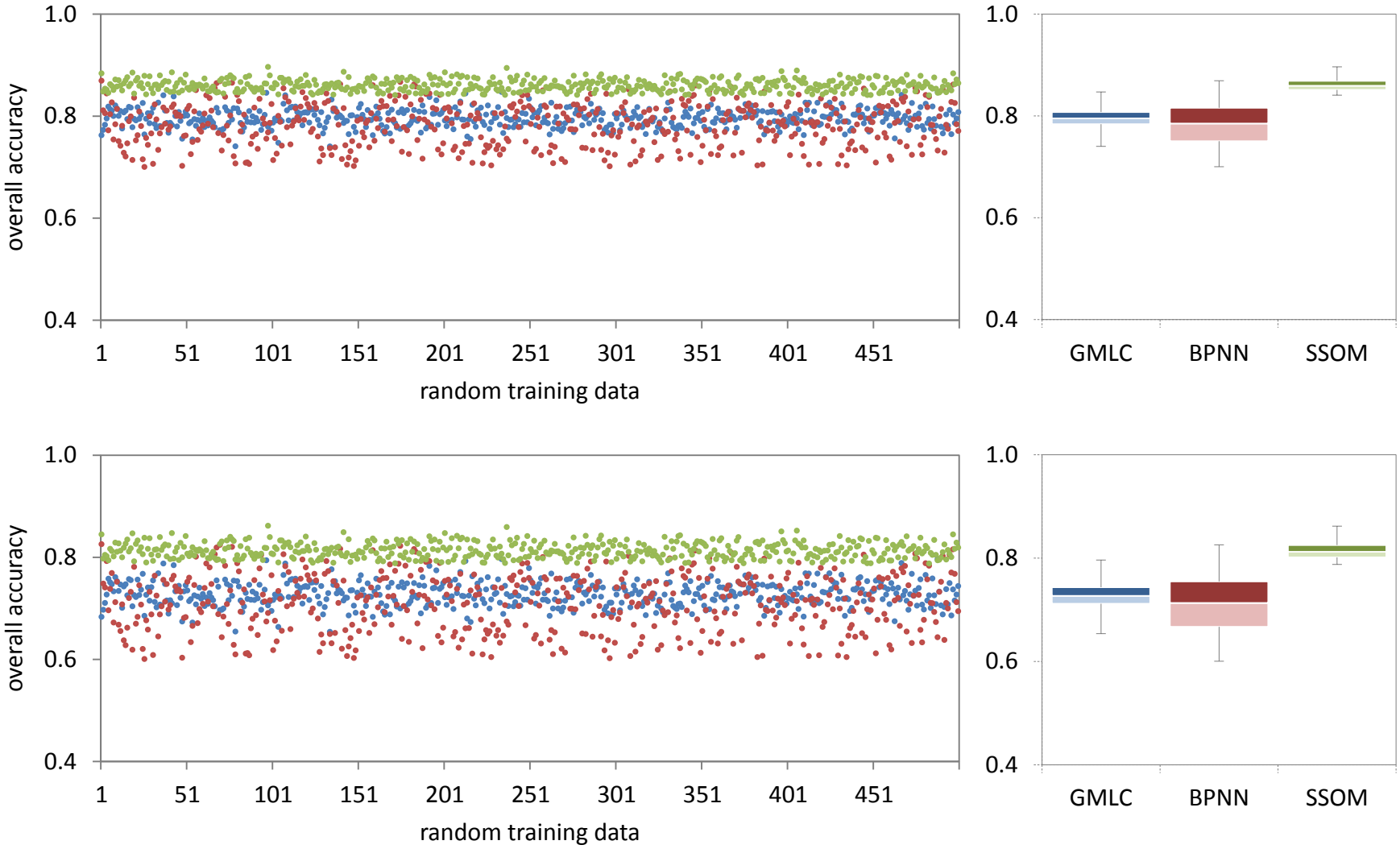
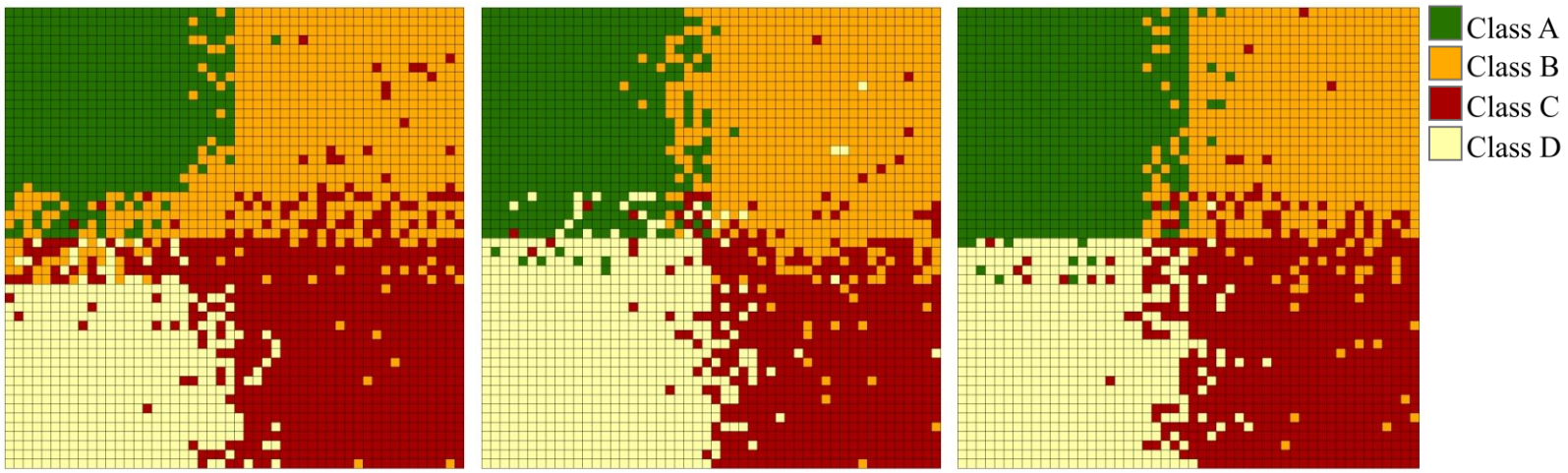
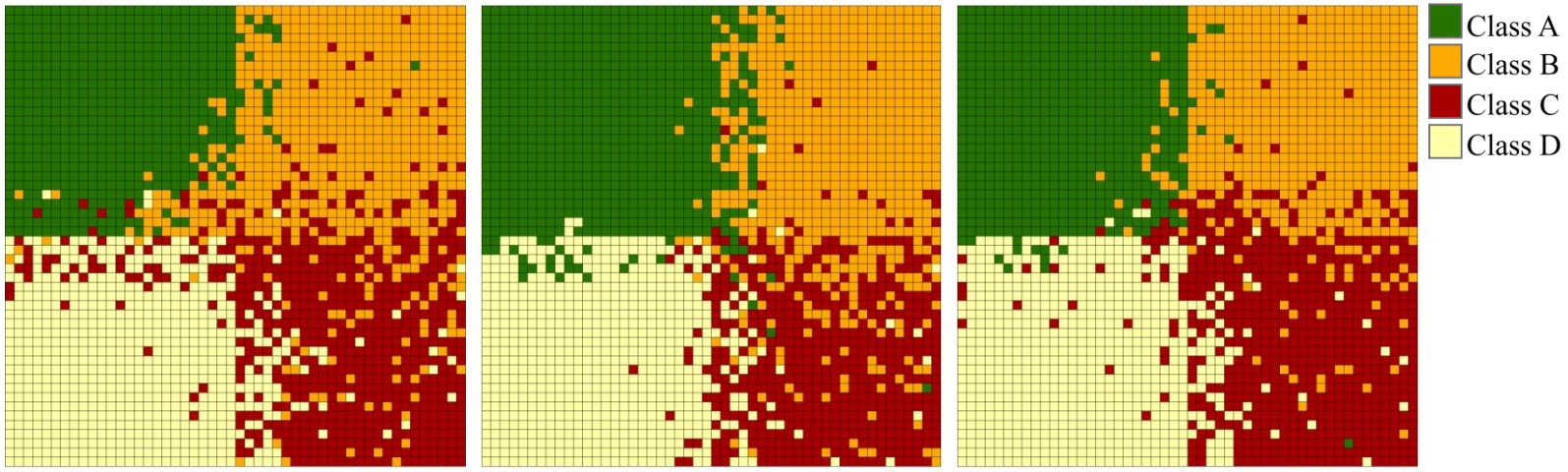


Figure 4.9. Distribution of classification accuracy of GMLC, BPNN, and SSOM in different random training data



(a)



(b)

Figure 4.10. The classified images of GMLC, BPNN, and SSOM providing the highest accuracy (a) in different simulated input data and (b) in different random training data.

4.4 Comparative evaluation of 3SOM with SSOM

This experiment involves a comparative evaluation between soft and hard classification. In general, the supervised self-organizing map classifier (SSOM) is employed as a hard classifier. To improve the efficiency and accuracy of the classification, the SSOM is modified and implemented as a soft classifier in order to improve the classification accuracy. This new classifier is called a soft-supervised self-organizing map classifier (3SOM).

For comparative purposes, the 3SOM and SSOM are utilized with the same input and training data. Both of these classifiers are also implemented with the same suitable neural network configuration, which is comprised of 23 input layer neurons, 4 output layer neurons, 6 × 6competitive layer neurons, an initial learning rate of 0.075, and 50 iterations, as shown in Table 4.4.

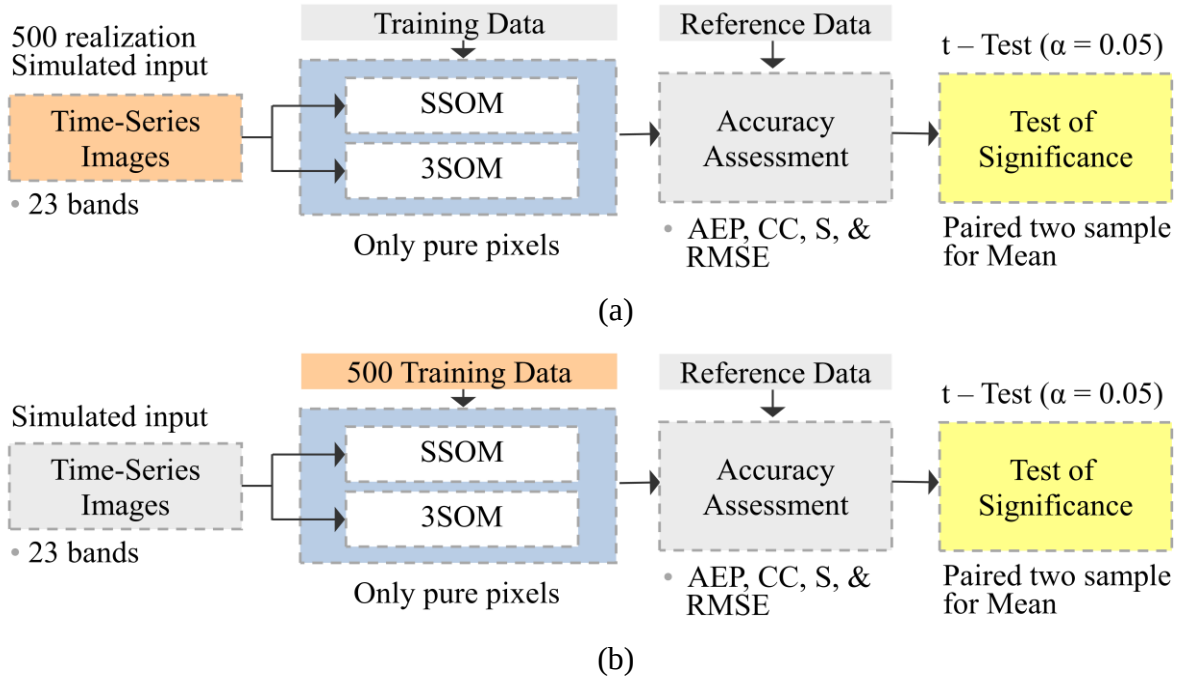


Figure 4.11 Experimental procedures of comparative evaluation between 3SOM and SSOM (a) in different simulated input data and (b) in different random training data.

Similar to the previous experiment, two tests are also performed in this experiment (Figure 4.11). The first test is a comparative evaluation of the 3SOM and SSOM with different simulated input data, and the second test is a comparative evaluation between the 3SOM and SSOM with different random training data. The details of each test are the same as described in section 4.3.

The output of the 3SOM (a soft classification) is a multi-layer image, one layer for each land cover class, whereas the SSOM (a hard classification) produces a single-layer image. To compare the performance of both classifiers, the SSOM output is broken down into a multi-layer image. Then, outputs of both classifiers are evaluated by using four accuracy assessment measures for soft classification: 1) area error proportion (AEP), 2) correlation coefficient (CC), 3) root mean square error (RMSE), and 4) closeness (S). In addition, a t-test is performed to determine whether there is a statistically significant difference between the 3SOM and SSOM at $\alpha = 0.05$.

Examination of the accuracy assessment measures reveals that increasing CCs, means of S (MS), and absolute AEPs, and decreasing RMSEs, are positively related to classification accuracy.

The results of both tests clearly illustrates that the classification accuracy of the 3SOM is higher than the SSOM in the CC, RMSE, and MS (Figure 4.12 and 4.13). The 3SOM provides higher means of the CC and lower means of the RMSE than the SSOM in all classes. Moreover, the 3SOM also yields smaller mean of MS than the SSOM (Table 4.7). Table 4.8 shows results of the t-statistic of paired difference between both classifiers. The results clearly show that the mean accuracy in all measures of both classifiers are statistically significantly different at an

alpha level of 0.05 in both tests. Moreover, the mean difference of the CC, RMSE, and MS also indicates that the 3SOM results are considerably more accurate than that of SSOM.

Table 4.7. Mean of classification accuracy of SSOM and 3SOM.

(a) Derived from 500 different simulated input data

	AEP		CC		RMSE		MS	
	SSOM	3SOM	SSOM	3SOM	SSOM	3SOM	SSOM	3SOM
Class A	-0.02	0.11	0.89	0.96	0.21	0.10	-	-
Class B	0.09	0.01	0.81	0.88	0.25	0.17	-	-
Class C	-0.10	-0.15	0.78	0.83	0.28	0.20	-	-
Class D	0.03	0.03	0.86	0.93	0.22	0.13	-	-
Entire image	-	-	-	-	-	-	0.06	0.02

(b) Derived from 500 different random training data

	AEP		CC		RMSE		MS	
	SSOM	3SOM	SSOM	3SOM	SSOM	3SOM	SSOM	3SOM
Class A	0.04	0.14	0.88	0.96	0.21	0.11	-	-
Class B	0.09	0.00	0.80	0.86	0.25	0.18	-	-
Class C	-0.15	-0.17	0.75	0.81	0.30	0.21	-	-
Class D	0.02	0.03	0.84	0.92	0.24	0.14	-	-
Entire image	-	-	-	-	-	-	0.06	0.03

However, different results are found for AEP in both tests. The first test shows that the SSOM provides more accurate results than the 3SOM in class A and C. The second test shows that the SSOM provides more accurate results than the 3SOM in all classes except class B. Although overall the SSOM tends to achieve more accurate results than the 3SOM in AEP, it also provides a larger variation of AEP in all classes. In addition, although AEP is a valuable

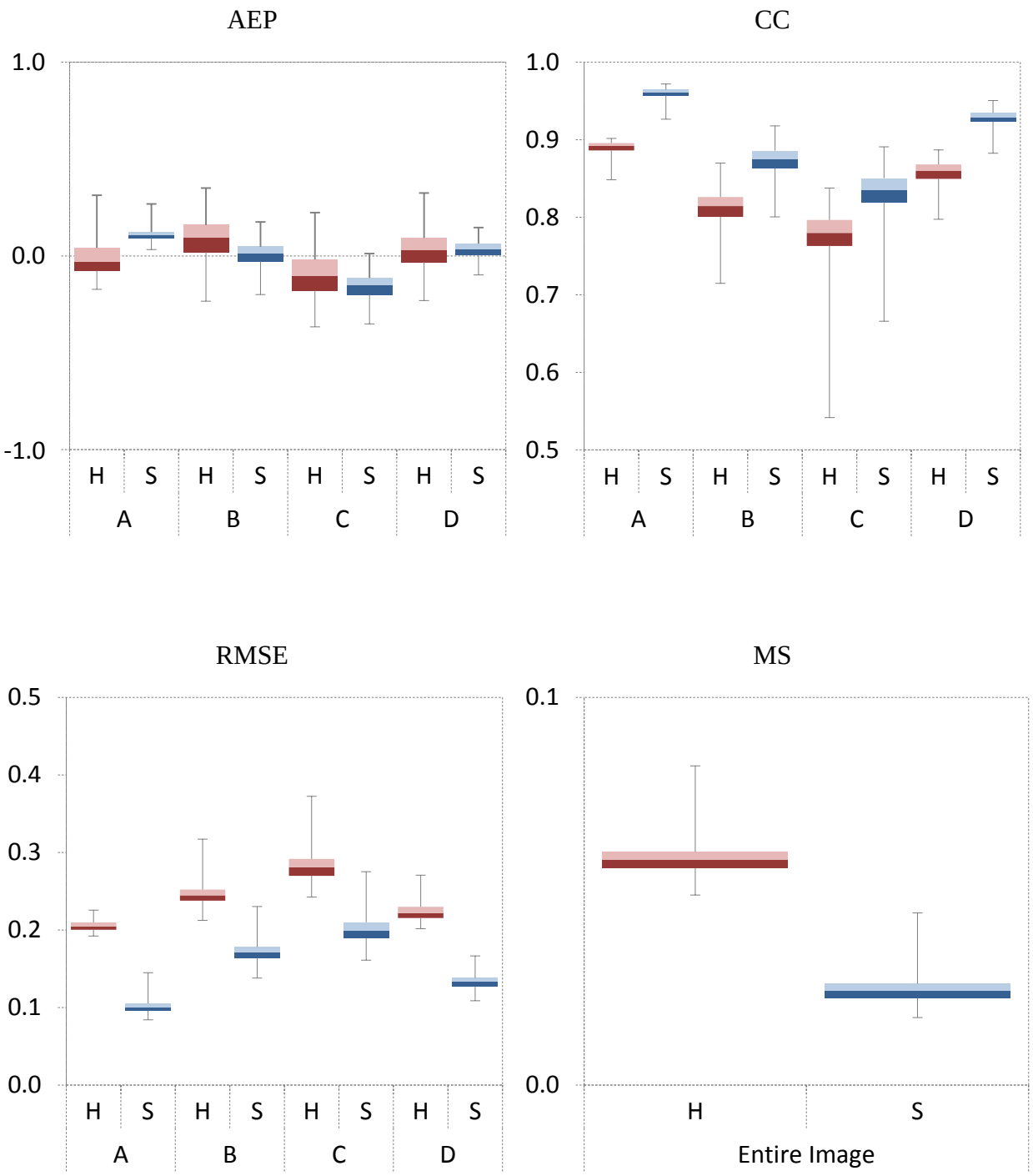
measure to assess the degree of the area error proportion, the use of only AEP is not sufficient to assess the performance and accuracy of soft classification. This is due to the fact that AEP does not take the spatial distribution of omission and commission errors into account.

Consequently, according to this experiment, it can be concluded that the 3SOM, which is employed as a soft classification, enables a more accurate classification than the SSOM, which is applied for a hard classification. The supported reason is that hard classification uses the same principle by assigning each pixel to a single class. In reality, many pixels in an image may represent more than one land cover class on the ground. To allocate a mixed pixel to a single land cover class not only provides an unrealistic result, but also leads to an inaccurate representation of land cover (Thornton *et al.*, 2006).

4.5 Comparative evaluation of fully-3SOM with partially-3SOM

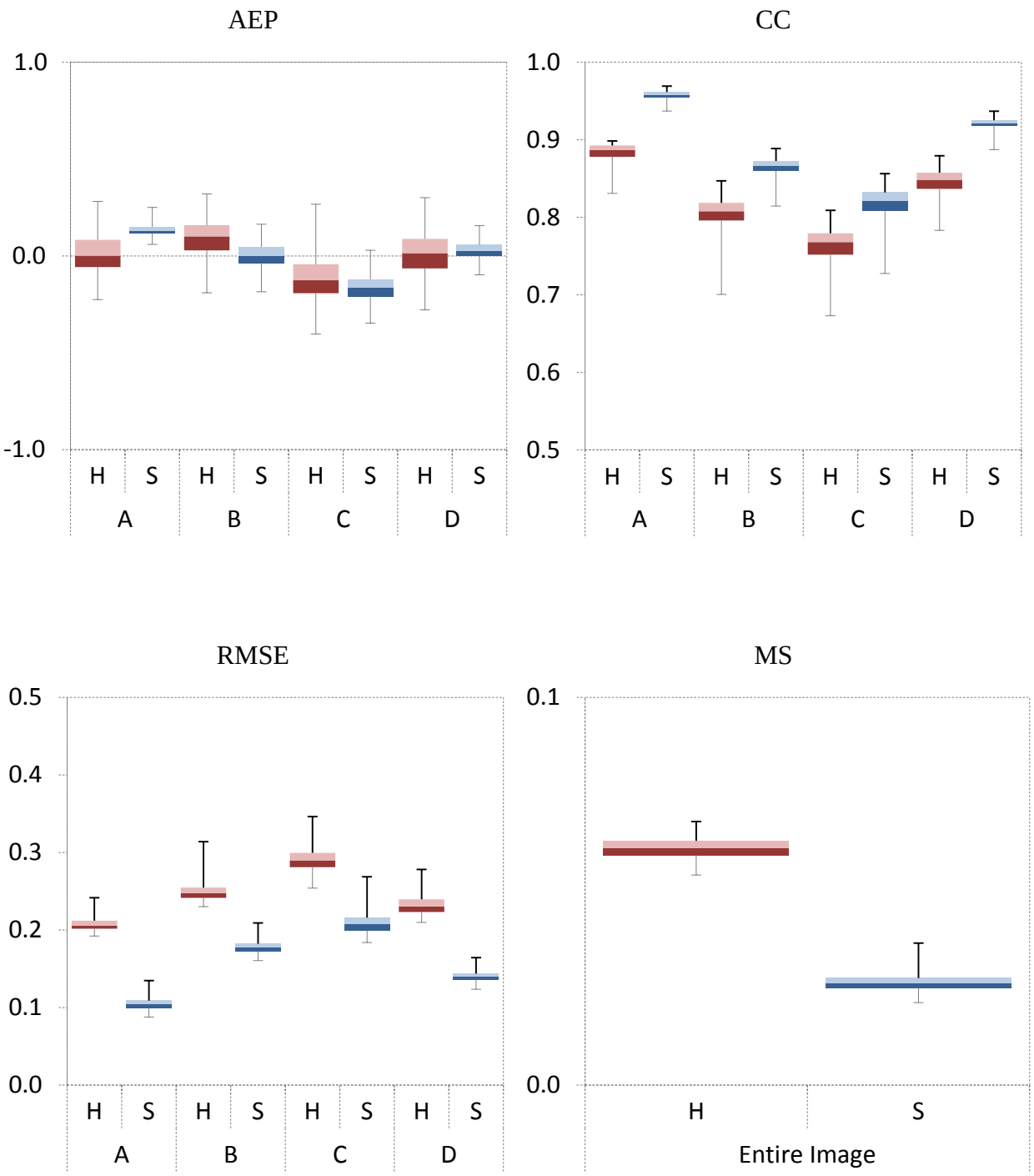
In the previous experiment, the results demonstrate that a soft classification is a more effective method for land cover classification as compared to a hard classification. However, in most studies, soft classifications generally deal with class mixing only in the classification stage, but do not accommodate class mixing in the ground reference data used in training and testing stages.

