Article

Integration of Satellite Imagery, Topography and Human Disturbance Factors Based on Canonical Correspondence Analysis Ordination for Mountain Vegetation Mapping: A Case Study in Yunnan, China

Zhiming Zhang [1],*, Frieke van Coillie [2], Xiaokun Ou [1],* and Robert de Wulf [2]

[1] Institute of Ecology and Geobotany, School of Life Science, Yunnan University, Kunming 650091, China
[2] Laboratory of Forest Management and Spatial Information Techniques, Ghent University, 9000 Ghent, Belgium; E-Mails: Frieke.VanCoillie@UGent.be (F.V.C.); Robert.DeWulf@UGent.be (R.D.W.)

* Authors to whom correspondence should be addressed;
E-Mails: zhiming_zhang76@hotmail.com (Z.Z.); xkou@yun.edu.cn (X.O.);
Tel.: +86-8716-5033-785; Fax: +86-8716-5165-581.

Abstract: The integration between vegetation data, human disturbance factors, and geo-spatial data (Digital Elevation Model (DEM) and image data) is a particular challenge for vegetation mapping in mountainous areas. The present study aimed to incorporate the relationships between species distribution (or vegetation spatial distribution pattern) and topography and human disturbance factors with remote sensing data, to improve the accuracy of mountain vegetation maps. Two different mountainous areas located in Lancang (Mekong) watershed served as study sites. An Artificial Neural Network (ANN) architecture classification was used as image classification protocol. In addition, canonical correspondence analysis (CCA) ordination was applied to address the relationships between topography and human disturbance factors with the spatial distribution of vegetation patterns. We used ordinary kriging at unobserved locations to predict the CCA scores. The CCA ordination results showed that the vegetation spatial distribution patterns are strongly affected by topography and human disturbance factors. The overall accuracy of vegetation classification was significantly improved by incorporating DEM or four CCA axes as additional channels in both the northern and southern study areas. However, there was no significant difference between using DEM or four CCA axes as extra channels in

the northern steep mountainous areas because of a strong redundancy between CCA axes and DEM data. In the southern lower mountainous areas, the accuracy was significantly higher using four CCA axes as extra bands, compared to using DEM as an extra band. In the southern study area, the variance of vegetation data explained by human disturbance factors was larger than the variance explained by topographic attributes.

Keywords: mountain vegetation; DEM; topography; human disturbance factors; CCA; ordinary kriging

1. Introduction

Remote sensing of vegetation in rugged and mountainous areas is severely affected by topographic effects [1,2]. Environmental factors such as topography, soil, water, and climate influence vegetation distribution in mountainous area [3–5]. Numerous authors have shown the relationships between ecosystem structure and composition and topographic features, such as elevation, slope angle, aspect and indices of relative moisture based on potential solar radiation and topographic redistribution of precipitation [6–9]. Topographic effects cause a difference in radiance values between inclined and horizontal surfaces. This causes the same land cover type on opposite facing slopes to have different spectral values and different land cover types to have similar spectral values [10].

Incorporating ancillary data has been proven to be efficient in reducing topographic effects and enhancing image classification accuracy. Digital Elevation Model (DEM) data are a commonly used type of ancillary data in mountainous areas. To improve image classification in mountainous terrain, DEM data have been used in four different ways:

➢ To reduce the topographic effect by topographic normalization techniques [1,11–21];
➢ As an additional "channel" increasing forest map accuracy [9,22–25];
➢ Combined with expert knowledge or a decision tree enhancing classification accuracy [26–28];
➢ Integrating the prior probability of the relationship between elevation and vegetation distributions improving image classification [29–34].

The spatial distribution pattern of mountainous vegetation is the result of complex interactions between natural processes and human influences [4,35]. Besides the topography as such human disturbance factors are also a key factor influencing distribution of vegetation types [36–39]. Identifying the factors controlling the distribution, abundance and diversity of species at local scale is one of the central objectives in community ecology [40,41]. Ordination and classification are the two main techniques for analyzing effects of environmental factors on plant community data [42]. Both techniques organize community data on species abundances independent of the habitat template [43]. These studies, however, merely examined the relationships between plant community or vegetation data and the environmental or topographic attributes. Interactions between the spatial components of the vegetation, such as spatial distribution pattern and the geographical distribution of the ecological and human disturbance factors have hardly been addressed [35]. Bio [44] recommended that the spatial

components of plant species response should be an integral part of the exploratory data analysis and mapping.

Remote sensing methods mainly focus on developing linkages between spectral response and surface features (such as vegetation types or land-use/land-cover types) at a broad range of spatial scales. Mapping plant communities by image classification however, is often limited by spectral similarities among essentially different communities [9,45,46]. To overcome this issue, environmental response models have been used to predict or produce potential vegetation maps in high mountainous areas [3–5,35,37]. However, environmental response models are insufficient to map actual vegetation patterns due to their limited capability to handle vegetation distribution prone to disturbance and perturbation [45,47,48]. In mountainous areas the integration between plant community data, human disturbance factors and geo-spatial data (DEM and image data), all featuring different spatial scales, is a particular challenge for vegetation mapping [35,43,45]. Hence there is an obvious need to establish the link between theories and practices of ecological research strategies on one side and remote sensing image analysis on the other side [43,49].

Zhang *et al.* [34] reported that incorporating the relationships between vegetation and environmental factors into image classification requires further investigation. This study aims to add relationships between vegetation spatial distribution pattern (SDP), and environmental/human disturbance factors to the remote sensing image classification process. Two mountainous study sites were selected that differ in terms of climate, ecology, geography, and population density. Selection of two study sites rather than one enables to establish whether the specific characteristics of these mountain zones are a determining factor in the success of the experimental approach.
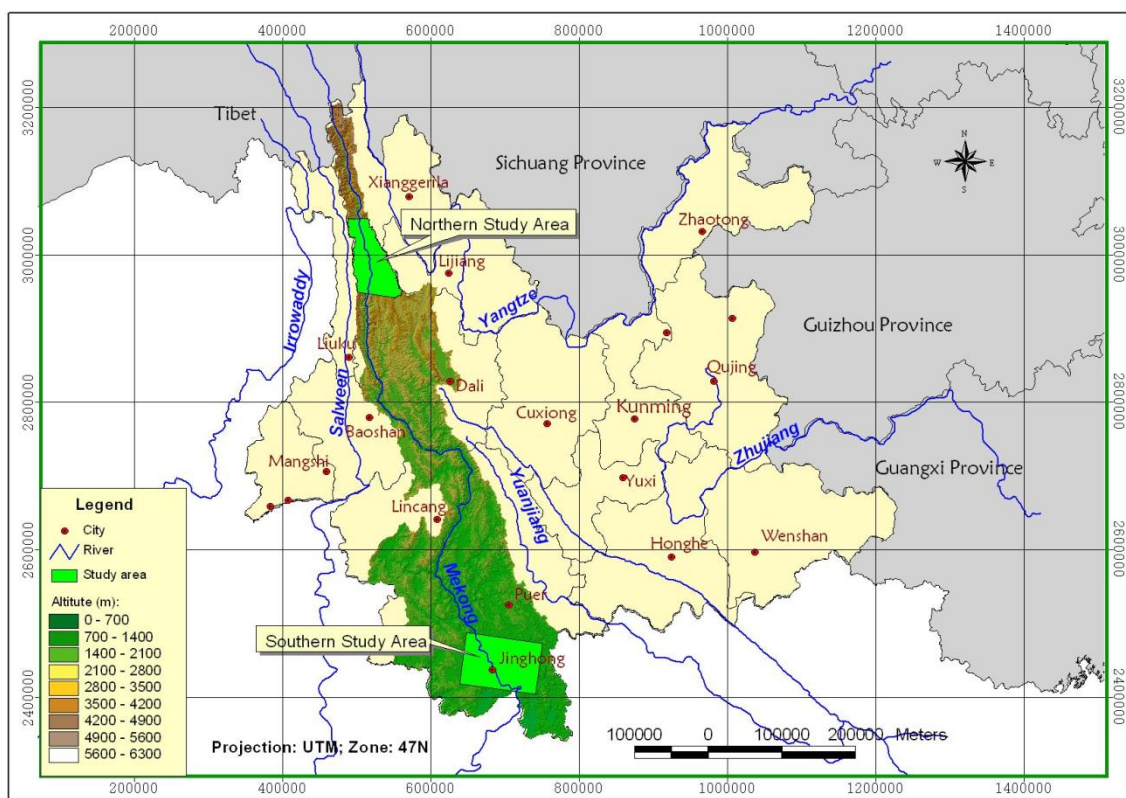
The objectives of this study are