For that reason, this experiment is aimed to investigate the advantages of the soft-supervised self-organizing map (3SOM) by comparing the accuracy derived from the fully-soft classification (3SOM-F) to the partially-soft classification (3SOM-P). The 3SOM-P is provided with a hard training dataset consisting of only pure pixels, whereas the 3SOM-F is provided with a soft training dataset containing both pure and mixed pixels.



Remark: H = SSOM; S = 3SOM

Figure 4.12. Distribution of classification accuracy of SSOM and 3SOM in different simulated input data.



Remark: H = SSOM; S = 3SOM

Figure 4.13. Distribution of classification accuracy of SSOM and 3SOM from different random training data.

Table 4.8. Test of significance difference in accuracy between SSOM and 3SOM.

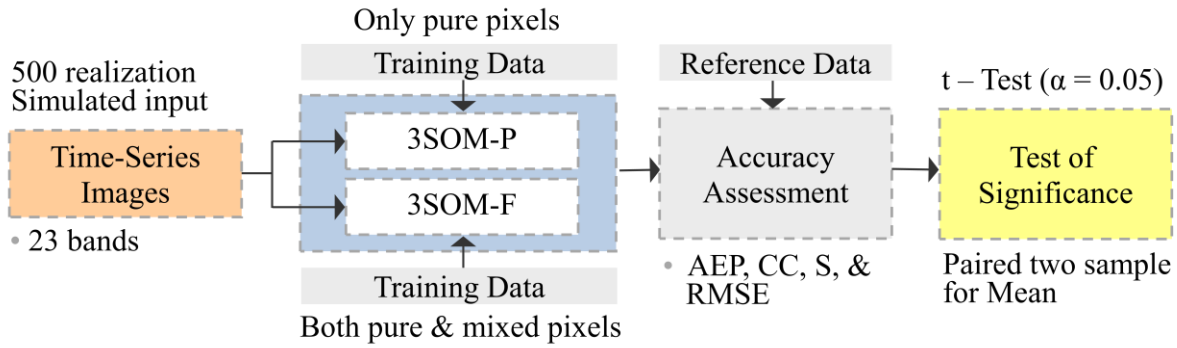
(a) Derived from 500 different simulated input data

	Paired Difference				
	Mean	Std. Deviation	Std. Error Mean	t	Sig. (2-tailed)
Area error proportion					
Class A	0.0393	0.0520	0.0023	16.89	< .001
Class B	-0.0625	0.0755	0.0034	-18.50	< .001
Class C	0.0331	0.0869	0.0039	8.50	< .001
Class D	-0.0338	0.0650	0.0029	-11.64	< .001
Correlation coefficient					
Class A	0.0703	0.0094	0.0004	167.73	< .001
Class B	0.0611	0.0147	0.0007	93.09	< .001
Class C	0.0549	0.0188	0.0009	65.26	< .001
Class D	0.0704	0.0138	0.0006	114.39	< .001
Root mean square error					
Class A	-0.1046	0.0087	0.0004	-267.87	< .001
Class B	-0.0740	0.0116	0.0005	-143.21	< .001
Class C	-0.0816	0.0147	0.0007	-124.22	< .001
Class D	-0.0905	0.0109	0.0005	-185.16	< .001
Mean of closeness					
Entire image	-0.0336	0.0026	0.0001	-290.14	< .001

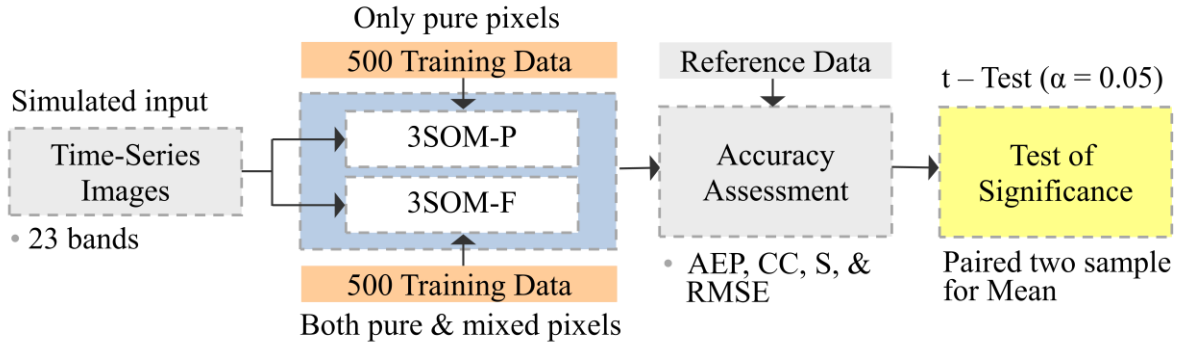
Table 4.8. (cont'd).

	Paired Difference				
	Mean	Std. Deviation	Std. Error Mean	t	Sig. (2-tailed)
Area error proportion					
Class A	0.0539	0.0625	0.0020	27.25	< .001
Class B	-0.0653	0.0749	0.0024	-27.56	< .001
Class C	0.0301	0.0929	0.0029	10.25	< .001
Class D	-0.0485	0.0707	0.0022	-21.70	< .001
Correlation coefficient					
Class A	0.0736	0.0124	0.0004	187.98	< .001
Class B	0.0595	0.0157	0.0005	120.14	< .001
Class C	0.0552	0.0215	0.0007	81.13	< .001
Class D	0.0748	0.0148	0.0005	159.49	< .001
Root mean square error					
Class A	-0.1024	0.0103	0.0003	-313.98	< .001
Class B	-0.0718	0.0123	0.0004	-184.55	< .001
Class C	-0.0828	0.0153	0.0005	-171.29	< .001
Class D	-0.0922	0.0117	0.0004	-249.15	< .001
Mean of closeness					
Entire image	-0.0347	0.0028	0.0001	-395.31	< .001

For comparative purposes, the sample size in the hard training dataset is the same as in the soft training dataset, and the pure pixels in the soft training data are memberships in the hard training data. The sample size of both training datasets is 240 samples. The 3SOM-P is trained by all samples (100% training data) for only pure pixels, while the 3SOM-F is trained by 96 samples (40% training data) for pure pixels and 144 samples (60% training data) for mixed pixels.



(a)



(b)

Figure 4.14. Experimental procedures of comparative evaluation between 3SOM-F and 3SOM-P (a) in different simulated input data and (b) in different random training data.

Two tests are also performed to evaluate and compare the classification accuracy between the 3SOM-F and 3SOM-P (Figure 4.14). The first test utilizes 500 different input datasets. All input data are classified with the same training data. For the second test, 500 classification simulations with the same input data are trained by different training datasets. In the classification stage, all simulations are classified by the same suitable neural network configuration, which is comprised of 23 input layer neurons, four output layer neurons, 6×6 competitive layer neurons, an initial learning rate of 0.075, and 50 iterations, as shown in Table 4.4.

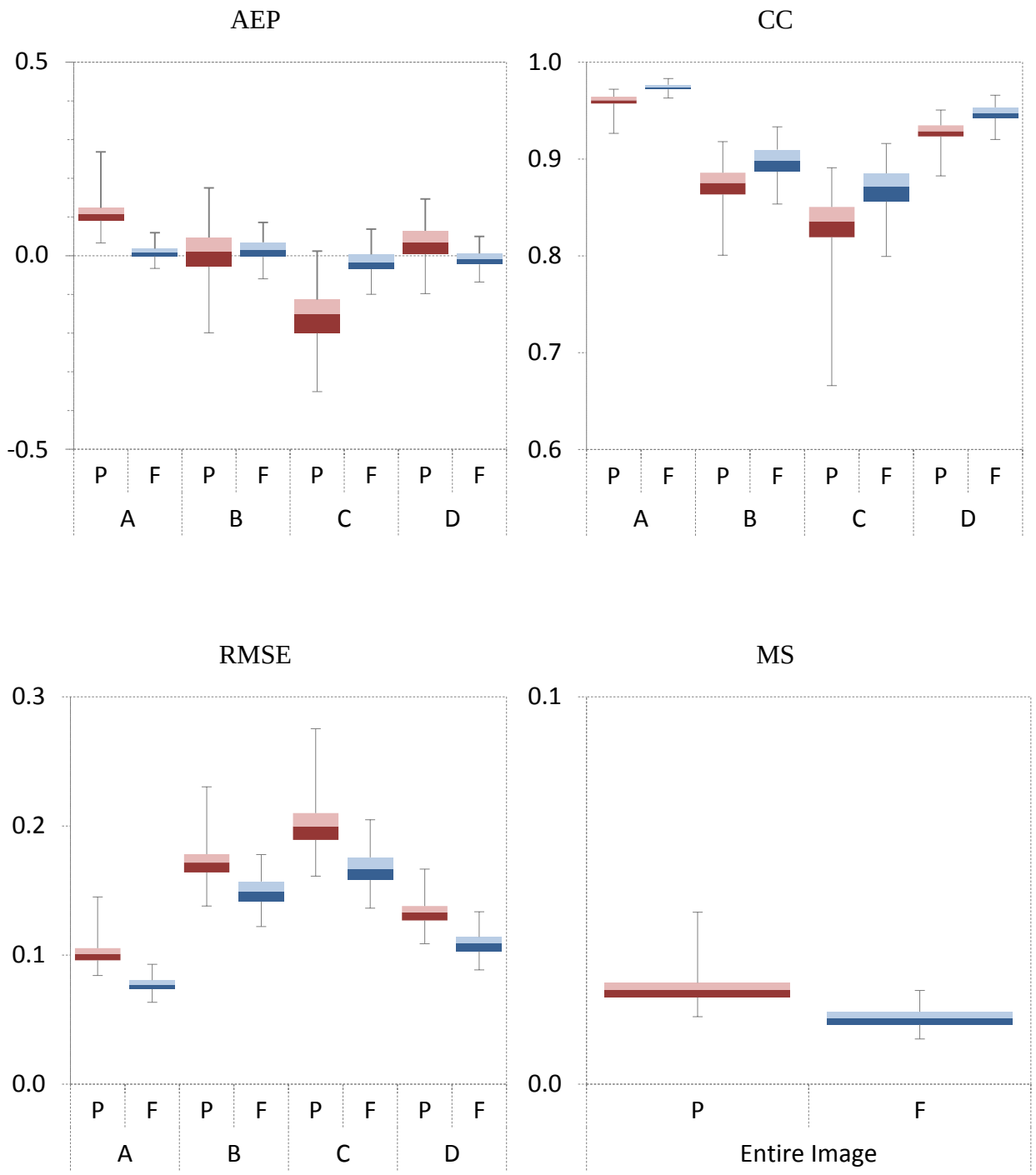
The performance of these two techniques is assessed by evaluating the accuracy of the soft classification. Four measures are calculated including the area error proportion (AEP), correlation coefficient (CC), root mean square error (RMSE), and closeness (S). Finally, a t-test is performed to determine whether there is a statistically significant difference between the two techniques at $\alpha = 0.05$ and which technique is more efficient.

The classification accuracy results of the first test are shown in Figure 4.15 and Table 4.9.a., while the results of the second test are illustrated in Figure 4.16 and Table 4.9.b. Results of both tests indicate that the 3SOM-F, which employs both pure and mixed pixels in the training process, produces more accurate classification results than the 3SOM-P, which utilizes only pure pixels in the training process. Moreover, Table 4.10 presents the test results of the significance differences of mean accuracy between the 3SOM-F and 3SOM-P. The results show that in both the 3SOM-F and 3SOM-P, the mean classification accuracies are strongly significantly different at the 95% confidence level for all classes. Additionally, the mean differences also indicated that the 3SOM-F provides a considerably higher accuracy than the 3SOM-P in all measures of soft classification accuracy assessment.

Furthermore, the results of visual interpretation are supported by the classification accuracy assessment. The classified proportional images of the 3SOM-F and 3SOM-P, which provide the lowest MS of all simulations in both tests, are shown in Figure 4.17 and 4.18. Results of both tests illustrate that classified proportional images produced by the 3SOM-F are more accurate and realistic when comparing to the reference images, especially in the areas of mixed pixels.

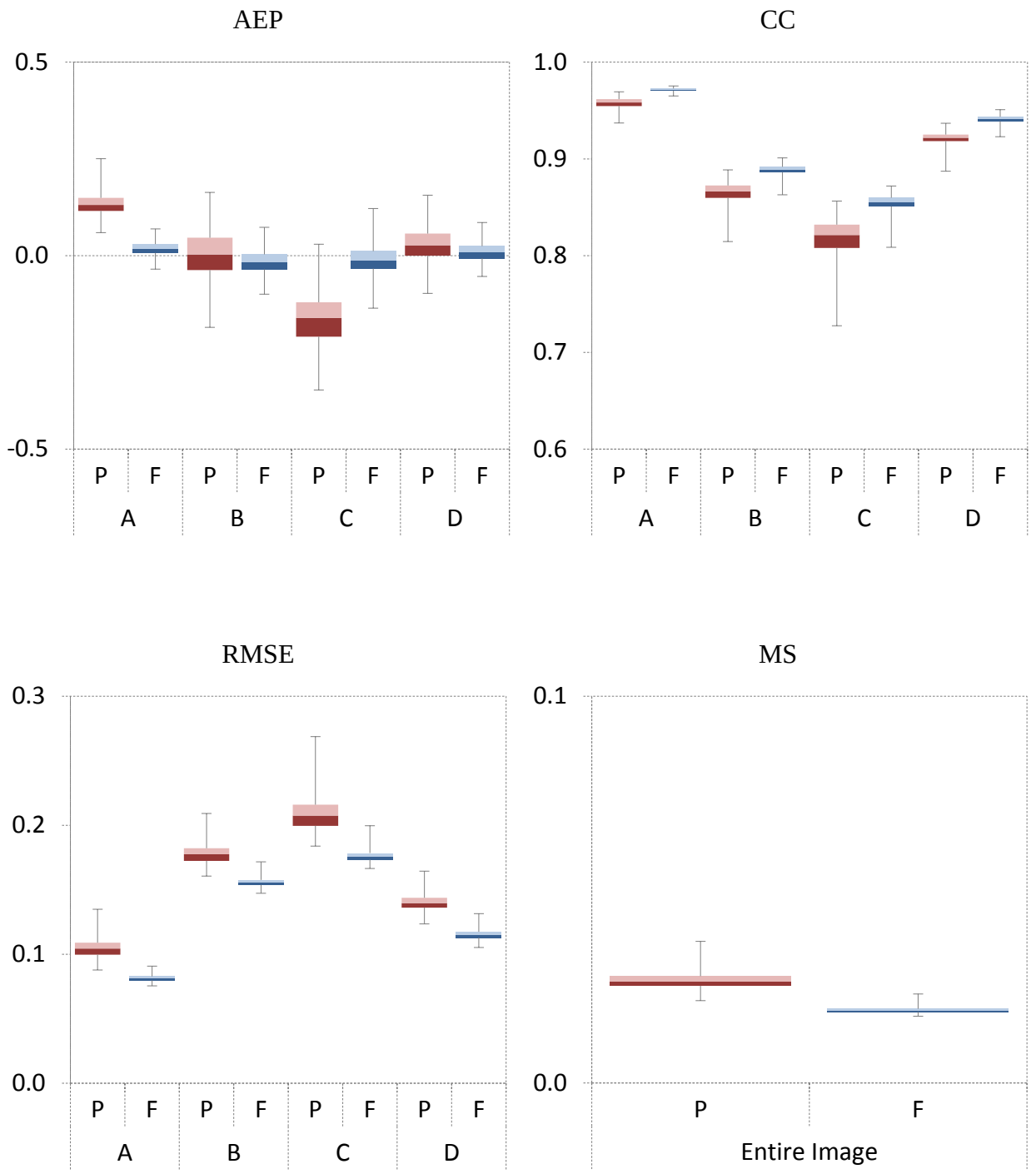
Closeness in an image is an ideal measure for showing the separation between classified images and reference images based on the relative proportion of each class in the pixel. In both tests, the 3SOM-F produces low values of closeness in most of the study area, but the 3SOM-P shows a high value of closeness in the areas of heterogeneity. This indicates that the 3SOM-F achieves higher performance than the 3SOM-P.

As a result, this experiment shows that to perform a soft classification with the 3SOM, the network should be trained with soft training data. This is due to the fact that the dominant classes and subsidiary classes in a pixel can be well recognized if it is trained with pure and mixed pixels. This suggests that the additional variability of spectral signatures introduced to the classifier by these mixed pixels help the network to generalize.



Remark: P = 3SOM-P; F = 3SOM-F

Figure 4.15. Distribution of classification accuracy of 3SOM-F and 3SOM-P in different simulated data.



Remark: P = 3SOM-P; F = 3SOM-F

Figure 4.16. Distribution of classification accuracy of 3SOM-F and 3SOM-P in different random training data.

Table 4.9. Mean of classification accuracy of 3SOM-F and 3SOM-P.

(a) Derived from 500 different simulated input data

	AEP		CC		RMSE		MS	
	P	F	P	F	P	F	P	F
Class A	0.11	0.01	0.96	0.97	0.10	0.08	-	-
Class B	0.01	0.02	0.88	0.90	0.17	0.15	-	-
Class C	-0.15	-0.02	0.83	0.87	0.20	0.17	-	-
Class D	0.03	-0.01	0.93	0.95	0.13	0.11	-	-
Entire image	-	-	-	-	-	-	0.02	0.01

(b) Derived from 500 different random training data

	AEP		CC		RMSE		MS	
	P	F	P	F	P	F	P	F
Class A	0.14	0.02	0.96	0.97	0.10	0.08	-	-
Class B	0.00	-0.02	0.87	0.89	0.18	0.16	-	-
Class C	-0.18	-0.01	0.82	0.85	0.21	0.18	-	-
Class D	0.03	0.01	0.92	0.94	0.14	0.12	-	-
Entire image	-	-	-	-	-	-	0.02	0.01

Remark: P = 3SOM-P; F = 3SOM-F

Table 4.10. Test of significance difference in accuracy between 3SOM-F and 3SOM-P.

(a) Derived from 500 different simulated input data

	Paired Difference				
	Mean	Std. Deviation	Std. Error Mean	t	Sig. (2-tailed)
Area error proportion					
Class A	-0.0942	0.0256	0.0012	-82.12	< .001
Class B	-0.0205	0.0379	0.0017	-12.06	< .001
Class C	-0.1261	0.0627	0.0028	-44.97	< .001
Class D	-0.0277	0.0356	0.0016	-17.37	< .001
Correlation coefficient					
Class A	0.0137	0.0046	0.0002	65.92	< .001
Class B	0.0239	0.0124	0.0006	43.02	< .001
Class C	0.0371	0.0197	0.0009	42.14	< .001
Class D	0.0189	0.0068	0.0003	62.02	< .001
Root mean square error					
Class A	-0.0239	0.0066	0.0003	-81.15	< .001
Class B	-0.0229	0.0097	0.0004	-53.11	< .001
Class C	-0.0337	0.0135	0.0006	-55.77	< .001
Class D	-0.0241	0.0069	0.0003	-78.25	< .001
Mean of closeness					
Entire image	-0.0075	0.0022	0.0001	-74.95	< .001

Table 4.10. (cont'd.).

	Paired Difference				
	Mean	Std. Deviation	Std. Error Mean	t	Sig. (2-tailed)
Area error proportion					
Class A	-0.1118	0.0293	0.0009	-120.59	< .001
Class B	-0.0200	0.0382	0.0012	-16.60	< .001
Class C	-0.1376	0.0717	0.0023	-60.67	< .001
Class D	-0.0204	0.0348	0.0011	-18.50	< .001
Correlation coefficient					
Class A	0.0138	0.0053	0.0002	83.02	< .001
Class B	0.0231	0.0114	0.0004	63.83	< .001
Class C	0.0359	0.0199	0.0006	57.05	< .001
Class D	0.0195	0.0065	0.0002	95.02	< .001
Root mean square error					
Class A	-0.0235	0.0072	0.0002	-103.61	< .001
Class B	-0.0222	0.0083	0.0003	-85.02	< .001
Class C	-0.0325	0.0135	0.0004	-76.23	< .001
Class D	-0.0246	0.0066	0.0002	-117.17	< .001
Mean of closeness					
Entire image	-0.0077	0.0022	0.0001	-112.39	< .001

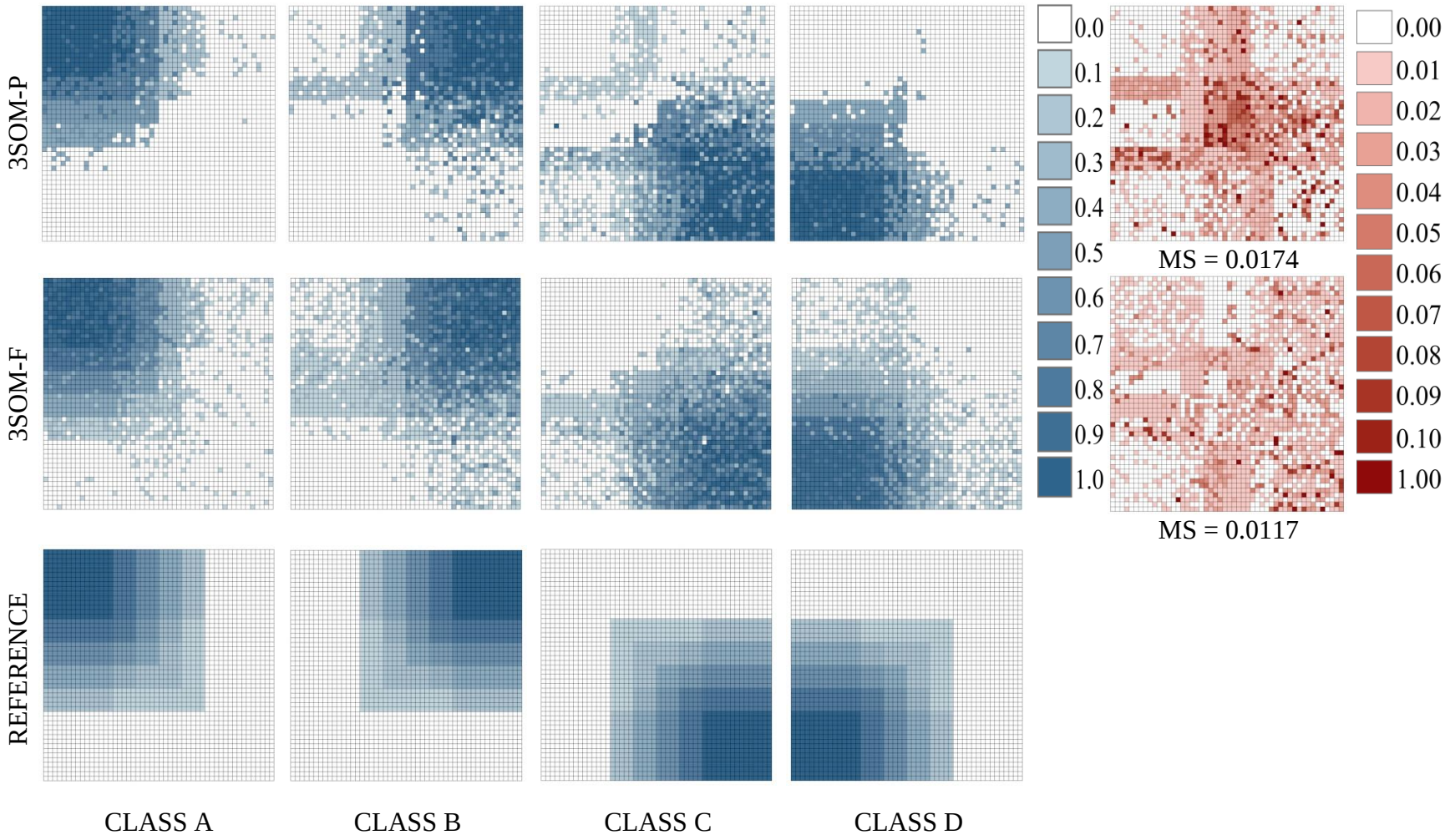


Figure 4.17. The classified proportional images of 3SOM-P, and 3SOM-F providing the lowest MS of all simulations in different simulated input data.

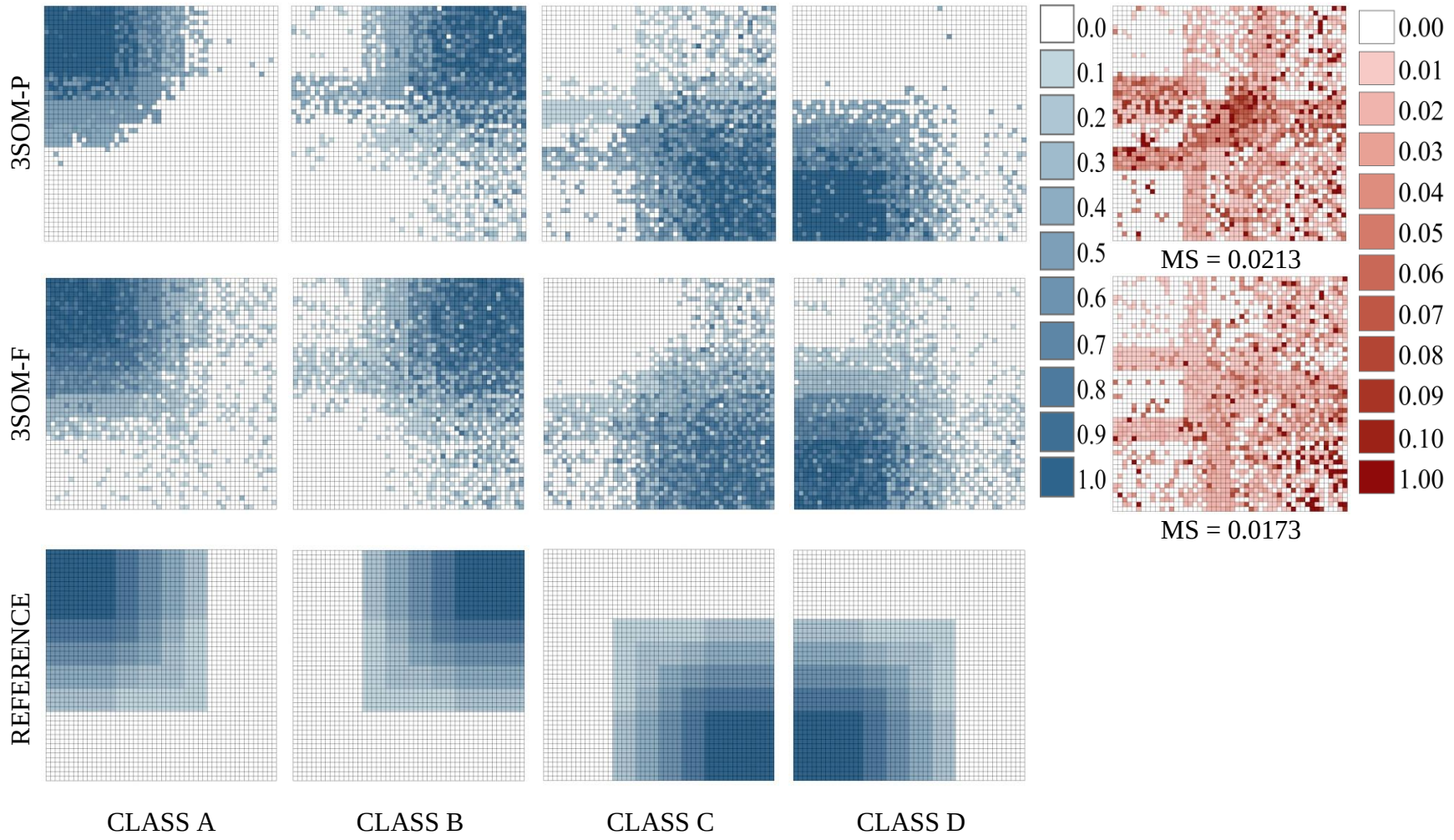


Figure 4.18. The classified proportional images of 3SOM-P, and 3SOM-F providing the lowest MS of all simulations in different random training data.

4.6 SSOM with uncertainty in classification accuracy

Land cover maps derived from remotely sensed data classification is universally used and are arguably the most important terrestrial data available. Previous experiments confirm that the supervised self-organizing map (SSOM) is an efficient method for image classification and also is often used to create land cover maps. However, the classification accuracy in an inference process is always less than the desired accuracy in the actual classification process, thus this marginalized difference is considered to be an element of uncertainty in the classification results. Failure to recognize uncertainty may lead to erroneous and misleading interpretations. Therefore, the aim of this experiment is to evaluate the uncertainty in the classification accuracy by considering the impact of possible factors on the spatial variation of the SSOM classification accuracy.

In this research, only the synthetic data is used to evaluate the classification uncertainty. A Monte Carlo simulation technique is applied to assess the reliability of the classification output by focusing on the uncertainty associated with the input data, training data, and the classifier itself.

4.6.1 Uncertainty associated with input data

The uncertainty in the input data is associated with the variations of environmental conditions (e.g. land management practices, climate change, atmosphere interactions, soil fertility) and data preprocessing. These variations have an influence on the classification accuracy. To evaluate the uncertainty in classification accuracy associated with input data, different levels of noise represented for the degrees of variations are added to the original input images. The noise is derived from a random number generator based on a normal distribution

using the extracted mean and standard deviation of each class. Five levels of noise are based on the degree of standard deviation from ±1.0 to 2.0.

The process is shown in Figure 4.19. The classification is performed using the SSOM classifier and is run 500 times with the same training data and varied input images. Each time, the classification provides a different realization of the classification accuracy. The 500 accuracy results are used to generate the classification accuracy distribution by using box plots and to create the accuracy possibility for each pixel. Then, the process is repeated for other standard deviation values.

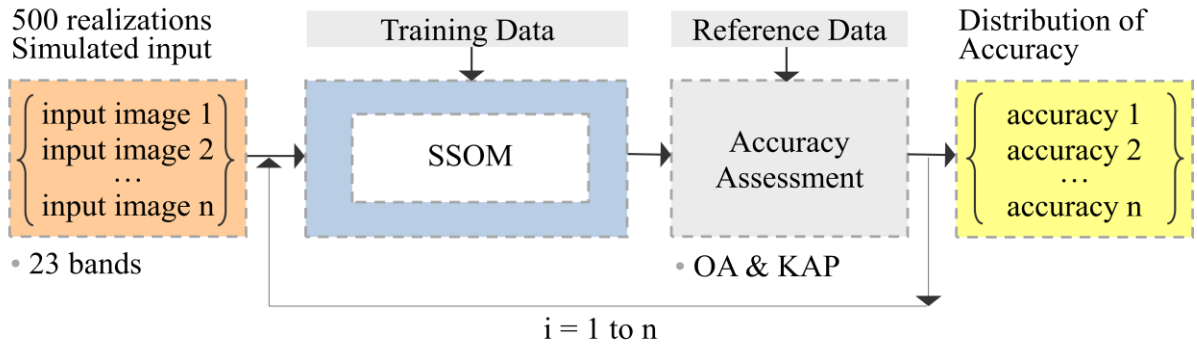


Figure 4.19. Experimental procedure to evaluate the classification uncertainty associated with the input data.

The results of an analysis on how the deviation in input data affects the uncertainty in the SSOM classification accuracy are shown in Figure 4.20 and Figure 4.21. The distributions of overall accuracy (OA) and Kappa coefficient (KAP) in Figure 4.20 illustrate that there is a negative relationship between the levels of noise and the classification accuracy. In other words, when levels of noise increase, classification accuracy decreases. Moreover, images of accuracy possibility (Figure 4.21) show that the possibility of accuracy decreases when the levels of noise increase. The results clearly show this strong effect particularly in the areas of mixed pixels.

These issues are typically found in image classification and they are major challenges for improving classifier effectiveness.

Interestingly, increasing the levels of noise does not tend to impact the uncertainty of classification accuracy after repeating 500 classifications with different levels of noise from ±1.0 to 2.0 standard deviations. The box plots show that there are small variations in the accuracy distribution, even though there is an increase in the levels of noise. These results indicate that the uncertainty in an input image has a small effect on the SSOM; therefore, it is a stable and robust classifier that provides precise accuracy.

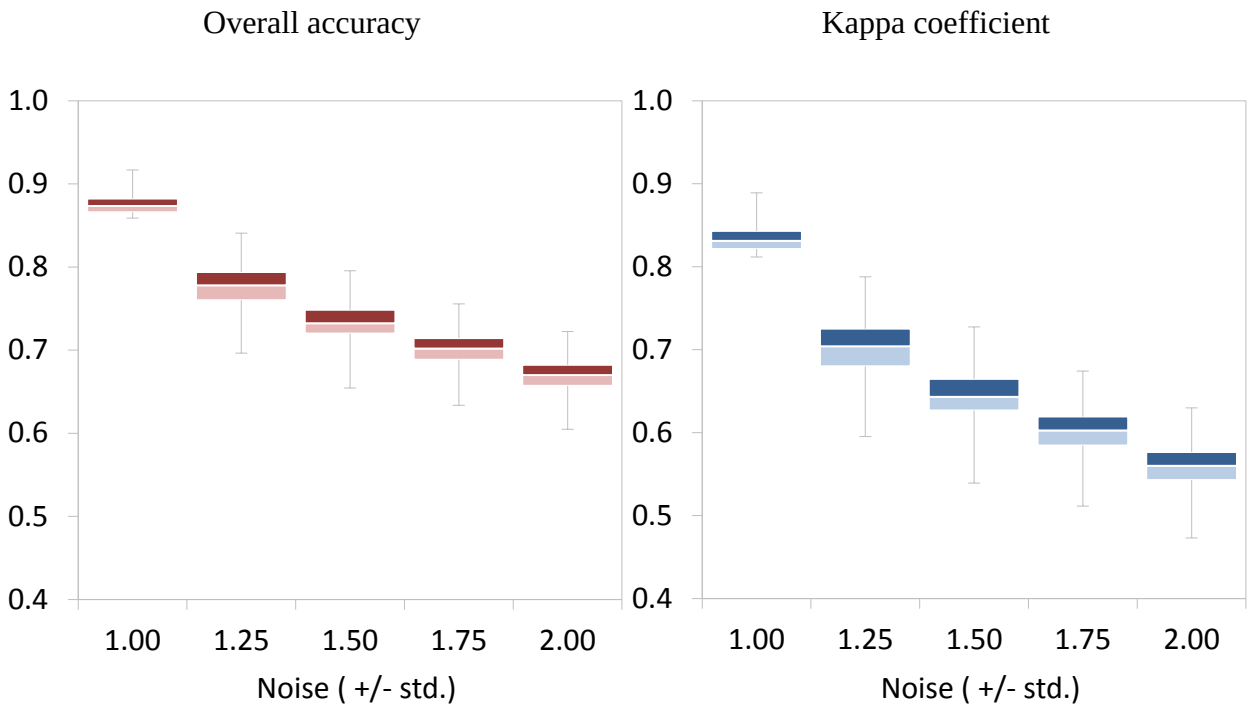


Figure 4.20. Distribution of classification accuracy of SSOM in different levels of noise.

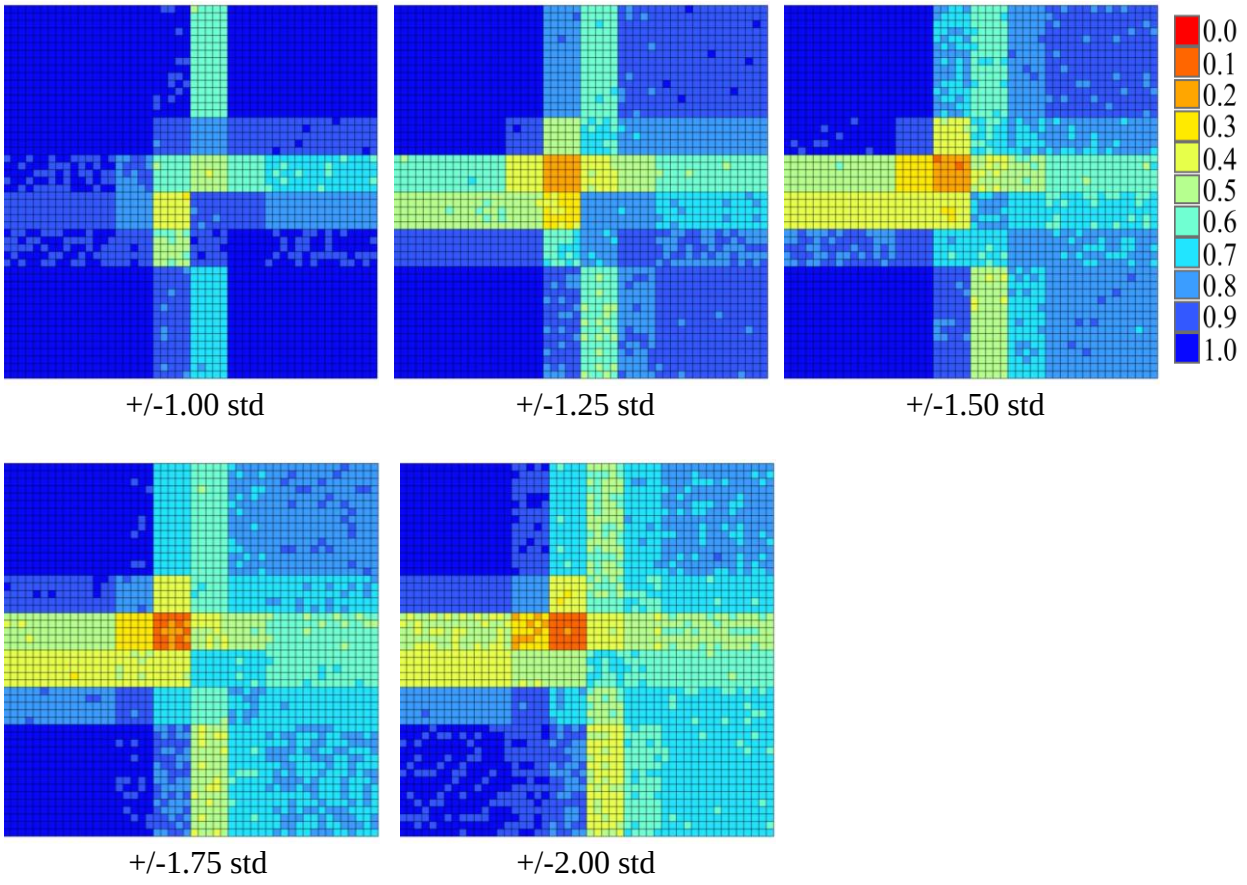


Figure 4.21. Images of accuracy possibility derived from SSOM in different levels of noise.

4.6.2 Uncertainty associated with training data

Due to the fact that training data is not only an important component of the classification process but also to the accuracy of the classification, two criteria are adopted for generating training datasets: random selection and a random shuffling sequence. These criteria are established to study the impact of training data on the classification accuracy. There are 500 training datasets for each criterion. The first set of 500 is generated by randomly selecting the training data, and the second set is created by a random shuffling sequence within the same training samples. For each set of training data, the classification is performed using the SSOM and run 500 times with the same input data but different training data. Classification accuracy is

assessed by comparing the output of each classification with the reference data. Then, the classification accuracy distribution and accuracy possibility image are generated to evaluate the uncertainty in classification accuracy associated with training data. The process is shown in Figure 4.22.

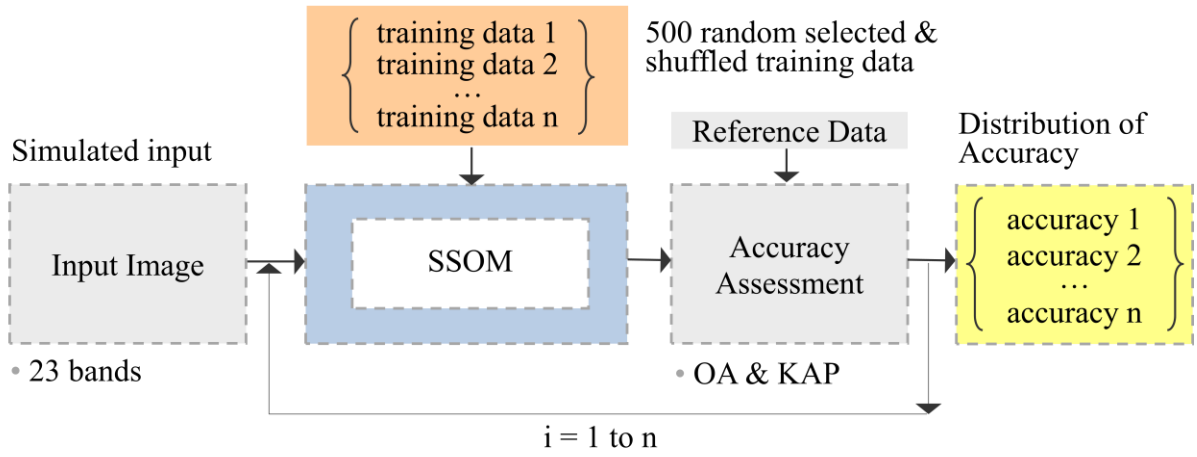


Figure 4.22. Experimental procedure to evaluate the classification uncertainty associated with the training data.

Figure 4.23 shows the classification accuracy distribution of the random selection and shuffling sequence training datasets. The box plots show that both criteria for generating set of training data have small variations of classification accuracy. Additionally, Figure 4.24 illustrates the accuracy possibility images, which signify that both criteria generate similar uncertainty in classification accuracy. The results reveal that both training datasets have little or no impact on the homogenous areas, whereas the accuracy possibility in heterogeneous areas tends to be sensitive to both datasets.

Although the training samples are randomly selected for creating the training data, the number of training samples for each class is restricted and associated with the total number of class samples in order to maintain the class distribution corresponding to that of the full dataset.

This procedure is applied to all randomly selected training datasets. According to this controlled condition, the random selection training data has a small influence on classification accuracy of the SSOM. Moreover, selecting training samples from the synthetic image can assure that each sample correctly corresponds to the actual class. Therefore, based on synthetic data, randomly selecting training the data slightly affects the uncertainty in classification accuracy. In general, the uncertainty in classification accuracy likely depends on the correctness of the training sample more than the random selection of the training sample. For that reason, selecting incorrect training samples has a considerable impact on the accuracy possibility and causes high uncertainty in classification accuracy.

Interesting results are found in the random shuffling sequence training data samples because the SSOM with different training data sequences produces variation in the classification accuracy. It implies that the SSOM is sensitive to the sample sequence although the same training data is used for SSOM.

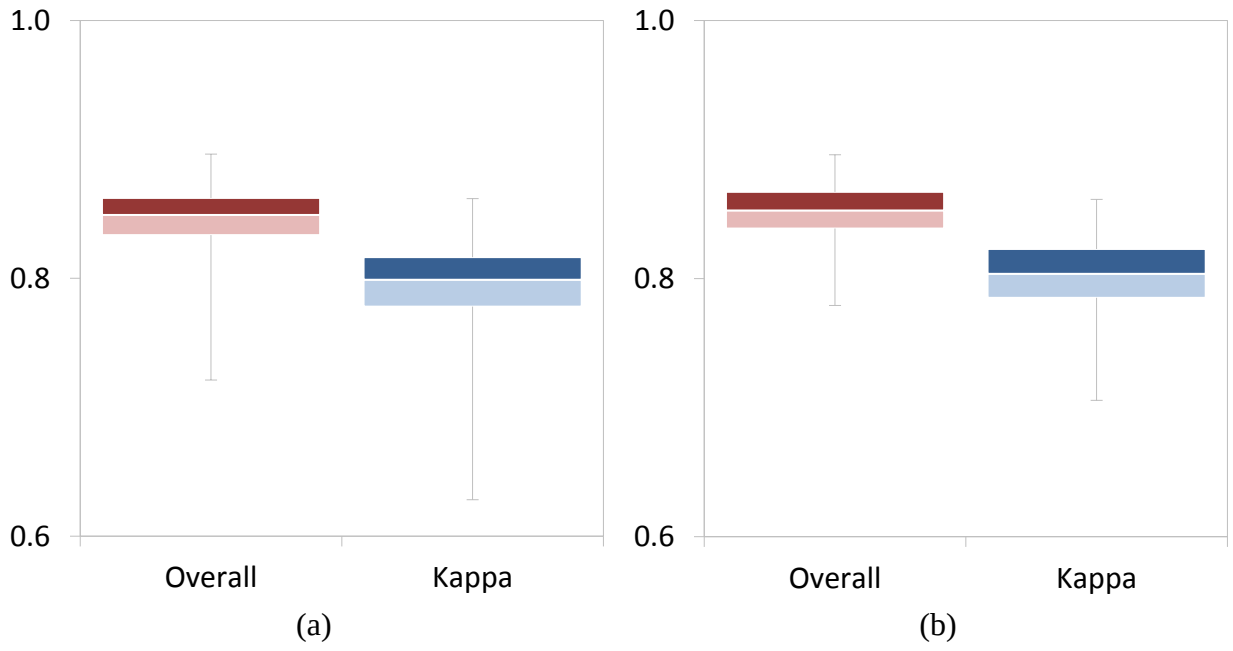
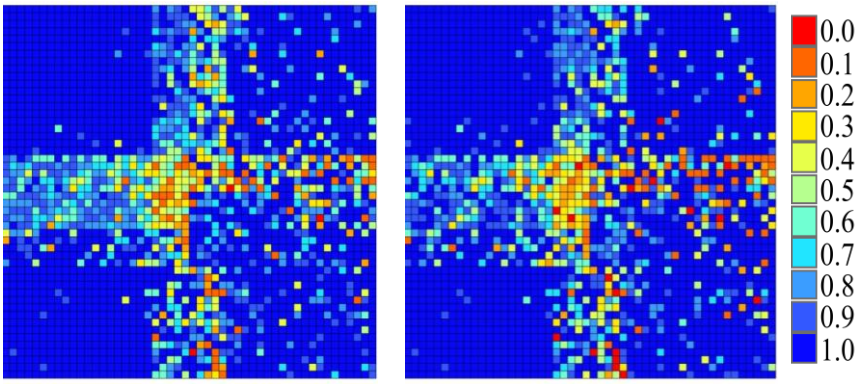


Figure 4.23. Distribution of classification accuracy of SSOM in different (a) random selecting training data (b) shuffling sequence training data.