➢ To quantify the relationships between vegetation SDP and topography and human disturbance attributes by using ordination analysis;
➢ To map mountain vegetation by integrating ordination models into remote sensing (Landsat Thematic Mapper) image analysis;
➢ To test the effectiveness of this mapping approach by evaluating its accuracy against two alternative classification strategies: (1) ordinary image classification without ancillary information (null model); and (2) classification with a DEM as extra input channel; and finally;
➢ To determine whether the significant differences of the two mountain zones are determinant in the outcomes of the experimental procedures.

2. Material and Methods

2.1. Study Area

The two research sites are situated in the Lancang (Mekong) watershed, Yunnan province, China (Figure 1). They are located in the northern and southern part of this watershed and feature large differences. The northern part is a very steep mountainous area marked by the presence of four of Southeast Asia's largest rivers: Yangtse, Mekong, Irrawaddy and Salween. The southern part of the watershed is a tropical region containing Xishuangbanna Dai Autonomous Prefecture and is one of the most important areas in China in terms of biodiversity [50].

Figure 1. Location of the two study areas.

The northern study area is located in the Weixi and Lanping counties and ranges from latitude: 26°7'12"N–27°53'24"N; longitude: 98°58'12"E–99°37'48"E (Figure 1). This study area covers the northern part of the Hengduan Mountain range of the eastern Himalayas and belongs to the World Heritage site of the Three Parallel Rivers of Yunnan Protected Area. The Hengduan Mountain region is one of the most biologically diverse temperate ecosystems on Earth [51–53]. It is important not only for biodiversity but also for its water resources. This area is a typical steep mountainous area, with altitudes ranging between 1,500 and 4,500 m. The yearly average temperature 11.3 °C, and average annual rainfall is 954 mm. Because of the altitude differences, the climate varies significantly with elevation. The climate can be divided into four zones from the bottom to the top of the mountain: warm temperate, temperate, cold temperate and sub-frigid [54]. This vertical climate zonation highly influences the vegetation spatial distribution pattern. In addition, many minority ethnic groups, such as Tibetans, Naxi, Bai, Lisu, are living in the northern study area [51].

The southern study area in the Xishuangbanna prefecture of China's Yunnan Province is located between 22°00'N–23°50'00"N and 100°00'12"E–102°00'E and covers about 50,000 ha (Figure 1). It belongs to the lower catchment of the Lancang River. Xishuangbanna is home to the richest biological and ethnic diversity (*i.e.*, Thai, Lahu, Jinuo, and Jingpo people) in China. The maximum altitude is about 2,000 m. The climate of the region is strongly seasonal. The monthly average temperature ranges between 16.4 °C and 22.0 °C. Between May and October the South-West Monsoon air masses brings about 80% of annual rainfall, whereas the dry and cold air of the Southern edges of the jet stream dominates the weather pattern between November and April. Annual rainfall varies between 1,200 mm in the Lancang valley and 1,900 mm at altitudes above 1,500 m.

2.2. Software

Canoco 4.5 was used to perform ordination. ArcView 3.2 and ArcGIS 9.1 were used to calculate topographic attributes and to perform ordinary kriging. Artificial Neural Networks (ANN) classification was conducted with LNNS, a neural network simulation program developed at the Laboratory of Forest Management and Spatial Information Techniques (FORSIT), Ghent University, Belgium. LNNS is freely available at http://dfwm.ugent.be/forsit [55]. Idrisi Andes was used to assess accuracy comparison between different classification strategies.