(a) (b)

Figure 4.24. Images of accuracy possibility derived from SSOM in different (a) random selecting training data (b) shuffling sequence training data.

4.6.3 Uncertainty associated with classifier

This section attempts to examine how the classifier itself affects the classification accuracy. The accuracy of the SSOM is determined by the neural network architecture and internal parameters. Different neural network design and settings can lead to uncertainty in the classification accuracy. In this study, several parameters involved in training the SSOM including the number of competitive layer neurons (NET), the initial weights (W), the number of iterations (ITER), and the initial learning rate (LR), are selected to evaluate the sensitivity of the classification accuracy derived from the SSOM.

Figure 4.25 shows the process of evaluating the uncertainty in the classification accuracy associated with a particular classifier. In this research, 300 different neural network configurations are conducted by combining different values of each parameter. The classification is repeated 300 times using the varied configurations, while the input and training data are kept constant. To assess the classification accuracy, the output images are compared with the reference data. Then, the accuracy distributions and accuracy possibility images are generated.

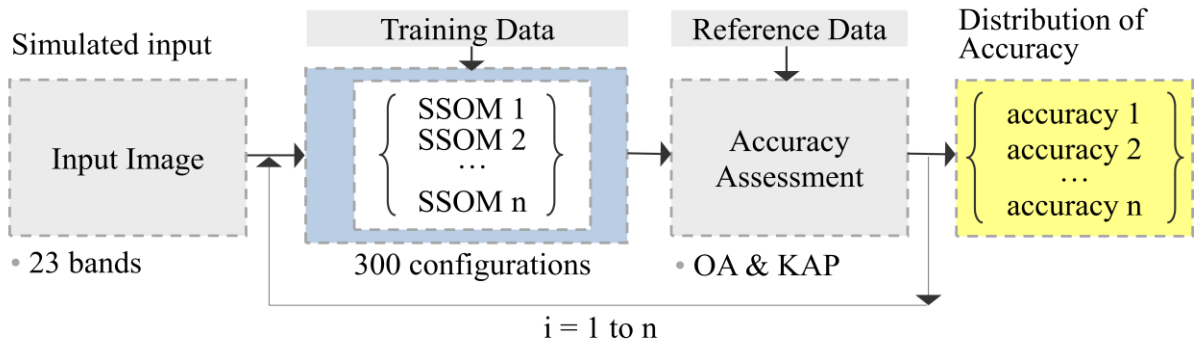


Figure 4.25. Experimental procedure to evaluate the classification uncertainty associated with the classifier.

1) Number of competitive layer neurons (NET)

In this study, NETs of 2 × 2, 4 × 4, 6 × 6, 10 × 10, 15 × 15 and 20 × 20 are applied to evaluate the uncertainty in the classification accuracy. Each NET is fixed, while the SSOM is trained by various initial learning rates (0.001 to 0.9) and numbers of iterations (50 to 1000). Fifty simulations total of each NET are conducted using the same input and training data.

Figure 4.26 shows that both the OA and KAP increase when the size of NET increase and then decrease after reaching their maximum at NET of 6 × 6 (36 neurons). A NET of 2 × 2 provides the highest variation and the lowest median classification accuracy, while a NET of 4 × 4 provides the lowest classification accuracy variation. The highest median classification accuracy is obtained with a NET of 6 × 6. These results indicate that the SSOM with extremely am small NET size is not effective because it will lead to a low accuracy possibility, particularly in areas of mixed pixel as shown in Figure 4.27.

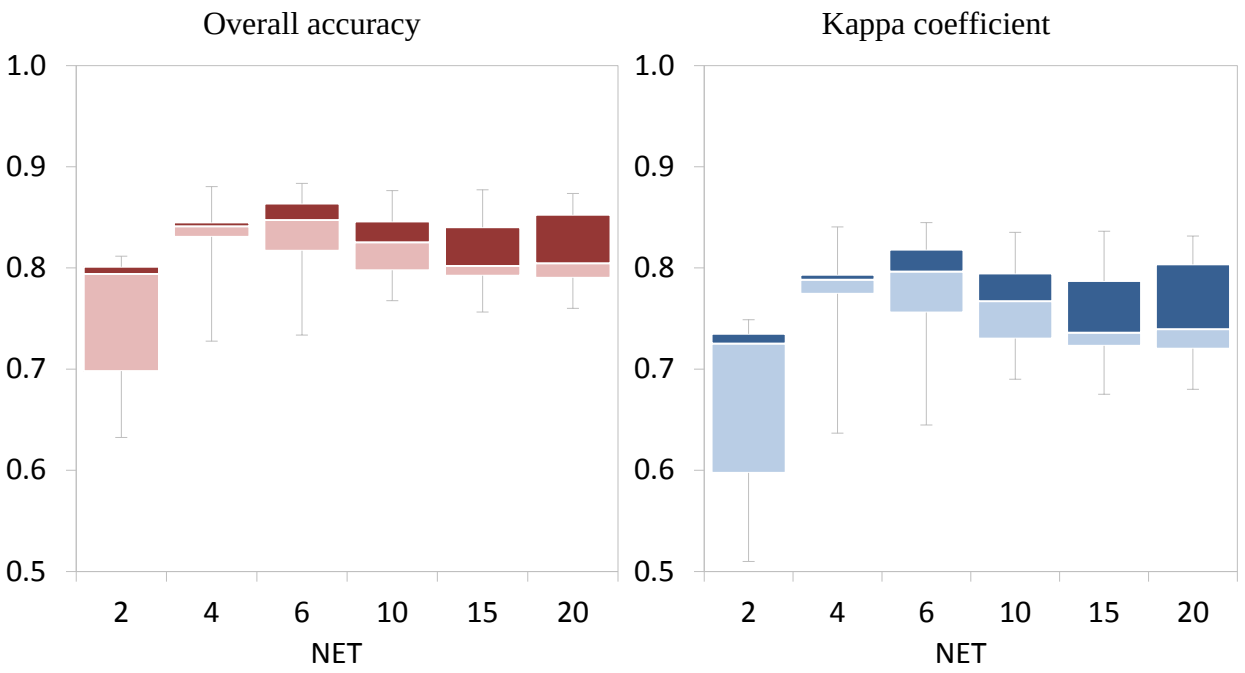


Figure 4.26. Distribution of classification accuracy of SSOM in different NET.

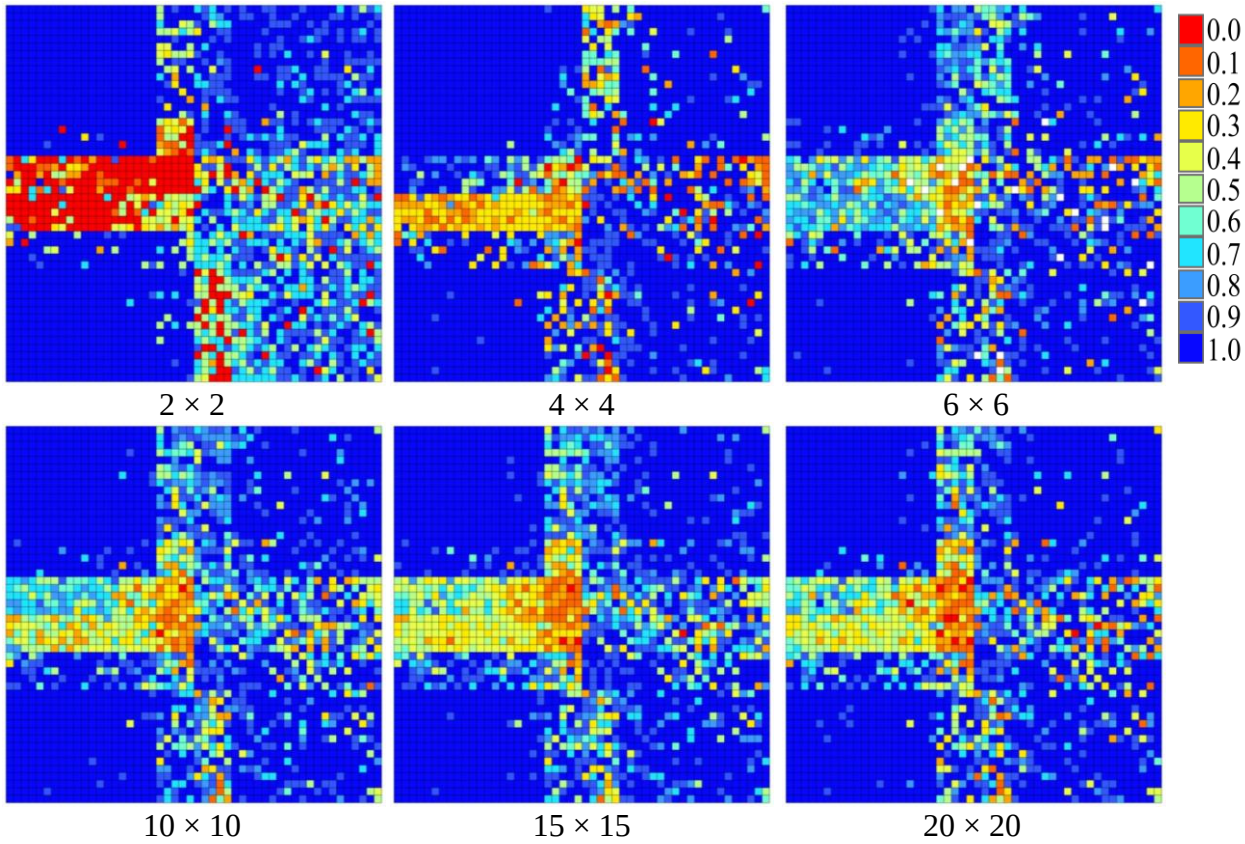


Figure 4.27. Images of accuracy possibility derived from SSOM in different NET.

Moreover, this study also found that the minimum NET corresponds to the number of land cover diversities. In a small NET, there is not enough space to cluster land cover classes; on the other hand, an extremely large NET can minimize the capability of the SSOM to generate unknown patterns. Similar findings in the accuracy possibility images also show that an unsuitable NET selection can drop the efficiency of the SSOM resulting in increased uncertainty in classification accuracy.

2) Initial weight (W)

Since the initial weight (W) of the SSOM will affect the degree of convergence, the classification accuracy uncertainty associated with W is evaluated in this study. The SSOM is trained by varied initial weights (0-1), whereas all other parameters (50 iterations, a NET of 6 × 6, a learning rate of 0.1, and training data) are fixed.

After running 500 classifications of different random initial weights, the SSOM can maintain constant accuracy. The accuracy distribution clearly shows a standard deviation of zero, in other words, initial weight does not have an influential effect on the SSOM classification accuracy. This implies that the SSOM is robust under the condition of varied initial weight. However, setting an appropriate W close to the center of each class will heighten the speed of convergence in learning process.

3) Number of iteration (ITER)

This study also examines how the number of ITER in the training process affects the uncertainty in classification accuracy of the SSOM. Five ITERs consisting of 50, 100, 200, 500,

and 1000 are employed. Classification is then performed by changing other training parameters and run for 60 simulations of each ITER.

The results illustrate that there is not an extensive change in both distributions of OA and KAP when increasing the level of ITER (Figure 4.28). This may suggest that changing the ITER has only a slight impact on the uncertainty in the SSOM classification accuracy. Images of accuracy possibility for different ITER also confirm that the SSOM is only slightly affected by changing different ITER settings as shown in Figure 4.29. Additionally, the results show that increasing ITER does not improve the classification accuracy; conversely, it extensively increases computational time in the learning process.

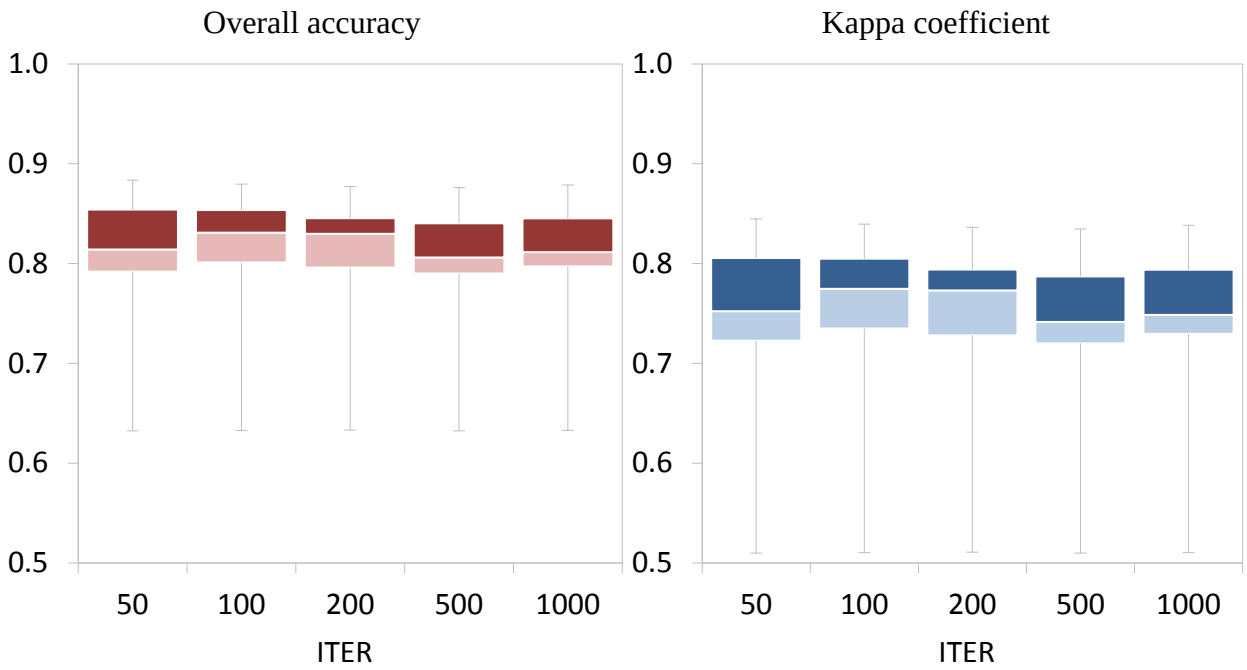


Figure 4.28. Distribution of classification accuracy of SSOM in different ITER.

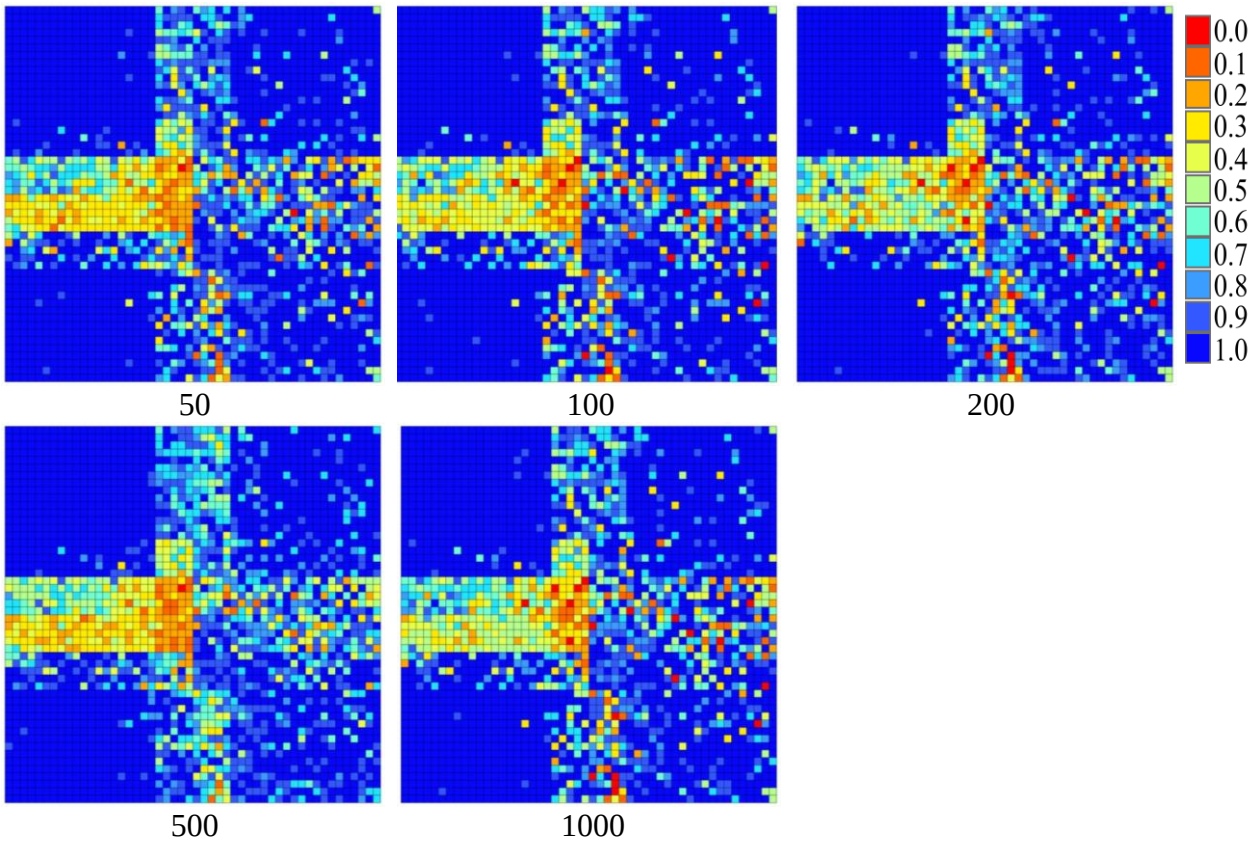


Figure 4.29. Images of accuracy possibility derived from SSOM in different ITER.

4) Initial learning rate (LR)

Another goal of this study is to investigate how each LR impacts the uncertainty of the SSOM classification accuracy. The value of the learning rate ranges between 0 and 1. Generally, the learning rate starts with a comparatively large value (close to unity), then gradually declines to a small value, then a learning rate close to zero is selected for fine adjustment in the final training cycles. LRs of 0.001, 0.005, 0.01, 0.025, 0.05, 0.075, 0.1, 0.25, 0.5, and 0.9 are employed in this study. Each LR is fixed and utilized in 30 simulations where the other training parameter values are varied.

The results of the classification accuracy uncertainty are quantified at different LRs (Figure 4.30 and 4.31). Interestingly, the results show that both the OA and KAP of the SSOM with LRs from 0.001 to 0.05 are higher than with LRs from 0.075 to 0.9.

According to this study, the SSOM provides higher accuracy using low LR values. The results show that an LR of 0.001 provides the lowest median classification accuracy. In contrast, increasing LRs are negatively related to the variation of classification accuracy. The results indicate that a larger LR value tends to produce low uncertainty in the SSOM classification.

It should be noted that the SSOM with different LR values can provide either high or low accuracy and uncertainty depending on the diversity of the study area and the complexity of the input data.

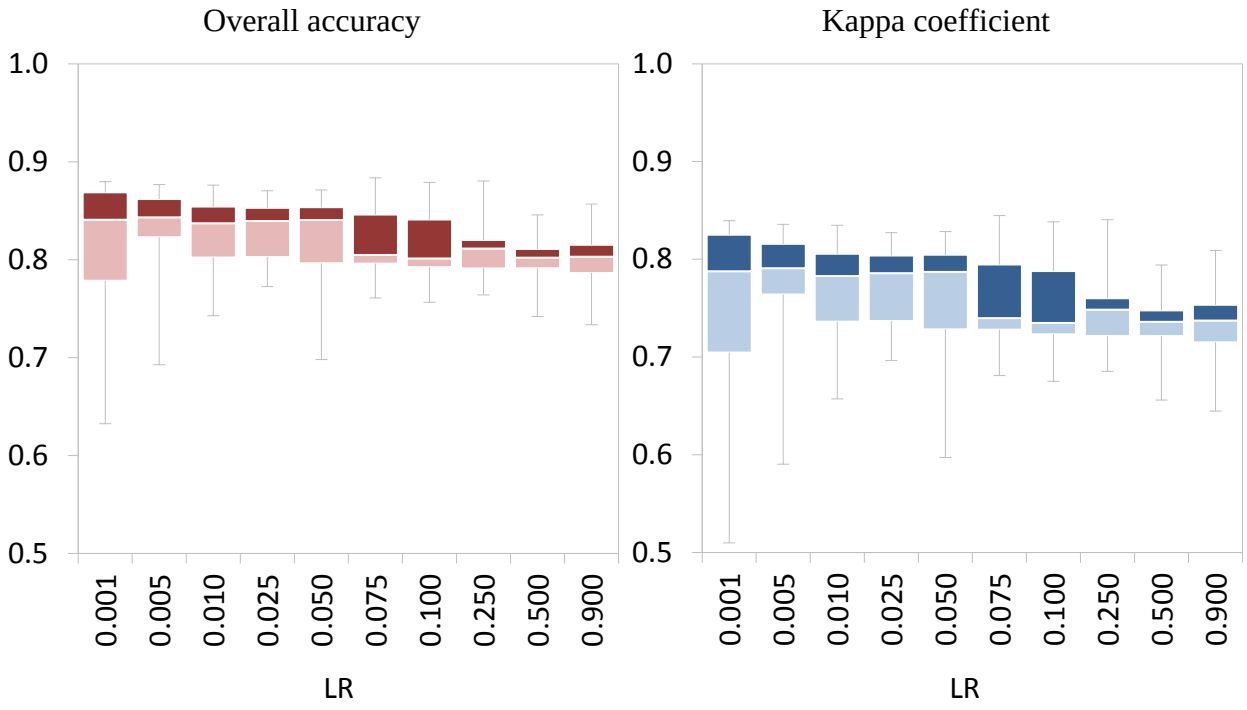


Figure 4.30. Distribution of classification accuracy of SSOM in different LR.

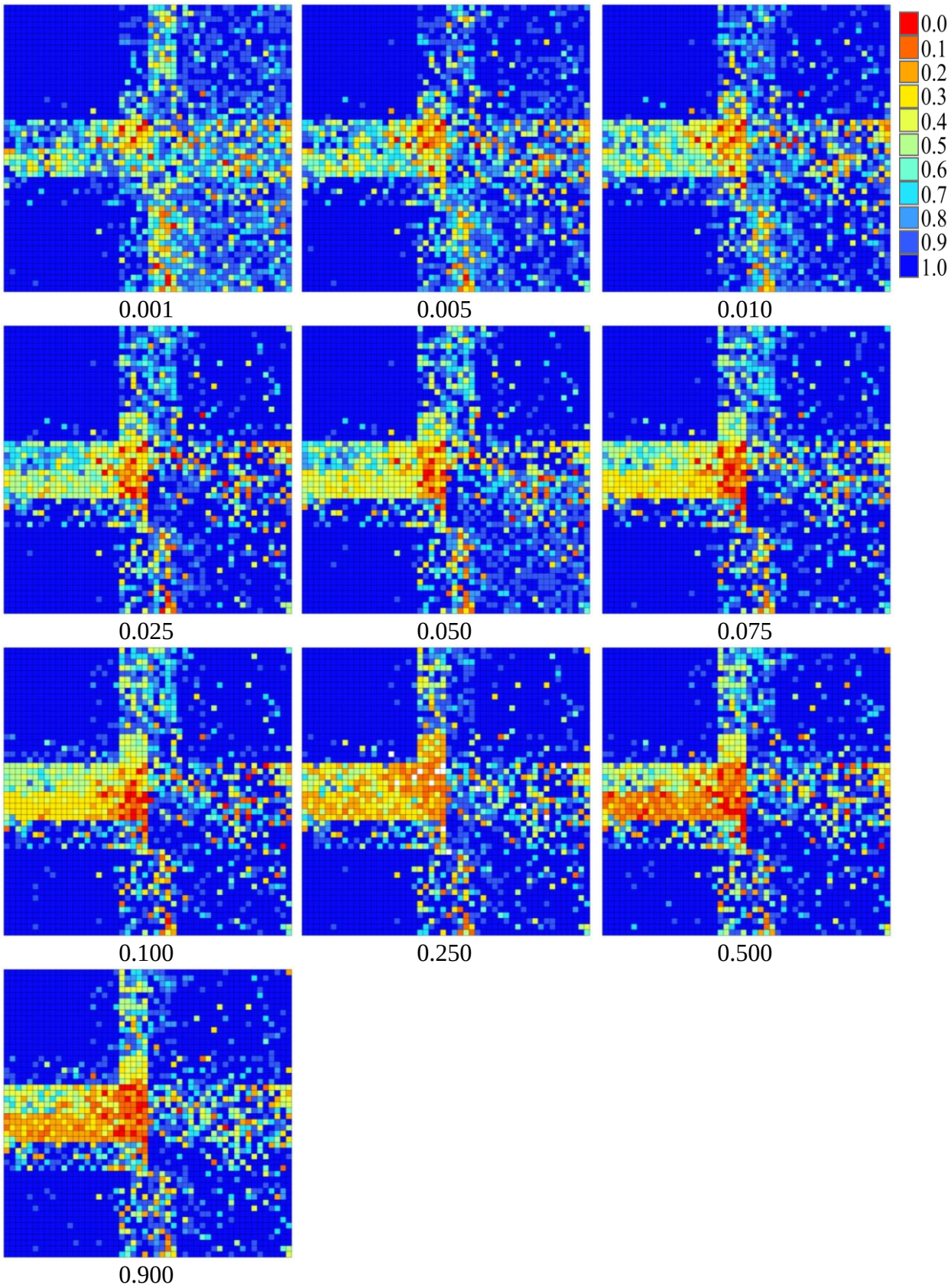


Figure 4.31. Images of accuracy possibility derived from SSOM in different LR.

4.7 Conclusion and discussion

This chapter demonstrates the suitable classification method using synthetic data. This data is generated from a remotely sensed time-series image of MODIS-EVI using a linear mixture model for realization. The synthetic data is utilized to develop and test new classification algorithms.

Appropriate input data are important for the classification process. The results in the first experiment indicate that TIME is a useful input dataset for land cover classification although this data has a high-dimensional space and may lead to the curse of dimensionality. However, the robustness of the SSOM handles the deficiencies of TIME so that it can be a suitable input data for classification. The SSOM can map high dimensional inputs onto low dimensional units and preserve the topology of input patterns in the low dimension after dimension reduction (Bagan *et al*., 2005). In contrast, PHEN extraction provides low dimensional space of the input data but leads to information loss in image classification. Therefore, the SSOM is an extremely useful classifier to apply to large datasets and can produce high classification accuracy.

In addition to input data, issues concerning the network structure and parameters need to be addressed for the neural network-based classification because the speed and effectiveness of the learning process are significant components of successful neural network-based classifications. The results indicate that a better classification performance could be produced by making certain adjustments to the network structure and the parameters used in the neural network classification. The results of the second experiment suggest that the accuracy of classification varies across different neural network models associated with different internal parameter settings. The network size is one example of this issue. Increasing the network size does not improve classification accuracy; on the other hand, it leads to intensive computational

time. Therefore, classification accuracy significantly varies according to the selection of the network structure and parameter values. However, the network structure and the parameters used in this study are limited and only appropriate for similar problems as well as a small number of output classes.

In the third experiment, the performance of three hard classifiers (GMLC, BPNN, and SSOM) is evaluated. The results show that the SSOM with the time-series image significantly outperforms the conventional classification techniques: GMLC and BPNN. The SSOM is able to learn from input data and thereafter be able to generalize and predict unknown patterns based on the data source. The SSOM architecture is composed of the number of neurons that can be chosen accordingly to the complexity of the problem. As the number of complex trends or groups increases, the number of neurons required increases as well. More importantly, the SSOM provides topology-preserving mapping from a high-dimensional input space onto a low-dimensional map space. Also, it can eliminate the problem of a local optimum in the learning process, which is a major problem in traditional neural network-based classifications. With these capabilities, the SSOM in this experiment shows accurate and stable results with small variation of accuracy when comparing to the GMLC and BPNN.

However, hard classification using the SSOM is less successful than soft classification. The fourth experiment shows that the 3SOM, which is operated as a soft classifier, generates more accurate classifications than the SSOM, which is utilized for hard classification. Hard classification is performed by assigning each pixel to a single class. In reality, many pixels in an image may represent more than one land cover class on the ground. To allocate a mixed pixel to a single land cover class not only provides an unrealistic result, but also leads to an inaccurate representation of land cover (Thornton *et al.*, 2006); therefore, it leads to an inaccurate

classification. This problem is heightened in the areas of mixed pixels because the image may contain more than one land cover class, particularly in coarse satellite images such as MODIS. The proportion of mixed pixels generally increases with a coarsening of the spatial resolution of the sensing system (Foody, 1996b). The 3SOM is able to address these problems. With the basic principle of soft classification, the strength of class membership derived in the classification should be related to its land cover composition (Foody, 1996b). Soft classification can allocate a mixed pixel by decomposing a collection of class component spectra or endmembers into a collection of corresponding fractions or abundances. Therefore, the 3SOM with soft classification provides more realistic land cover and a more accurate representation than the SSOM with hard classification.

Not only does soft classification perform in classification stage, but class mixing in the ground reference data used in the training and testing stages is also very important. The fifth experiment demonstrates that the 3SOM-F, which applies both pure and mixed pixels for the training process, produces more accurate and realistic classification results than the 3SOM-P, which utilizes only pure pixels for the training process. This indicates that training with pure and mixed pixels can improve the performance of classifiers. It should be noted that there is a minimum difference in classification accuracy between the 3SOM-F and 3SOM-P in the areas of homogeneity. However, the classification accuracy of the 3SOM-P is less than that of the 3SOM-F in the areas of heterogeneity, where there are both more classes and similar temporal profiles among classes.

Interestingly, the classification results from these experiments show that class A yields the highest classification accuracy, while there is some confusion between class B and C. This is because the classification performance was affected by the variation of the standard EVI

temporal profiles. Figure 3.2 shows that the EVI temporal profile of class A is distinctly different from other classes resulting in the highest accuracy. Similar EVI temporal profiles of class B and C are identified as spectral confusion, which leads to difficulty in classification and consequently, results in lower accuracy compared to class A.

There are a few important notes regarding the accuracy assessment of soft classification. Although soft classification appears to provide a more accurate result, a major limitation to its use and interpretation is the evaluation of the accuracy of the land cover representation (Goodchild, 1994 as cited in Foody, 1996a). In this study, the accuracy assessment shows that AEP does not take the spatial distribution of omission and commission errors into account. Although the AEP results show high accuracy, they do not indicate high proportion correspondence between the output and reference images. Appropriate accuracy evaluation of soft classification should be brought into consideration in order to allow the comparison of different classifications and assess their advantages and disadvantages.

It is undeniable that the reliability of the classification output is an important subject for image classification. Consequently, the last experiment is conducted to investigate how uncertainty in the input data, the training data, and classifier affects the classification accuracy of the SSOM. For uncertainty associated with the input data, increasing the levels of noise has an extensive influence on the classification accuracy, particularly in areas of mixed pixels, but has a marginal effect on the uncertainty in the classification accuracy. This indicates that the SSOM is a stable and robust classifier providing precise accuracy. For uncertainty associated with the training data, randomly selecting training datasets has a small impact on the uncertainty in classification accuracy due to the use of synthetic data and the procedure of selecting training samples. Interestingly, although the same training data is applied, the SSOM with different

sequences of training data still produces a variation in classification accuracy. For uncertainty associated with the classifier, the SSOM shows its effectiveness with a NET of 6 × 6. The small NET size leads to a low accuracy possibility. Conversely, a large NET size does not improve classification accuracy; in addition, it can lower the performance of the SSOM. Therefore, the minimum NET should correspond to the number of land cover diversities. Initial weight does not apparently have an influence on the classification accuracy, whereas ITER has a minimal effect on the uncertainty in classification accuracy of the SSOM. The results of this study also indicate that SSOM provides high classification accuracy at low values of LR, but a large value of LR tends to provide low uncertainty in classification accuracy. An appropriate LR value depends on the study area diversity and the complexity of the input data. Most importantly, the SSOM is likely to produce low accuracy and high uncertainty in areas of heterogeneity and large diversity. These results enhance the conceptual understanding of the uncertainty in classification accuracy associated with the input data, training data, and classifier. Moreover, results in this study can also be a guideline for a appropriate configuration of the SSOM to improve classification results.

Therefore, these results affirm that the 3SOM-F is a potential method for land cover classification with time-series imagery. The effectiveness of the 3SOM-F depends on selecting the suitable neural network configuration. Furthermore, uncertainty can have a significant influence on the reliability of the 3SOM-F output.

Chapter 5

Applying Identified Method Using Real Landscape Dataset

5.1 Introduction

In chapter 4, the process of developing a suitable classification method using synthetic data is described. This process involved conducting the following six experiments : 1) comparison of input images between time-series and phenology images by using the self-organizing map classifier (SSOM), 2) selection of the suitable neural network configuration of the back-propagation neural network classifier (BPNN) and SSOM, 3) comparative evaluation of the SSOM with the Gaussian maximum likelihood classifier (GMLC) and BPNN, 4) comparative evaluation of the SSOM with the soft-supervised self-organizing map classifier (3SOM), 5) comparative evaluation of the fully-3SOM with the partially-3SOM, and 6) accessing the uncertainty in classification accuracy of SSOM.

The results exhibit that time-series imagery is a potentially useful input dataset for land cover classification. Moreover, the SSOM with time-series data significantly outperforms the conventional classification techniques of the GMLC and BPNN. Additionally, the 3SOM, which is employed as a soft classification, generates a more accurate classification than SSOM, which is applied as a hard classification. Furthermore, the 3SOM-F, which is applied using pure and mixed pixels for the training process, accomplishes classification results more accurate and realistic than the 3SOM-P, which utilizes only pure pixels in the training stage. Therefore, these results suggest that the 3SOM-F is an appropriate method for land cover classification with time-series imagery. However, there is uncertainty in the classification accuracy associated with

network architecture design and internal parameters settings. As a result, the suitable neural network configuration should be investigated for the best performance of the classifier.

Although 3SOM-F is proven to be an appropriate method, it is very important to test its capability with real landscapes because the diversity of the landscape has a considerable impact on the performance and ability of the classifier. In order to demonstrate the effects of landscapes on the proposed method, this chapter applies the 3SOM-F to real landscape datasets derived from MODIS time-series images. Two study areas in Thailand and the Midwest region of the U.S. are selected based on differences in land cover characteristics. The agricultural areas in Thailand are small in size and tend to be more diverse, which cause mixed pixels containing more land cover classes in a given pixel and greater confusion between classes. The agricultural areas in the Midwestern U.S. are less diverse than those in Thailand because there are only two major crops: corn and soybeans; however, both crops have very similar EVI temporal profiles leading to difficulties in classification.

5.2 Description of MODIS-EVI time-series dataset

The MODIS-EVI 16-day 250 m (MOD13Q1) dataset for 2010 was acquired for this study. Each pixel contains an EVI (Enhanced Vegetation Index) value. MOD13Q1 is designed to provide consistent spatial and temporal comparisons of vegetation conditions. Blue, red, and near-infrared reflectance are used to determine the MODIS daily vegetation indices. The MODIS-EVI minimizes canopy background variations and maintains sensitivity over dense vegetation conditions.

In this study, the MODIS-EVI time-series in 23 composite images is utilized. The products are downloaded from the USGS Land Processes Distributed Active Archive Center (LP

DAAC, 2008). These EVI products have a 250-m resolution in a level 3, grid projection. Each pixel contains the best possible daily observation during a 16-day period. These version-5 EVI products are validated stage 1, which means that accuracy has been estimated using a small number of independent measurements obtained from selected locations and time periods and a ground-truth/field program. Prior to the analysis, these data are converted to GeoTIFF format and reprojected to a Geographic coordinate system with the WGS1984 (the World Geodetic Survey System of 1984) datum using the MODIS Reprojection Tool (MRT) from USGS/LPDAAC (LP DAAC, 2008). The Savitzky-Golay filter is then applied, which uses a moving 5-point window for each pixel in the time-series profile, and in each window, noise values are approximated by a polynomial function to smooth EVI values in the window. Then, these reprojected and smoothed time-series images are used as the input dataset in the 3SOM-F classification.

In this research, two real-landscape datasets are used to investigate the ability of the 3SOM-F classification. Two study areas provide different spatial and temporal mixture problems. To simplify the analysis, only 12.5 × 12.5 sq.km (50 × 50 pixels) subsets of these regions are selected. The characteristics of these two study areas are illustrated as follows:

5.2.1 Characteristic of Thailand dataset

The first dataset is focused on major crops grown in the Lopburi Province, central Thailand. This study area is comprised mainly of four land cover classes: forest, sugar cane, cassava, and paddy rice. All other classes (e.g., buildings, asphalt, and water) are excluded from the analysis. The vegetation development in this area is affected by a great heterogeneity that has consequences for the mixture of land cover types. Moreover, the phenology is very dependent on

agricultural practices, which may greatly vary among farmers. The variations of soil property, water supply, and fertilizer management are the major factors of intra-class variability.

Figure 5.1.a shows the distributions of the EVI temporal profiles of each land cover class, which are extracted from the MODIS-EVI time-series imagery in this study area. The median EVI temporal profiles are represented by the center lines. The shaded regions represent the minimum, lower quartile, upper quartile, and maximum of distribution around the median. The timing and value of the peak EVI diverge among each land cover class. This variability reflects the regional variations in environmental conditions and crop management practices. High intra-class variability can increase the overlap in the EVI temporal profiles among the crops and reduce their separability.

Figure 5.1.b shows the median EVI temporal profiles corresponding to the four crops of interest. They correctly represent the agricultural calendar. Forest tends to have the highest EVI profile because of high density of green cover, while lowest EVI profile is obtained from paddy rice. However, the start of the growing season for paddy rice shifts toward the later dates when comparing to other crops. During the green-up phase, forest and paddy rice have distinctly different profiles while sugar cane and cassava have very similar profiles. The profiles of forest, cassava, and paddy rice are similar during the beginning of the senescence phase but the profiles of sugar cane and cassava are different in the general timing of senescence.

5.2.2 Characteristic of the Midwestern U.S. dataset

The second dataset is a part of the state of Iowa in the U.S. Corn Belt. This area consists of predominantly corn and soybeans andis one of the most important corn and soybean production areas in the U.S. Iowa is located in the Midwest and has very little terrain variability as it is mainly dominated by plains. According to Doraiswamy (2007), this state is intensively

cultivated, with approximately 75% of the land devoted to corn and soybean farming. These crops are grown under rain-fed conditions where soil moisture is normally adequate in the growing season; however, moisture stress conditions can occur in the early stages of crop development and more often during the latter part of the season. Seasonal rainfall ranges between 800 mm in the north to 450 mm in the southern part of the region. Soil moisture is generally at field capacity at the start of the season. However, because of spatial variability in the spring rainfall, planting dates across the region are variable. Crop planting is completed by mid-May, with corn generally planted about 2 weeks earlier than soybeans. Crop maturity is occurred by late September.

Corn and soybeans are categorized as '*summer crops*' because most of their growth cycle occurs during the summer. Although these summer crops have similar crop calendars, unique spectral–temporal responses that represent subtle differences in their growth cycles are reflected in their respective EVI temporal profiles.

Figure 5.2.a shows the distributions of EVI temporal profiles for corn and soybeans. The center lines are the median EVI temporal profiles, while the distributions (minimum, lower quartile, upper quartile, and maximum) are illustrated by the shaded regions around the median line. Corn and soybeans have low EVI profile variation compared to the Thailand dataset because the timing and value of the peak EVI are quite consistent over the growing season of both corn and soybeans.

Figure 5.2.b is the median EVI temporal profiles showing differences between corn and soybeans. The corn and soybean profiles exhibit a rapid increase then decrease over the growing season. The start of the growing season for the corn profile begins in April and continues until mid-May. Then, the large decrease occurs starting mid- to late August. The soybean crop is

planted several weeks after corn and the maturity follows that of corn. Thus, the profiles of soybeans are slightly shifted toward the later dates when comparing to the corn profiles.

5.3 Derivation of proportional reference image

To evaluate the accuracy of soft classification, the reference data of land cover proportions is created and identified. However, the derivation of the land cover proportions is not straightforward. Land cover data such as those extracted from higher resolution images are conventionally used and represented in the form of discrete polygons or grid cells of classes. In order to generate the reference data of land cover proportions for this research, the land cover data of both study areas (Thailand and the Midwestern U.S.), which are extracted from higher resolution imagery with ground truth data, are used to derive sub-pixel land cover proportions related to the MODIS image (250m). The details of the procedure to derive the reference data for each study area are described below.

5.3.1 Thailand

National Land Use Dataset of Thailand from the Land Development Department, Ministry of Agriculture, Thailand is used as reference land cover data. This data is generated from digital color aerial images from 2004 at a scale of 1: 4,000, SPOT-5 images with a spatial resolution of 5 m from 2007, and THEOS images with 2 m spatial resolution from 2010. The digital color aerial images were geo-referenced and ortho-rectified using field-collected ground control points (GCPs) and fine-resolution digital elevation models (DEM). Visual interpretation of digital color aerial images, which is updated by SPOT-5 and THEOS images with field support data, was conducted to generate land cover data in vector format at a scale of 1:50,000. This data is rasterized to a fine grid of pixels 25 m in size (Figure 5.3.a), then the grid data is

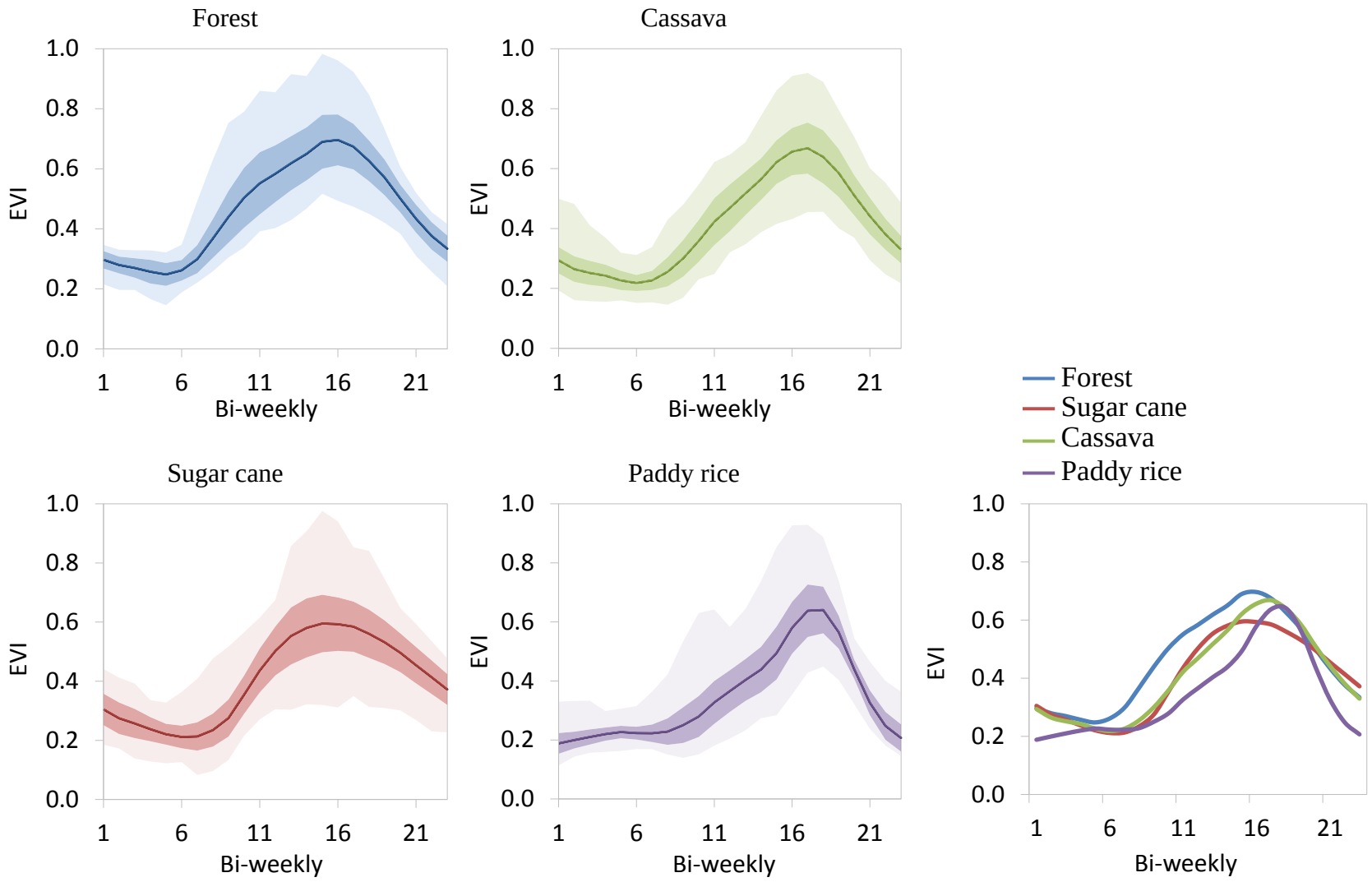
aggregated to correspond with the MODIS pixel size (100 pixels of fine grid = 1 pixel of MODIS image).

Proportions of land cover for each class in a pixel are calculated on a pixel-by-pixel basis which includes the classes of forest, sugar cane, cassava, and paddy rice. For example, for 100 pixels of the land cover image that are spatially associated with one pixel of MODIS image, if 70, 25 and 5 pixels belong to forest, sugar cane and paddy rice, respectively, this proportion reference pixel contains 70%, 25% and 5% of forest, sugar cane and paddy rice, respectively. To facilitate subsequent analysis, the proportional reference data derived from proportions of the sub-pixel component land cover are also stored as a four-layer image, one for each class (Figure 5.3.b).

5.3.2 The Midwestern U.S.

The U.S. Cropland Data Layer (CDL) derived from the LANDSAT 5-TM is utilized to evaluate the classification accuracy in this study area. This data is published by the National Agricultural Statistics Service (NASS), which is part of The United States Department of Agriculture (USDA). The CDL is produced in raster format with a 30 m spatial resolution. This data is then resampled to the resolution of the fine grid data at 25 m (Figure 5.4.a).