2.3. Data

2.3.1. Image Data

Each study area is covered by a 30 m resolution Landsat Thematic Mapper 5 (TM) image (Table 1). The available Landsat images are third level products of the China Remote Sensing Satellite Ground Station (http://www.rsgs.ac.cn/English/Standard%20Product.htm). Basic radiometric and geometric preprocessing was done by the image provider. Gitas and Devereux (2006) [56] mentioned that topographic correction should be regarded as an essential element of any classification methodology. We compared six different topographic normalization methods (Cosine correction, Minnaert correction, C-correction, SCS correction, two-stage topographic normalization, and slope matching technique) for their effectiveness in enhancing vegetation classification in mountainous environments. However, none of these topographic correction methods could significantly improve overall classification accuracy (Zhang *et al.*, 2011) [57]. Therefore, in this study, none of the topographic correction techniques was applied.

Table 1. Satellite image data and number of training classes and subclasses.

Study Areas	Path/Row	Date	Sensor	Number of Classes	Number of Subclasses
Northern study area	132-41	26 December 2003	Thematic Mapper 5	10	41
Southern study area	130-45	1 March 2004	Thematic Mapper 5	10	46

2.3.2. Vegetation Data

Different land cover classes including vegetation and non-vegetation classes (*i.e.*, water, burned land, and cast shadow) were discerned (Tables 2 and 3). In view of this study, their selection was based on a twofold rationale. The first relates to the major vegetation communities in both study areas, while the second is driven by the vegetation classes that can be discerned from Landsat TM imagery.

2.3.3. DEM, Topographic Maps and Derived Attributes

The 25 m resolution DEM data produced by the China State Bureau of Surveying and Mapping was made available for this study (http://www.sbsm.gov.cn/). The DEM was re-sampled to 30 m and projected into the Universal Transverse Mercator (UTM) projection system, zone 47, WGS84 (World Geodetic System 1984). In this study, topographic maps produced in the 1970s by Chinese military at a scale of 1:100,000 were used providing useful information with respect to villages, towns, rivers, roads and land cover.

Table 2. The dominant vegetation classes and dominant species in the northern study area.

Code	Classes	Dominant Species
FSF	Fir and spruce forest	*Picea likiangensis, Picea brachytyla, Abies georgei, Abies georgei var. smithii, Abies forrestii, Abies ferreana.*
PF	Pine forest	*Pinus yunnanensis, Pinus densata, Pinus armandi, Corylus yunnanensis*
OF	Oak forest	*Quercus aquifolioides, Quercus gilliana, Quercus pannosa.*
MF	Mixed forests	*Pinus yunnanensis, Pinus armandii, Quercus aquifolioides, Quercus gilliana, Betula utilis, Acer cappadocicum, Acer davidii, Picea likiangensis, Picea brachytyla, Sorbus* sp., *Abies georgei, Pseudotsuga forrestii, Tsuga* sp., *Abies georgei var. smithii, Abies forrestii.*
LDF	Low density forest and tall shrubs	*Pinus yunnanensis, Salix myrtillacea, Corylus yunnanensis.*
DSM	Dwarf shrub and meadow	*Elsholtzia capituligera, Incarvillea arguta, Bauhinia brachycarpa, Rhododendron tapetiforme, Rhododendron telmateium, Rhododendron phaeochrysum, Salix annulifera, Salix hirticaulis, Salix myrtillacea, Vitex microphylla, Potentilla* sp., *Polystichum* sp., *Juncus* sp., *Carex* sp., *Poa* sp., *Plantago* sp., *Heleocharis yokoscensis, Polygonum lapathifolium, Polygonum calostachyum.*
AL	Agricultural land	*Juglans* sp., *Zea mays* L.
SN	Snow	
WT	Water	
CS	Cast shadow	

Table 3. The dominant vegetation classes and dominant species in the southern study area.