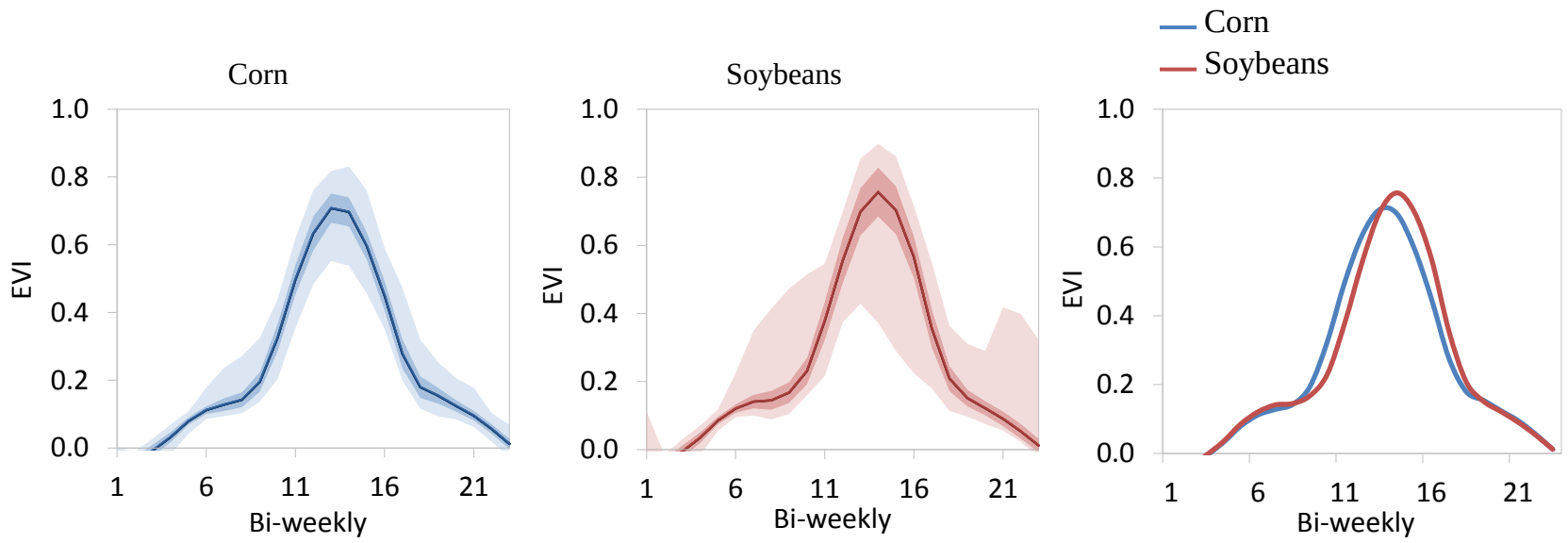
Proportions of land cover are calculated the same way as discussed in 5.3.1. The proportions of two land cover classes (corn and soybeans) are calculated on a pixel-by-pixel basis and correspond with the MODIS pixel size. Sub-pixel component land cover proportions of this reference data are extracted for the proportional image in two layers, one for each class (Figure 5.4.b).



(a) Distribution of temporal profiles for each class

(b) Median EVI temporal profiles

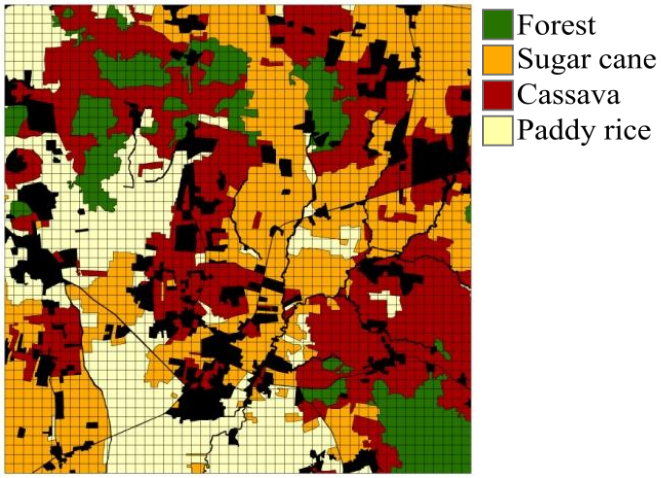
Figure 5.1. The characteristics of EVI time-series images of the study area in Thailand.



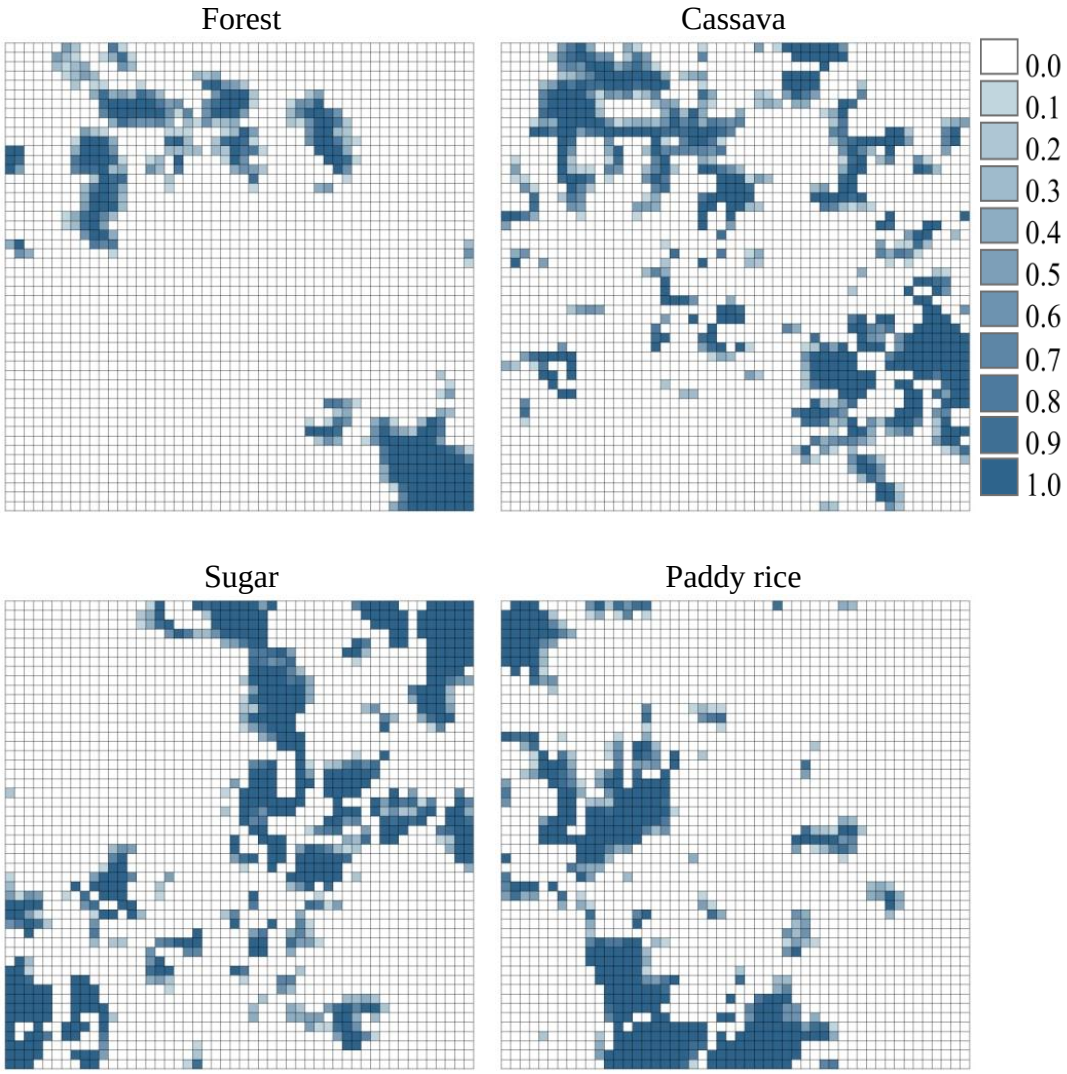
(a) Distribution of temporal profiles for each class

(b) Median EVI temporal profiles

Figure 5.2. The characteristics of EVI time-series images of the study area in the Midwestern U.S.

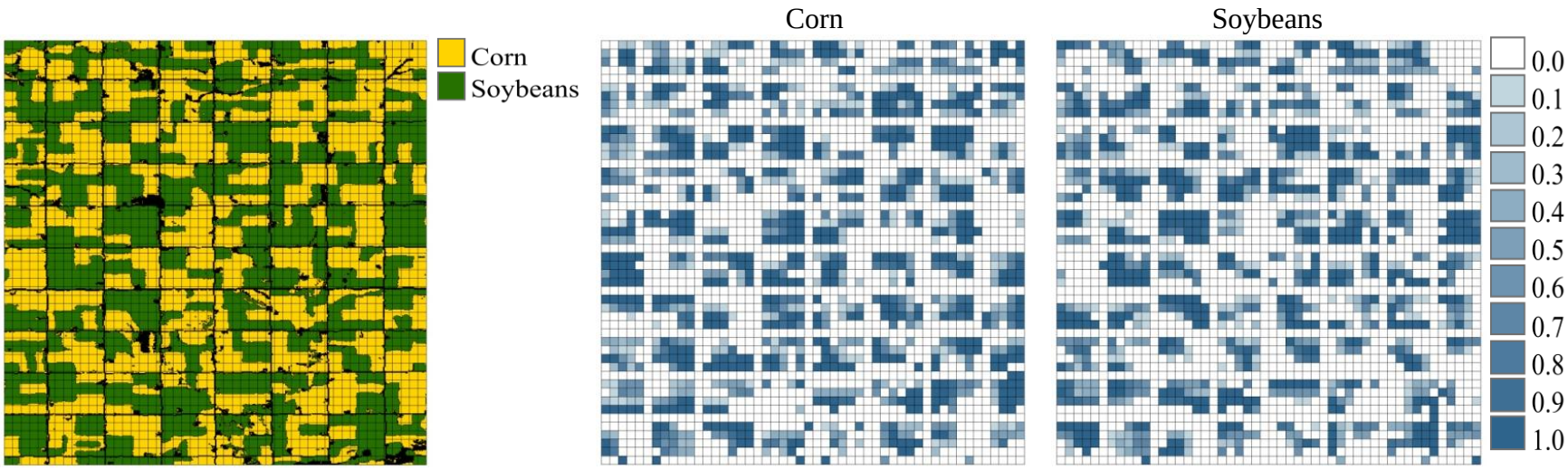


(a) Fine land cover image



(b) Proportional land cover images of each class

Figure 5.3. The reference images of the study area in Thailand.



(a) Fine land cover image

(b) Proportional land cover images of each class

Figure 5.4. The reference images of study area in the Midwestern U.S.

5.4 Classification procedures

The fully soft-supervised self-organizing map (3SOM-F) is designed to be an appropriate classifier for time-series images of both sub-areas in Thailand and the Midwestern U.S. The experimental procedure is shown in Figure 5.5. MODIS-EVI time-series images from 2010 with 50 x 50 pixels are used as the input images. a Savitzky-Golay filtering technique is applied to both time-series images using TIMESAT software (Jonsson & Eklundh, 2004), as described in Chaper 3.2, to remove the spikes and irregular values of original images caused by atmospheric and cloud conditions.

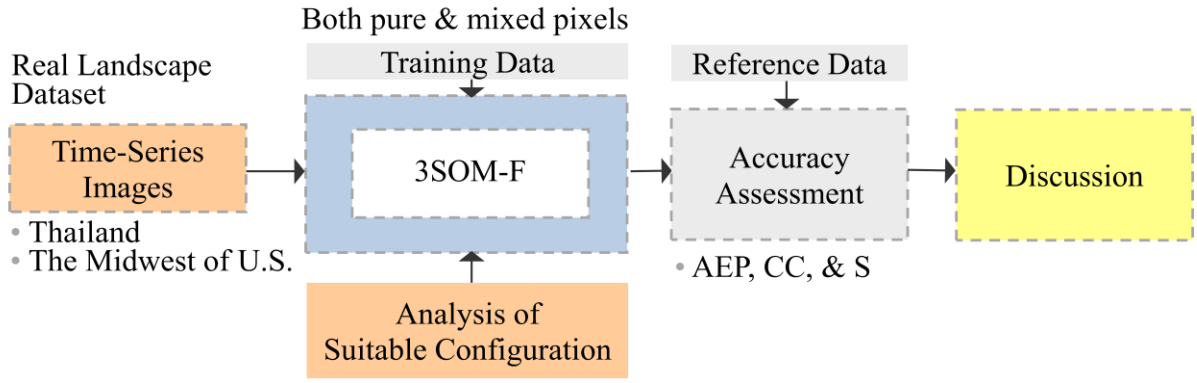


Figure 5.5. Experimental procedure of applying 3SOM-F using real landscape dataset.

In order to determine the suitable neural network configuration, the 3SOM-F with 300 different configurations is executed on the same training and testing datasets under the settings listed in Table 3.5. The number of input and output layer neurons is determined by the number of time-series images and land cover classes, respectively.

All 23 EVI layers (each corresponding to a consecutive two-week period) are utilized for the study area in Thailand since the growing season is year-round. Only 11 layers are employed for the study area in the Midwestern U.S. because the rest of the layers fall into the winter

period, which does not have a crop growing season, thus there is no useful information for land cover classification.

The number of competitive layer neurons and other network properties are selected subjectively on the basis of empirical results from the trial runs. There are six different numbers of competitive layer neurons (NET). For each NET, ten trials with 10 different initial learning rates (LR) are performed and five variant numbers of iterations (ITER) are examined.

All configurations are run on a case-by-case basis with trial-and-error analysis because there is no standard procedure for choosing the optimal configuration. All trials are carried out on the same training and testing datasets. The measure of accuracy used to evaluate the output from different configurations of the soft classification is the mean of closeness (MS). The configuration providing the lowest MS is determined to be the suitable configuration.

In this study, a sample size of both training datasets is 20% of all class samples. The class samples consist of pixels corresponding to the major crops in each study area and other pixels are excluded from classification. There are 1,519 class samples in Thailand, while the Midwestern U.S. has 1,383 class samples. A total of 304 and 277 samples are randomly selected as training datasets for the study areas of Thailand and the U.S., respectively.

Since the 3SOM-F is identified as a fully-soft classification, soft training datasets comprised of both pure and mixed samples are employed. The number of pure and mixed samples is related to the proportions between the total number of pure and mixed samples. Additionally, the number of pure samples for each class is associated with the total number of pure samples in each study area to keep the class distribution similar to that of the full dataset. All training numbers, which are selected from study areas in Thailand and the U.S., are listed in Table 5.1.

In the classification stage, both real landscape datasets are classified using the 3SOM-F with the suitable neural network configurations. These configurations are different between study areas in Thailand and the U.S. Then, the accuracy of land cover classifications of both study areas are evaluated using four measures of soft classification accuracy consisting of area error proportion (AEP), correlation coefficient (CC), root mean square error (RMSE), and mean of closeness (MS).

Table 5.1. Number of class and training samples.

a) Number of class and training samples for Thailand dataset

Sample	**Number of samples**	
	Class sample	**Training sample**
Pure sample	1,088	213
Mixed sample	431	91
Total	1,519	304

b) Number of class and training samples for the Midwestern U.S. dataset.

Sample	**Number of samples**	
	Class sample	**Training sample**
Pure sample	851	166
Mixed sample	532	111
Total	1,383	277

5.5 Results and discussions

Table 5.2 shows the suitable configurations of the 3SOM that provide the highest accuracy. These configurations are used to classify the MODIS-EVI time-series images of both study areas. For the Thailand dataset, the network consists of 23 input layer neurons, 20×20

competitive layer neurons, and four output layer neurons. In the learning process, the parameter defining the initial learning rate is set at 0.5 for 100 iterations. For the U.S. dataset, the network contains 11 neurons in the input layer, 10×10 competitive layer neurons, and two neurons in the output layer. The 3SOM-F is again used, but here the training is constrained to 50 iterations of the algorithm with the parameters defining the initial learning rate set at 0.001.

Table 5.2. The suitable configuration of 3SOM-F for Thailand and the Midwestern U.S. datasets.

Parameters	**Suitable values**	
	Thailand	**Midwest, U.S.**
Number of input layer neuron	23	11
Number of output layer neuron	4	2
Number of competitive layer neuron	20 x 20	10 x 10
Initial learning rate	0.5	0.001
Iterations	100	50

These networks with the suitable neural network configurations are trained to classify the proportional coverage of the classes in the pixels for each study area. Land cover classification results of Thailand and the Midwestern U.S. using the 3SOM-F are described below.

5.5.1 Thailand

Visual inspection of Figure 5.6 demonstrates that the 3SOM-F produces highly accurate and realistically classified proportional images when comparing to the reference images. More importantly, paddy rice clearly appears to be separated from the surrounding classes. However, there are some scattered misclassified pixels distributed among the other classes, particularly in areas between cassava and sugar cane. These erroneous and misleading results are due to some

confusion or mixing between classes because of the similar temporal EVI profiles of these two classes.

The statistics shown in Table 5.3 also confirm the above visual assessments. The table shows the accuracy assessment consisting of the AEP, CC, RMSE, and MS. For all classes, the 3SOM-F generally produces high accuracy for all measures. The AEP of cassava (0.001) is closest to zero which could be interpreted as the highest accuracy for maintaining the class area. In contrast, the area proportion of forest and paddy rice seem to be underestimated with positive AEP values of 0.104 and 0.034, respectively, while a negative AEP value of -0.061 obtained from sugar cane indicates the overestimation of area proportion. Therefore, the 3SOM-F is able to maintain the area of cassava more accurately than the areas of other classes.

Moreover, the 3SOM-F also produces each land cover class with highly satisfactory CC values of 0.679 for forest, 0.705 for sugar cane, 0.607 for cassava, and as high as 0.876 for the paddy rice class. Such results are also apparent in the RMSE. There is only a slight difference of RMSE values for each class, and all classes produce low RMSE values, including 0.235, 0.311, 0.333, and 0.194 for forest, sugar cane, cassava, and paddy rice, respectively.

In addition, Figure 5.7 illustrates that the 3SOM-F produces closeness values near zero in most areas. This indicates that there are little to no differences in class proportions between each pixel of the classified image and the reference images.

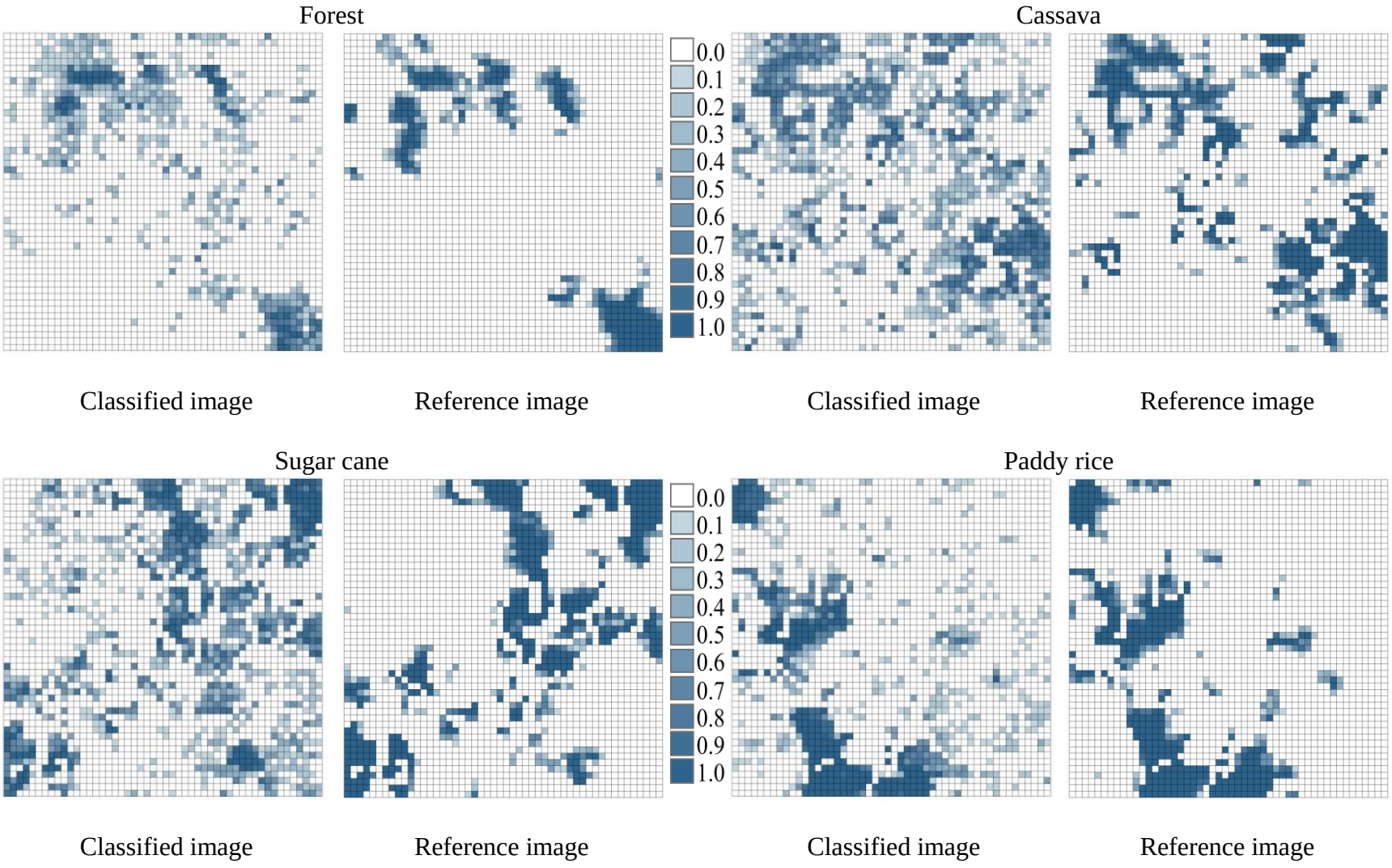


Figure 5.6. The proportional classified images of Thailand using 3SOM-F classification.

Furthermore, the MS of 0.075 for the entire image also signifies that the overall error of the 3SOM-F for classifying proportions of four crops in this area is substantially low. Consequently, the results indicate that the 3SOM-F, which is operated as a fully-soft classifier, successfully classifies land cover in study area of Thailand; especially for the paddy rice class.

Table 5.3. Classification accuracy assessment of study area in Thailand.

	Measures			
	AEP	CC	RMSE	MS
Forest	0.103769	0.679441	0.235212	-
Sugar cane	-0.06106	0.704618	0.310682	-
Cassava	0.001234	0.607234	0.333427	-
Paddy rice	0.033548	0.875902	0.193706	
Entire image	-	-	-	0.075136

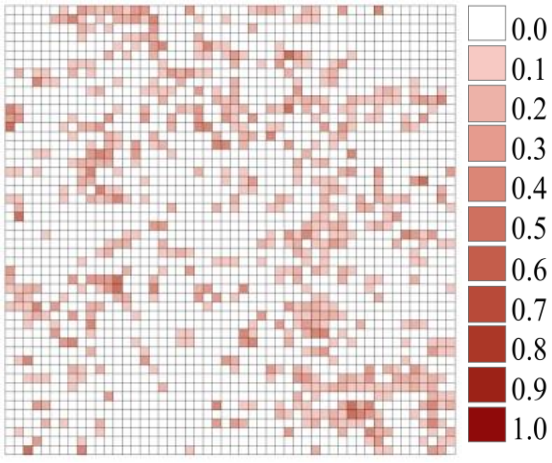


Figure 5.7. Closeness images of the study area in Thailand.

5.5.2 The Midwestern U.S.

The results of fully-soft classification using the 3SOM-F with the suitable neural network configuration are compared with the reference images. A visual assessment of Figure 5.8 illustrates that the 3SOM-F achieves meaningful classification results for corn and soybeans. The visual depiction of the results also demonstrates that both corn and soybeans are uniformly allocated when comparing to the reference images. However, confusion between corn and soybeans is apparent in the areas of pure pixels. The results show that the 3SOM-F underestimates area proportions for each crop. This is because of the similar temporal EVI profile of these two classes.

Table 5.4 shows the statistics of the soft classification accuracy assessment for each class. The results are consistent with the visual interpretation that the 3SOM-F performs well in the classification of corn and soybeans with high accuracy for all measures. However, the 3SOM-F tends to classify corn more accurately than soybeans as seen from its slightly higher classification accuracy in all measures.

The statistics show positive AEP values for corn and soybeans indicating that the area of both classes is underestimated compared to the reference images. Corn has a lower value of AEP (0.022) than soybeans (0.067), suggesting a slightly higher accuracy of corn than soybeans. The results also illustrate that the 3SOM-F classifies each land cover class more precisely with CC values of 0.752 and 0.719 for the corn and soybean classes, respectively. The CC values are only slightly different for both land cover classes. This indicates that the 3SOM-F produces a close correspondence of class proportions between the classified and reference images.

Similar findings are observed in the RMSE values. Corn and soybeans are mapped accurately, with low RMSE values of 0.311 and 0.323, respectively.

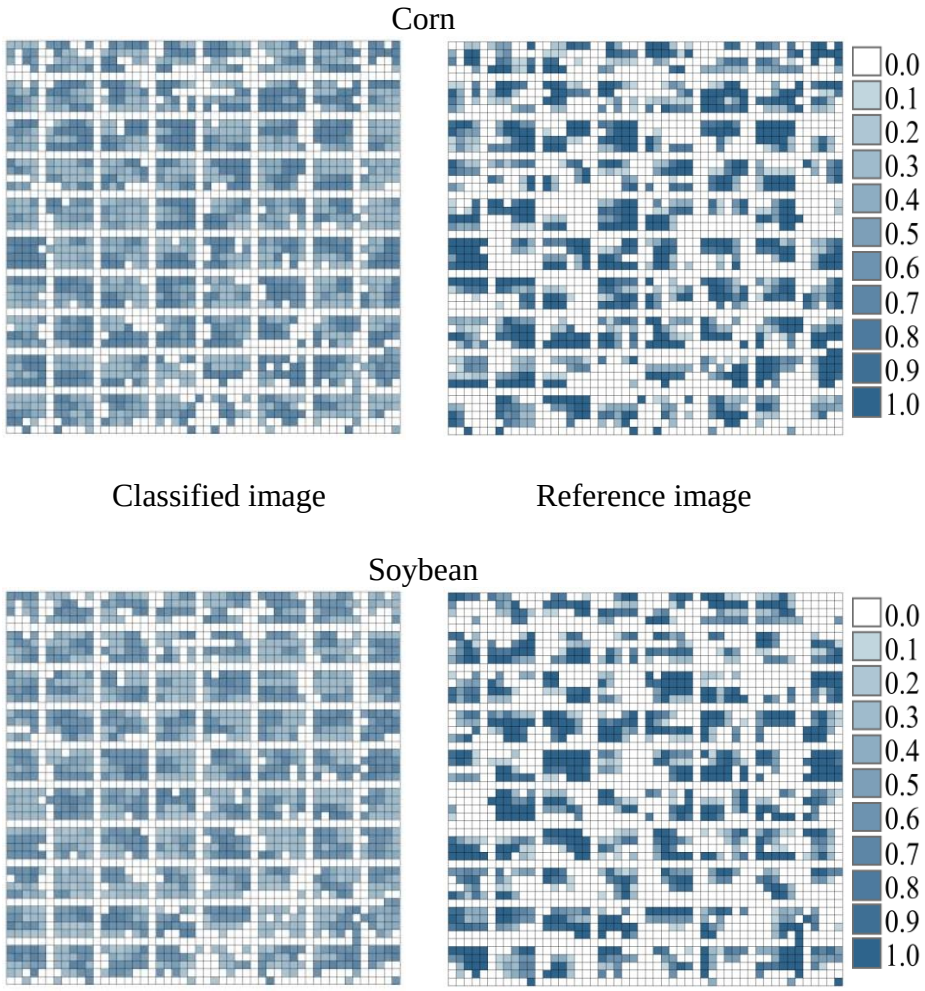


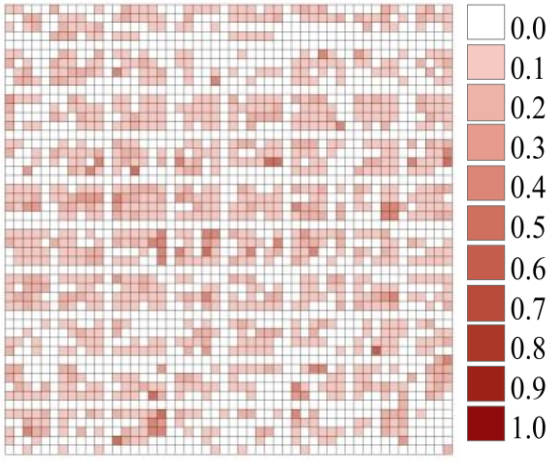
Figure 5.8. The proportional classified images of the Midwestern U.S. using 3SOM-F classification.

Moreover, Figure 5.9 shows the closeness image, which represents the classification error by pixel. Most pixels have a closeness value of less than 0.1 indicating that little to no difference in class proportions between each pixel of the classified and reference images.

Additionally, overall image results show a small MS value of 0.101, which signifies a marginally low overall error of the 3SOM-F for classifying proportions of corn and soybeans in this area. Therefore, the results indicate that 3SOM-F employed as a fully-soft classifier achieves successful land cover classification in study area in the Midwestern U.S.

Table 5.4. Classification accuracy assessment of study area in the Midwestern U.S.

	Measures			
	AEP	CC	RMSE	MS
Corn	0.021788	0.752031	0.311189	-
Soybeans	0.067235	0.718645	0.323053	-
Entire image	-	-	-	0.100601



MS = 0.100601

Figure 5.9. Closeness images of study area in the Midwestern U.S.

5.6 Conclusion and Discussion

In this chapter, the capability of the 3SOM-F is tested with real landscape data because the diversity of real landscapes has a considerable impact on the performance and ability of this classifier. As compared with other methods in the previous chapter, the 3SOM-F is proven to be an appropriate method for image classification based on time-series imagery.

All results in this chapter confirm that the 3SOM-F successfully classifies land cover in both areas of Thailand and the Midwestern U.S. In the Thailand study area, the 3SOM-F performs well in classifying four land cover classes as shown by the high accuracy values for all measures. Paddy rice yields the highest classification accuracy, while there is some confusion between cassava and sugar cane. Some scattered misclassification is found in the areas of heterogeneity or mixed pixels. In the U.S. study area the 3SOM-F produces high accuracy for all measures for both corn and soybeans. However, confusion between corn and soybeans, which is apparent in the areas of homogeneity or pure pixels, causes erroneous and misleading interpretations.

Dissimilarities in the results from two study sites are related to spatial and spectral confusion. The agricultural areas in Thailand are small in size and tend to be highly heterogeneous, resulting in mixed pixels containing more land cover classes in each pixel. This cause of spatial confusion in this area diminishes the classification accuracy. However, spatial confusion is likely to have only a small impact on the performance of the 3SOM-F when the EVI temporal profiles of each class are distinctly different from each other. Similar EVI temporal profiles of corn and soybeans in the Midwestern U.S. are identified as spectral confusion, which leads to difficulty in classification and consequently, results in erroneous and misleading

interpretations. Due to the confusion of similar EVI temporal profiles, the 3SOM-F tends to underestimate area proportions in homogenous or pure pixel areas.

In addition to spatial and spectral confusion, the number of classes is another confusion factor. The greater number of classes present within a pixel, the more errors is found. Therefore, the number of classes in the image affects the classification ability of the 3SOM-F. This problem is more noticeable if the classes have similar EVI temporal profiles. Accordingly, the classification accuracy of the 3SOM-F will decrease as the number of classes increase in a given pixel. In other words, 3SOM-F allocates fewer classes within a pixel more effectively than when there are many classes within a pixel.

Although the 3SOM-F shows acceptable results, its effectiveness depends on the study area, data resolution, and data dimensions. Consequently, when this method is applied to other regions, it is essential to investigate the performance of this method.

Chapter 6

Conclusions and Further Research

6.1 Conclusions

Land cover maps derived from satellite images are widely used as inputs for environmental models and they are also a valuable resource for decision makers in environmental management. Therefore, up-to-date, highly accurate land cover data with current detailed and timely information is required for the global environmental change research community to support natural resource management, environmental protection, and policy making. Remotely sensed image classification has long been a fundamental technique for studying vegetation and land cover (Richard, 1993 and Mclver, 2002). However, there appears to be a number of limitations associated with data utilization such as weather conditions, data availability, considerable costs, and time for acquiring and processing large numbers of images. Additionally, improving the classification accuracy and reducing the classification time have long been goals of remote sensing research and they still require further study.

A primary goal of this research is to manage the challenges described above. To accomplish this goal, improvements must be made to the classification algorithms that can be applied to MODIS-EVI time-series imagery.. A supervised self-organizing map (SSOM) and a soft supervised self-organizing map (3SOM) are modified and improved to increase classification efficiency and accuracy.

This research is designed be comprised of two parts for the purpose of thorough investigation. The first part is to test and develop the suitable classification method by using the synthetic data. Six experiments, which are performed in this first part, consist of: 1) comparison

of input images between time-series and phenology images using SSOM, 2) selection of the suitable neural network configuration of the back-propagation neural network classifier (BPNN) and SSOM, 3) comparative evaluation of the SSOM with the Gaussian maximum likelihood classifier (GMLC) and BPNN, 4) comparative evaluation of the SSOM with the 3SOM, 5) comparative evaluation of the fully-3SOM with the partially-3SOM, and 6) accessing the uncertainty in classification accuracy of SSOM.

In addition to the synthetic data component, the second part applies the identified suitable method to real landscape data derived from MODIS-EVI time-series images. Two study areas in Thailand and the Midwestern U.S. are selected based on differences of land cover characteristics. The 3SOM-F is employed in both study areas to confirm that its classification performance of is effective even when it is applied to real landscape data. Moreover, the classification results are used to examine how the characteristic of land cover affect the capability of this method.

The main results of this dissertation are as follows:

6.1.1 Testing and developing a suitable method using synthetic data

The synthetic data is utilized to test and develop a suitable method of classification. The first experiment is to determine appropriate input data to be used for land cover classification. In this experiment, TIME and PHEN are utilized in the SSOM classification. With 300 simulations, the results of classification derived from TIME clearly demonstrate that the SSOM produces considerably higher classification accuracy than the classification results derived from PHEN. The SSOM is effectively applicable to large datasets due to its ability to map high dimensional

inputs onto low dimensional units and preserve the topology of input patterns in the low dimension after dimension reduction (Bagan *et al*., 2005).

The second experiment is designed to investigate the suitable neural network configuration of the SSOM and BPNN. The performance of the BPNN and SSOM classifiers are examined by setting different network structures and internal parameter values to find suitable values that produce the highest classification accuracy. To investigate the performance of the BPNN classifier, 455 neural network configurations are formed based on three primary parameters: number of hidden layer neurons (HN), learning rate and momentum factor (LR&MF), and number of iterations (ITER). To investigate the performance of the SSOM classifier, 300 different neural network configurations are generated based on three primary parameters: number of competitive layer neurons (NET), number of iterations (ITER), and initial learning rate (LR). Both series of configurations are operated on a case-by-case basis by trial-and-error analysis. The results suggest that making some adjustments to the network structure and parameter values will improve the performance of the neural network classification. The suitable neural network configurations are applied in the subsequent experiments.

The accuracy of three hard classifiers including the GMLC, BPNN, and SSOM are evaluated in the third experiment. Two tests in this experiment consist of the comparative evaluation of the GMLC, BPNN, and SSOM with different simulated input data and different random training data. The first test utilizes 500 different simulated input datasets, while the second test uses 500 different training datasets. The results demonstrate that the SSOM achieves more meaningful classification results than those obtained from the GMLC and BPNN for both tests. With the robust architecture and effective learning process, the SSOM is able to provide stable results with only a small variation in classification accuracy.

The fourth experiment compares the performance between the SSOM and 3SOM using different simulated input data and different random training data. The results indicate that the 3SOM employed as a soft classification delivers a more accurate classification than the SSOM applied as a hard classification for both tests. The classification accuracy of the 3SOM is higher than the SSOM in all measures. The supported reason is that the SSOM assigns each pixel to a single class. In reality, many pixels in an image may represent more than one land cover class on the ground. To allocate a mixed pixel to a single land cover class not only provides an unrealistic result, but can also lead to an inaccurate representation of land cover (Thornton *et al.*, 2006). On the other hand, the 3SOM can deal with the mixed pixel problem and present proportions of land cover classes in a pixel instead of a single class. Therefore, the 3SOM provides a more realistic and accurate land cover representation than the SSOM.

The fifth experiment compares two methods of soft classification. A hard training dataset containing only pure pixels is utilized with the 3SOM-P and a soft training data consisting of both pure and mixed pixels is employed with the 3SOM-F. Two tests are also performed to evaluate and compare the classification accuracy between the 3SOM-F and 3SOM-P. The first test is a comparative evaluation between the 3SOM-F and 3SOM-P using different simulated input data, while the second test is a comparative evaluation between them using different random training data. According to the results, the 3SOM-F presents superior performance over the 3SOM-P in both tests. The 3SOM-F accomplishes more accurate classification results than the 3SOM-P; however, discrepancies in the classification accuracy between the 3SOM-F and 3SOM-P are highlighted in heterogeneous areas. This is due to the fact that the dominant classes and subsidiary classes in a pixel can be well recognized if it is trained with pure and mixed pixels. This suggests that introducing additional variability of spectral signatures, which is

characteristic of mixed pixels in the classifier, helps the network to generalize unknown patterns of data.

In addition to improving the classification accuracy, the sensitivity and reliability of the classification output are also important subjects for image classification. Consequently, the last experiment in this section evaluates how uncertainty in the input data, the training data, and the classifier affects the classification accuracy of the SSOM. For uncertainty associated with the input data, increasing the levels of noise has an extensive influence on the classification accuracy, particularly in areas of mixed pixels but it has a marginal effect on the uncertainty in classification accuracy. This indicates that the SSOM is a stable and robust classifier providing precise accuracy. For uncertainty associated with the training data, randomly selecting training datasets has a small impact on the uncertainty in classification accuracy due to the use of synthetic data and the procedure of selecting training samples. Interestingly, although the same training data is applied, the SSOM with different sequences of training data still produces a variation in classification accuracy. For uncertainty associated with the classifier, several parameters including NET, W, ITER, and LR are selected to evaluate the sensitivity of the SSOM. The results indicate that a very small NET size leads to a low possibility of accuracy because there is not enough space to cluster land cover classes. A very large NET size does not only improve classification accuracy, but also minimizes the capability of the SSOM to generate unknown patterns but requires extensive computational time. As a result, this study found that the minimum NET should correspond to the number of land cover diversities. The initial weight does not apparently have an influence on the SSOM classification accuracy, while ITER has a slight effect on the classification uncertainty. Therefore, increasing ITER does not improve the classification accuracy; conversely, it extensively increases computational time in the learning

process. The results of this study also indicate that the SSOM provides high classification accuracy at low values of LR, but the large values of LR tends to provide low uncertainty in classification accuracy. Additionally, an appropriate LR value depends on the diversity of the study area and the complexity of the input data.

6.1.2 Applying identified method using real landscape data

The next part is aimed at applying and confirming that the classification performance of the identified method, the 3SOM-F, is effective even when it is applied to the real landscape data. The MODIS-EVI 16-day 250 m (MOD13Q1) data from 2010 is applied in this study.

Two study areas, Thailand and the Midwestern U.S., with different land cover characteristics are selected to investigate the performance of the 3SOM-F classification and to examine how the land cover characteristics affect the ability of this method. The first study area is focused on four major crops: forest, sugar cane, cassava, and paddy rice, grown in Lopburi, central Thailand and the second study area is focused on two summer crops: corn and soybeans planted in Iowa, which is located in the Midwestern U.S. To simplify the analysis, only subsets of 12.5 x 12.5 sq.km. (50 x 50 pixels) within these regions are selected. Both real landscape data subsets are classified using the 3SOM-F with the suitable neural network configuration. The classification accuracy is evaluated using the measures of accuracy assessment for soft classification. The results show that the 3SOM-F successfully classified land cover in both Thailand and the U.S. In the Thailand study area, all measures illustrate that the 3SOM-F presents high classification accuracy for all classes. Paddy rice yields the highest classification accuracy, while there is some confusion between cassava and sugar cane. However, some areas of scattered misclassification are found in heterogeneous or mixed pixel regions. In the Midwestern U.S. study area the 3SOM-F produces high accuracy in all measures for corn and

soybeans. However, confusion between corn and soybeans is apparent in homogenous or pure pixel areas, which resulted in erroneous and misleading interpretations.

Dissimilarities in the results from the two study sites are related to spatial and spectral confusion. The agricultural areas in Thailand are small in size and tend to be highly heterogeneous, resulting in mixed pixels containing multiple land cover classes in a given pixel. This type of spatial confusion diminishes the classification accuracy. However, spatial confusion is likely to have only a small impact on the performance of the 3SOM-F when the EVI temporal profiles for each class are distinctly different from each other. In the Midwestern U.S., spectral confusion is caused by similar EVI temporal profiles of corn and soybeans, which leads to classification difficulty and results in erroneous and misleading interpretations. When spectral confusion of similar EVI temporal profiles occurs, the 3SOM-F underestimates area proportions in heterogeneous or pure pixel areas.

In summary, eight research objectives have been successfully addressed in this dissertation. A supervised self-organizing map (SSOM) and a soft supervised self-organizing map (3SOM) are modified and improved to increase classification efficiency and accuracy. The results show that the 3SOM provides an alternative technique for land cover classification by using the MODIS-EVI time-series images. This research contributes to the field of remotely sensed image classification as follows:

1) When utilizing MODIS-EVI time-series images, the soft supervised self-organizing map (3SOM) is a significant alternative technique which is used to increase the efficiency and accuracy of land cover classification.

2) With the SSOM, land cover images derived from TIME achieves more meaningful classification result than those derived from PHEN.

3) The results of determining the optimal architecture and learning factor values of the neural network provide guidance to users regarding the selection of appropriate neural network parameters.

4) The SSOM can provide a promising alternative to the Gaussian maximum likelihood classifier (GMLC) and the backpropagation neural network (BPNN) for land cover classification regarding the applicability of the MODIS-EVI time-series images.

5) With coarse resolution images, (i.e., MODIS-EVI time-series images), the soft classification performs better than hard classification for land cover mapping using the SSOM, and provides more informative and accurate results,

6) In the learning process of the 3SOM, using both pure and mixed training data (fully-soft classification) can yield better decompositions of mixed pixels than using only pure pixels (partially-soft classification).

7) Classification uncertainty associated with the input data, training data, and the classifier can be used to explain the sensitivity of the classifier and the reliability of the classification output.

8) The classification performance of the 3SOM-F is effective even when it is applied to the real landscape data from both Thailand and the Midwestern U.S.

6.2 Benefits and limitations

Based on extensive experiments, the results suggests that the soft classification method is an option which takes the mixture in a pixel as a part of the classification process into consideration in order to model coarse spatial resolution remotely sensed data. This research affirms that the 3SOM-F is an efficient soft classification method for land cover classification

with time-series imagery. Moreover, this method also has the potential to describe and model real landscape variation within remotely sensed images.

In addition to the ability to handle large and diverse datasets, the 3SOM-F has increasingly adapted its performance and flexibility for multiple class analyses. As demonstrated in this research, this method not only provides significant information concerning the classification results, but it also allocates accurate proportions of land cover classes in a pixel.

The methods developed in this study can benefit researchers who employ coarse remote sensing imagery for detailed land cover image classification. Additionally, this method should be applicable to other images from any remote sensing system.

The proposed method, the 3SOM-F, will benefit land cover classification at the regional scale. The spatial pattern of land cover classes would be valuable information for managing and understanding the environment as well as monitoring land cover change. Furthermore, the advantages of this research will contribute to various disciplines such as map updating, agricultural area estimation, cartography, and urban planning.

However, the experimental results presented in this research clearly show that the performance of the 3SOM-F depends on the selection of network architecture and internal parameter settings. This selection has a significant influence on classification accuracy and uncertainty. Different datasets need different valid configurations. Consequently, it is necessary to determine network models by trial and error. However, this process is generally computationally intensive and time consuming.

Landscape characteristics have an influence on the performance of the 3SOM-F. This research finds that high classification confusion exists in the areas of heterogeneity or mixed pixels. This spatial confusion is found where the agriculture areas are small in and contain more

than one land cover classes in a pixel. On the other hand, erroneous and misleading interpretations are also found in the areas of homogeneity due to spectral confusion caused by similar EVI temporal profiles. Therefore, careful consideration of these characteristics is very important when classifying land cover by 3SOM-F.

In addition to spatial and spectral confusions, the number of classes is another factor of confusion. The greater number of classes mixed within a pixel, the more errors are found. Therefore, the number of classes in the image would affect the ability of the 3SOM-F. This problem is more noticeable if those classes have similar EVI temporal profiles. Accordingly, the classification accuracy of the 3SOM-F will decrease as the number of classes increase in one pixel. In other words, the 3SOM-F allocates fewer classes within a pixel more effectively than a larger number of classes. .

Appropriate accuracy evaluation measures are needed to assess the value of soft classification. According to this study, although area error proportion (AEP) is a valuable measure of error, the use of only AEP is not sufficient to assess the performance and accuracy of soft classification. This is due to the fact that AEP does not take the spatial distribution of omission and commission errors into account. Therefore, use of additional measures (i.e. correlation coefficient (CC), root mean square error (RMSE), and closeness (S)) is required to provide reliable output quantification and interpretation.

Furthermore, when applying the 3SOM-F to satellite time-series data, computational time should be a concern. The classification time of this method has a linear relationship with the size of the image and the number of data dimensions. In addition, the size of the training data and the number of classes have significant effects on computational time. The classification of large images employing large training datasets to classify several land cover classes is computationally

intensive and time consuming; therefore, computer performance should be deliberately considered for this classification process.

As a final point, it is important to note that the effectiveness of this proposed method depends on the study area, data resolution, and data dimension. Consequently, when this method is applied to other regions, it is essential to investigate the performance and classification results of this method.

6.3 Further research

This research shows that the 3SOM-F has great utility for land cover classification with time-series data and can also be applied different regions around the world. . Furthermore, the 3SOM-F is able to replace costly field surveys for crop area estimation and long-term monitoring of cropping intensity over large-scale areas. However, due to the limitations and challenges found in this research, several further studies are highlighted.

Future research will concentrate on extending the current algorithm to handle land cover features in other regions with different data resolutions and sensors as well as different crop types. Additionally, extensive development of the classification algorithm is the most important aspect of future research. Improvement of the 3SOM-F classifier should be developed, and its capability tested, in order to increase classification efficiency and accuracy.

Issues concerning suitable dates for land cover classification also need to be addressed. For example, selecting the appropriate dates of MODIS time-series data or the appropriate phenological parameters may improve classification accuracy and also reduce computation time.

In addition to suitable dates, investigation of training data size should be included for further studies. The suitable training data size may enhance the classification accuracy and

reduce processing time. Moreover, it is challenging to understand the sensitivity of classification accuracy among different training data size selections; therefore, the uncertainty in classification accuracy associated with the size of training data is still essential to be examined.

APPENDICES

Appendix A

EVI time-series data applied for simulating the synthetic data

Table A.1 EVI time-series data applied for simulating the synthetic data.

(a) Mean of EVI time-series data

Time (bi-weekly)	Land cover class			
	A	B	C	D
1	0.363	0.320	0.275	0.204
2	0.368	0.279	0.244	0.216
3	0.364	0.251	0.228	0.226
4	0.359	0.228	0.215	0.233
5	0.342	0.215	0.202	0.238
6	0.334	0.211	0.190	0.239
7	0.351	0.218	0.197	0.239
8	0.404	0.240	0.225	0.236
9	0.482	0.275	0.271	0.253
10	0.552	0.332	0.339	0.273
11	0.615	0.394	0.410	0.308
12	0.660	0.453	0.479	0.344
13	0.681	0.514	0.548	0.392
14	0.695	0.561	0.607	0.435
15	0.722	0.595	0.649	0.483
16	0.727	0.621	0.657	0.573
17	0.693	0.624	0.627	0.639
18	0.649	0.600	0.576	0.638
19	0.593	0.572	0.515	0.562
20	0.528	0.537	0.451	0.436
21	0.466	0.499	0.393	0.316
22	0.418	0.454	0.341	0.238
23	0.389	0.402	0.301	0.209

(b) Standard deviation of EVI time-series data

Time (bi-weekly)	Land cover class			
	A	B	C	D
1	0.052	0.065	0.047	0.044
2	0.052	0.064	0.042	0.042
3	0.052	0.057	0.040	0.041
4	0.054	0.044	0.039	0.038
5	0.056	0.039	0.039	0.033
6	0.058	0.042	0.041	0.031
7	0.063	0.047	0.042	0.032
8	0.066	0.052	0.044	0.040
9	0.072	0.058	0.052	0.046
10	0.078	0.061	0.066	0.054
11	0.081	0.067	0.078	0.060
12	0.086	0.074	0.087	0.065
13	0.087	0.085	0.094	0.071
14	0.087	0.089	0.096	0.081
15	0.086	0.087	0.094	0.081
16	0.083	0.081	0.084	0.077
17	0.077	0.078	0.081	0.086
18	0.073	0.078	0.081	0.085
19	0.063	0.075	0.077	0.071
20	0.057	0.066	0.076	0.046
21	0.061	0.064	0.070	0.038
22	0.063	0.060	0.061	0.041
23	0.059	0.056	0.051	0.044

Appendix B

Python code for 3SOM classification

```
from __future__ import division

##-------------------------------------------------------------------------------------------
##Soft-supervied self-organizing map (3SOM) Classification
##Written by Siam Lawawirojwong
##Created July 2011 (Modified December 2011)
##
##This program employs hard/soft classification
##
##Reference:
##    Kohonen, T. (1989). Self-Organization and Associatiive Memory (3rd ed.). Berlin:Srpinger.
##    Marvin Minsky (www.ai-junkie.com/ann/som/som1.html)
##-------------------------------------------------------------------------------------------

from random import *
from math import *
import numpy

class Network:

    def __init__(self, height, width, fv_size, cv_size, learning_rate):
        self.height = height
        self.width = width
        self.neurons = height*width
        self.net = numpy.zeros(((fv_size+cv_size),height,width)).astype(float)
        self.neuron_pos = self.neuron_position(height, width)
        self.learning_rate = learning_rate
        self.radius = 0.5*(height+width)
        self.time_constant = 0.0
        self.fv_size = fv_size
        self.cv_size = cv_size

    def train(self, iterations, train_vector, train_shuffle=False):
        self.time_constant = iterations/log(self.radius)
        for i in range(1,iterations+1):
            if train_shuffle:
                shuffle(train_vector)
            radius = self.radius_decay(i)
            learning_rate = self.learning_rate_decay(i)
```

```
        for  j in range(len(train_vector)):
            input_fv = numpy.array(train_vector[j][0])
            input_cv = numpy.array(train_vector[j][1])
            bmu = self.best_match(input_fv)
            dist = self.distance(bmu)
            influence = self.influence_decay(dist, radius, i)
            self.update(input_fv, input_cv, influence, learning_rate)

def radius_decay(self, t):
    """ returns the radius of influence for current epoch """
    return self.radius*exp(-float(t/self.time_constant))

def learning_rate_decay(self, t):
    """ returns the learning rate for current epoch """
    return self.learning_rate*exp(-float(t/self.time_constant))

def influence_decay(self, dist, radius, t):
    """ calculates the neiborhood function depending on dist to bmu and current radius"""
    return numpy.exp(-1.0*(dist**2/(2*radius*t)))*(dist < radius)

def best_match(self, fv):
    f_vec = fv.reshape(self.fv_size, 1, 1)
    f_arr = self.net[:self.fv_size]
    temp = numpy.sqrt(((f_arr-f_vec)**2).sum(0))
    pos = numpy.argmin(temp)
    return = pos/self.width, pos%self.width

def neuron_position(self, width, height):
    x = numpy.array([[i for i in range(width)]]*height)
    y = numpy.array([[i for i in range(height)]]*width).transpose()
    return numpy.array([y,x])

def distance(self, bmu):
    bmu = numpy.array(bmu).reshape(len(bmu),1,1)
    dist = numpy.sqrt(((self.neuron_pos - bmu)**2).sum(0))
    return dist

def update(self, fv, cv, influence, learning_rate):
    input_vec = numpy.append(fv, cv).reshape(self.fv_size+self.cv_size, 1, 1)
    self.net = self.net+influence*learning_rate*(input_vec-self.net)
    return

def classify_pattern(self, fv):
    fv = numpy.array(fv)
    best = self.best_match(fv)
    return self.net[:,best[0],best[1]][self.fv_size:]
```

REFERENCES

REFERENCES

Agrawal, S., Singh, S., Joshi, P. K., & Roy, P. S. (2006). Phenology based classification model for vegetation mapping using IRS-WiFS. In *GLC 2000 'First Results' Workshop.*

Atkinson, P. M. & Tatnall, A. R. L. (1997) Neural networks in remote sensing: introduction. *International Journal of Remote Sensing*, 18(4), 699-709.

Atkinson, P. M., & Foody, G. M. (2002). Uncertainty in remote sensing and GIS: fundamentals. In Foody, G.M., & Atkinson, P.M. (Eds.), *Uncertainty in remote sensing and GIS* (pp.1-18). England: John Wiley & Sons Ltd.

Atkinson, P. M. (2005). Sub-pixel target mapping from soft-classified, remotely sensed imagery. *Photogrammetric Engineering & Remote Sensing, 71*(7), 839-846.

Bagan, H., Wang, Q., Watanabe, M., Yang, Y., & Ma, J. (2005). Land cover classification from MODIS EVI times-series data using SOM neural network. *International Journal of Remote Sensing, 26*(22), 4999-5012.

Bardossy, A., & Samaniego, L. (2002). Fuzzy rule-based classification of remotely-sensed imagery. *IEEE Transaction on Geoscience and Remote Sensing, 40*(2), 362-374.

Benediktsson, J. A., Swain, P. H., & Ersoy, O. K. (1990) Neural network approaches versus statistical methods in classification of multisource remote sensing data. *IEEE Transactions on Geoscience & Remote Sensing*, 28(4), 540-552.

Bernard, A. C., Wilkinson, G. G., & Kanellopoulos, I. (1997). Training strategies for neural network soft classification of remotely-sensed imagery. *International Journal of Remote Sensing, 18*(8), 1851-1856.

Bischof, H., Schneider, W., & Pinz, A. J. (1992) Multispectral classification of landsat-images using neural networks. *IEEE Transactions on Geoscience & Remote Sensing*, 30(3), 482-490.

Canters, F. (1997). Evaluating the uncertainty of area estimates derived from fuzzy land-cover classification. *Photogrammetric Engineering & Remote Sensing 63*, 403-414.

Carpenter, G. A., Gjaja, M. N., Gopal, S., & Woodcock, C. E. (1997). ART neural networks for remote sensing: vegetation classification from Landsat TM and terrain data. *IEEE Transaction on Geoscience and Remote Sensing, 35*(2), 308-325.

Chen, J., Jonsson, P., Tamura, M., Gu, Z., Matsushita, B., & Eklundh, L. (2004). A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky-Golay filter. *Remote Sensing of Environment, 91*, 332-344

Civco, D. L. & Wang, Y. (1994) Classification of multispectral, multitemporal, multisource spatial data using artificial neural networks. Proceeding of the 1994 ASPRS/ACSM Convention 1, 123-133.

Crane, A. (1992) Example based filters: Post-processing classification image with neural networks (Master thesis). University of Wisconsin-Madison.

Dai, X., Guo, Z., Zhang, L., & Wu, J. Spatio-temporal pattern of urban land cover evolvement with urban renewal and expansion in Shanghai based on mixed-pixel classification for remote sensing imagery. *International Journal of Remote Sensing, 31*(23), 6095-6114.

Dalstra, H. (2008) Development of a Multi-temporal remote sensing classification methodology for nature classes in Dutch land-use database: a phenology-based approach [Abstract]. Laboratory of Geo-Information Science and Remote Sensing, Wageningen University. Retrived February 12, 2009. From http://www.grs.wur.nl/UK/newsagenda/agenda/Development_of_a_MultiTemporal_Remote_Sensing_Classification_Methodology_for_Nature_Classes_in_the_D.htm

Doraiswamy, P. C., & Stern, A. J. (2007) Crop classification in the U.S. corn belt using MODIS imagery. In *International Geoscience & Remote Sensing Symposium. Barcelona, Spain*

Dungan, J. L. (2002). Toward a comprehensive view of uncertainty in remote sensing analysis. In Foody, G.M., & Atkinson, P.M. (Eds.), *Uncertainty in remote sensing and GIS* (pp. 25-35). England: John Wiley & Sons Ltd.

Dymond, C. C., Mladenoff, D. J., & Radeloff, V. C. (2002). Phenological differences in tasseled cap indices improve deciduous forest classification. *Remote Sensing of Environment, 80,* 460-472.

Eastman, J. R., & Laney, R. M. (2002) Bayesian soft classification for sub-pixel analysis: a critical evaluation. *Photogrammetric Engineering & Remote Sensing, 68(*11), 1149-1154.

Fisher, P. F. (1994). Visualization of the reliability in classified remotely sensed images. *Phtogrammetric Engineering and Remote Sensing, 60*, 905-910.

Foody, G. M., & Cox, D. P. (1994). Sub-pixel land cover composition estimation using a linear mixture model and fuzzy membership functions. *International Journal of Remote Sensing, 15*(3), 619-631.

Foody, G. M. (1992). Derivation and applications of probabilistic measures of class membership from the maximum likelihood classification. *Phtogrammetric Engineering and Remote Sensing, 58*, 1335-1341.

Foody, G. M. (1995). Cross-entropy for the evaluation of the accuracy of a fuzzy land cover classification with fuzzy ground data. *ISPRS Journal of Photogrammetry and Remote Sensing, 50,* 2-12.

Foody, G, M. (1996a). Approaches for the production and evaluation of fuzzy land cover classification form remotely-sensed data. *International Journal of Remote Sensing, 17*(7), 1317-1340.

Foody, G, M. (1996b). Relating the land-cover composition of mixed pixels to artificial neural network classification output. *Photogrammetric Engineering and Remote Sensing, 62*(5), 491-499.

Foody, G. M. (1996c). Fuzzy modeling of vegetation from remotely sensed imagery. *Ecological Modelling, 85,* 3-12.

Foody, G. M. (1997). Fully fuzzy supervised classification of land cover from remotely sensed imagery with an artificial neural network. *Neural Computing Applications, 5,* 238-247.

Food, G. M. (2002). Status of land cover classification accuracy assessment. *Remote Sensing of Environment, 80,* 185-201.

Gao, X., Huete, A. R., Ni, W., & Miura, T. (2000). Optical-biophysical relationships of vegetation spectra without background contamination. *Remote Sensing of Environment, 74,* 609-620

Goodchild, M., & Gopal, S. (eds.). (1989). *Accuracy of Spatial Databases.* London: Taylor and Francis.

Gopal, S., Woodcock, C. E., & Strahler, A. H. (1999) Fuzzy neural network classification of global land cover from a 1^0 AVHRR data set. *Remote Sensing of Environment*, 67, 230-243.

Hagan, M. T., Demuth, H. B., & Beale, M. (1996). *Neural network design*. Boston: PWS Publishing Company.

Heermann, P. D., & Khazenie, N. (1990) Application of neural networks for classification of multi-source multi-spectral remote sensing data. Proceedings of 1990 IEEE International Geoscience & Remote Sensing Symposium.

Heuvelink, G. B. M., & Burrough, P. A. (1993). Error propagation in cartographic modelling using Boolean logic, and continuous classification. *International Journal of Geographic Information System, 7,* 231-246.

Heuvelink, G. B. M. (1998). *Error propagation in Environmental Modelling with GIS.* London: Taylor and Francis.

Heuvelink, G. B. M. (1999). Propagation of error in spatial modelling with GIS. *In Longley, P. A., Goodchild, M. F., Maguire, D. J, Rhind, D. W. (Eds), 1999, Geographical Information Systems* (2nd ed., pp. 207-217). New York: Wiley.

Hewitson, B. C., & Crane, R. G. (1994) *Neural nets: applications in geography*. Kluwer Academic Publishers, Boston.

Houghton, J. T., Jenkins, G. J., & Ephraens J. J. (1990). *Climate change*. The IPCC scientific assessment (pp. 365-266). Cambridge University Press.

Hrycej, T. (1992) *Modular learning in neural networks: a modularized approach to neural network classification*. A Wiley-Interscience Publication, New York.

Hu, X. (2009). Impervious surface estimation from remote sensing imagery using sub-pixel and object-based classifications in Indianapolis, USA. (Doctoral dissertation). Available from ProQuest Dissertation & Theses database. (UMI No. 3394733)

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment, 83*(1-2), 195-213.

Huete, A., Liu, H. Q., Batchily, K., & van Leeuwan, W. (1997). A comparison of vegetation indices over a global set of TM images for EOS-MODIS. *Remote Sensing of Environment, 59*, 440-451.

Ibrahim, M. A., Arora, M. K., & Ghosh, S. K.(2005). Estimating and accommodating uncertainty through the soft classification of remote sensing data. *International Journal of Remote Sensing, 26*(14), 2995-3007.

Jin, S. M., & Sader, S. A. (2005). MODIS time-series imagery for forest disturbance detection and quantification of patch size effects. *Remote Sensing of Environment, 99*, 462-470

Jonsson P., & Eklundh, L. (2004) TIMESAT—a program for analyzing time-series of satellite sensor data. *Computers & Geosciences, 30*, 833-845.

Jonsson P., & Eklundh, L. (2006). *TIMESAT—a program for analyzing time-series of satellite sensor data: users guide for TIMESAT 2.3*, MALMO AND LUND.

Kavzoglu T., & Mather P. M. (2003). The use of backpropagating artificial neural networks in land cover classification. *International Journal of Remote Sensing, 24*(23), 4907-4938.

Ke, J., Liu, X., & Wang, G. (2008) Theoretical and empirical analysis of the learning rate and momentum factor in neural network modeling for stock prediction. Advances in Computation and Intelligence (Lecture Notes in Computer Science), 5370, 697-706.

Key, J., Maslanik, J. A., & Schweiger, A. J. (1989) Classification of merged AVHRR and SMMR arctic data with neural networks. *Photogrammetric Engineering & Remote Sensing*, 55(9), 1331-1338.

Kohonen, T. (1989). *Self-Organization and Associative Memory* (3rd ed.). Berlin: Srpinger.

Kohonen, T. (1990). The *Self-Organization Map.* In *Proceedings of the Ieee* 78, pp.1464-1480.

Leite, P. B. C., Feitosa, R. Q., Formaggio, A. R., Costa, G. A. O. P., Pakzad, K., & Sanches, I. D. A. (2008) Crop type recognition based on hidden markov models of plant phenology. In *XXI Brazilian Symposium on Computer Graphics and Image Processing*, pp.27-34.

Lek, S. & Guegan, J. F. (1999) Artificial neural networks as a tool in ecological modelling, an introduction. *Ecological Modelling*, 120, 65-73.

Li, Z. & Eastman, J. R. (2006). The nature and classification of unlabelled neurons in the use of Kohonen's self-organizing map for supervised classification. *Transaction in GIS*, 10(4), 599-613.

Li, Z. (2007). Development of soft classification algorithms for neural network models in the use of remotely sensed imagery classification (Doctoral dissertation). Available from ProQuest Dissertation & Theses database. (UMI No. 3282765).

Li, Z. (2008) Fuzzy ARTMAP based neurocomputational Spatial Uncertainty measures. *Photogrammetric Engineering & Remote Sensing*, 74(12), 1573-1584.

Li, Z., & Eastman, J.R. (2010). Commitment and typicality measures for the self-organizing map. *International Journal of Remote Sensing*, *31*(16), 4265-4280.

Liu, L., Wang, B., & Zhang, L. (2010). An approach based on self-organizing map and fuzzy membership for decomposition of mixed pixels in hyperspectral imagery. *Pattern Recognition Letters, 31*, 1388-1395.

LP DAAC (2008). *MODIS Reprojection Tool User's Manual*. USGS Earth Resources Observation and Science (EROS) Center.

Maslanik, J., Key, J., & Schweiger, A. (1990) Neural network identification of sea ice seasons in passive microwave data. Proceedings of 1990 IEEE International Geoscience & Remote Sensing Symposium: 1281-1284.

Mather, P.M. (1987). *Computer Processing of Remotely-Sensed Images.* Chichester: Wiley.

Mclver, D. K. (2002) Adapting machine learning methods for coarse resolution land cover classification (Doctoral dissertation). Available from ProQuest Dissertation & Theses database. (UMI No. 3026421).

Merry, C., Wright, D., Wentz, E., Anderson, S., Budge, A., & Hepner, G. (2000). *Remotely acquired data and information in GIScience. 2000 Research White Papers*, UCGIS.

Mill, H., Cutler, M. E. J., & Fairbairn, D. (2006) Artificial neural networks for mapping regional-scale upland vegetation from high spatial resolution imagery. *International Journal of Remote Sensing*, 27(11), 2177-2195.

Myneni, R. B., Keeling, C. D., Tucker, C. J., Asrar, G., & Nemani, R. P. (1997). Increased plant growth in the northern high latitudes from 1981 to 1991. *Nature, 386*, 698-702.

Nelson, M. M., & Illingworth, W. T. (1991). *A practical guide to neural networks*. Addison-Wesely, New York.

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

Roehrig, J., Thamm, H. P., Menz, G., Porembski, S., & Orthmann, B. (2005). A phenological classification approach for the upper Oueme in Benin, West Africa using SPOT VEGTATION. In *Proceeding of the Second International VEGETATION User Conference.*

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

Schalkoff, R. J. (1992). *Pattern recognition: statistical, structural and neural approaches*. New York. Wiley.

Schwartz, M. D. (2003). *Phenology: an integrative environmental science*. The Netherlands: Kluwer Academic Publishers.

Seetha, M., Muralikrishna, I. V., Deekshatulu, B. L., Life Fellow, Malleswari, B. L., Nagaratna, & Hegde, P. (2008). Artificial neural networks and other methods of image classification. *Journal of Theoretical and Applied Information Technology*. 1039-1053.

Simonneaux, V., Duchemin, B., Helson, D., Er-Raki, S., Olioso, A., & Chehbouni, A. G. (2007) The use of high-resolution image time series for crop classification and evapotranspiration estimate over an irrigated area in central Morocco. *International Journal of Remote Sensing, 29*(1), 95-116.

Tatem, A. J., Lewis, H. G., Atkinson, P. M., & Nixon, M. S. (2002). Super-resolution land cover mapping from remotely sensed imagery using a Hopfield Neural Network. In Foody, G. M., & Atkinson, P. M. (Eds.), *Uncertainty in remote sensing and GIS* (pp77-98). England: John Wiley & Sons Ltd.

Thornton, M. W., Atkinson, P. M., & Holland, D. A. (2006). Sub-pixel mapping of rural land cover objects from fine spatial resolution satellite sensor imagery using super-resolution pixel-swapping. *International Journal of Remote Sensing, 27*(3), 473-491.

Tu, J. V. (1996). Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *Journal of Clinical Epidemiology*, 49(11), 1225-1231.

Villmann, T., Merenyi, E., & Hammer, B. (2003). Neural maps in remote sensing image analysis. *Neural networks*, 16, 389-403.

Wardlow, B. D., Egbert, S. L., & Kastens, J. H. (2007). Analysis of time-series modis 250 m vegetation index data for crop classification in the us central great plains. *Remote Sensing Of Environment, 108*(3), 290-310.

Watanachaturaporn, P. (2005). Classification of remote sensing images using support vector machines (Doctoral dissertation). Available from ProQuest Dissertation & Theses database. (UMI No. 3177025)

Xu, M., Watanachaturaporn, P., Varshney, P.K., & Arora, M.K. (2005). Decision tree regression for soft classification of remote sensing data. *Remote Sensing of Environment, 97*, 322-336.

Yang, C. C. (2005) Landmine detection and classification through hybrid neural networks and fuzzy set. Master of Science Thesis, Department of Electrical Engineering, Pennsylvania State University.

Zhan, Q., Molenaar, M., & Lucieer, A. (2002). Pixel unmixing at the sub-pixel scale based on land cover class probabilities: application to urban areas. In Foody, G.M., & Atkinson, P.M. (Eds.), *Uncertainty in remote sensing and GIS* (pp59-76). England: John Wiley & Sons Ltd.

Zhang, J., & Goodchild, M. F. (2002). *Uncertainty in Geographical Information.* London: Taylor and Francis.

Zhang, J., & Foody, G. M. (1998). A fuzzy classification of sub-urban land cover from remotely sensed imagery. *International Journal of Remote Sensing, 19*(4), 2271-2738.

Zhang, J., &Foody, G. M. (2001). Fully-fuzzy supervised classification of sub-urban land cover from remotely sensed imagery: Statistical and artificial neural network approaches. *International Journal of Remote Sensing, 22*(4), 615-628.

Zhang, X. X., Wang, M. J., Zheng, J. Y., Zhu, Q. K., & Ma, J. (2006). Building NDVI-Phenology comparison method to detect growing periods during 1982 – 1999 in Northeast China. In *Proceedings of the ASPRS Mid-term Symposium*.

Zhang, X. Y., Friedl, M. A., Schaaf, C. B., Strahler, A. H., Hodges, J. C. F., Reed, B. C., & Huete, A. (2003). Monitoring vegetation phenology using MODIS. *Remote sensing of environment, 84*(3),471-475.