Code	Classes	Dominant Species
ORT	Old rubber trees	*Hevea brasiliensis*
YRT	Young rubber trees	*Hevea brasiliensis*
EF	Evergreen forest	*Cyclobalanopsis delavayi, Castanopsis hystrix, Castanopsis mekongensis, Lithocarpus truncatus, Litsea glutinosa, Actinodaphne henryi, Schima wallichii, Syzygium yunnanensis, Elaeocarpus austro-yunnanensis, Paramichelia baillonii, Engelhardtia* sp. *Machilus salicina, Symplocos cochinchinensis, Olea rosea, Aporusa* sp. *Pinus khasya* var. *langbianensis, Lithocarpus* sp. *Quercus dentata, Betula alnoides, Quercus acutissima, Cyclobalanopsis kerrii, Quercus variabilis.*
LDF	Low density forest and tall shrubs	*Pinus khasya* var. *langbianensis, Pyrus pashia, Phoebe minutiflora, Myrica esculenta, Colona floribunda* and *Vaccinium bracteatum.*
DF	Deciduous forest	*Ficus altissima, Toona sinensis, Nephelium chryseum, Altingia excelsa, Bischofia javanica, Colona floribunda, Bombax ceiba, Erythrina stricta, Bauhinia variegata, Dendrocalamus strictus, D. brandisii, Cephalostachyum pergracile, Indosasa sinica, Schizostachyum funghomii,* and *Dinochloa puberula.*
SGL	Shrub and grass land	*Trema orientalis, Dalbergia obtusifolia, Docynia indica, Eurya groffii, Saccharum sinense, Leucosceptrum canum, Eupatorium coelestinum.*
AL	Agricultural land	
BL	Burned land	
WT	Water	
CS	Cloud and shadow	

The topographic maps can provide land cover information because the plant community's types did not change notably, especially in the remote mountainous areas. Moreover, human disturbance factors, such as road (small roads and trails), villages, rivers, and towns can easily be derived from topographic maps by digitization.

To analyze the relationship between topographic attributes and vegetation spatial distribution pattern, eight topographic attributes were derived from the DEM data. The applied topographic attributes can be divided into primary and secondary attributes. Primary attributes include elevation, slope, aspect, plan and profile curvature, and topography position index, and are calculated directly from the derivatives of a topographic surface [58]. The secondary attributes, computed from two or more primary attributes, include a solar and a wetness index. No erosion index could be computed as the rainfall data was not available. Details of the computed attributes are listed in Table 4 [59–65]. Table 3 lists the dominant vegetation classes and dominant species in the southern study area.

Table 4. Short description of the selected topographic factors.

Code	Factors	Descriptions
ELEV	Elevation	Elevation is one of most important topographic factors in regulating mountain vegetation patterns [39,59,60].
SLO	Slope	Slope also is one of important topographic factors for mountain vegetation patterns because it will influence features such as soil moisture, wind, and solar radiation [37,59–61].
ASP	Aspect	Vegetation spatial distribution can be affected by slope aspect [3,4,59].
PRF	Slope profile curvature	This index measures the rate of change of potential gradient and hence is important for characterizing changes in flow velocity and sediment transport processes [58]. It also potentially indicates soil moisture [37].
PLF	Planiform curvature	This index is related to converging/diverging flow and soil water content [62].
TPI	Topographic position index	Topographic position index is the basis of the topography classification system and is simply the difference between a cell elevation value and the average elevation of the neighborhood around that cell [63]. This index can affect the vegetation patterns in mountainous areas [4,59,64].
CTI	Compound topographic index	CTI is a steady state wetness index [62]. Wetness index has been shown to affect vegetation spatial patterns [3,4].
PADIR	Potential annual direct incident radiation	PADIR is a solar index, and was developed by McCune and Keon [65]. Solar radiation is the primary atmospheric control over soil moisture status between precipitation events in vegetation not receiving melt water and appears to influence the local adaptation of vegetation [3].

2.3.4. Human Disturbance Factors

Forestry, agriculture, economic development, transport infrastructure are some of the disturbance drivers that influence vegetation spatial distribution pattern and landscape pattern at large regional scales [38]. Numerous studies have shown that population density and proximate factors are important driving forces for both land cover and vegetation pattern change as well as for deforestation [66–69]. As it pertains to the distribution of the population over a region, the surface density of villages was computed through GIS analysis [70]. Secondly, proximate factors such as distance to settlements and

markets (villages and towns) were considered [66,68]. In rural areas in China, farmland and rural settlements are the most important land use classes [71]. Most of the farming activities tend to cluster into the villages [72]. In addition, mining, road and dam construction are important human interventions in the landscapes of both study areas. These activities are directly related to the occurrence of roads and rivers. Distance to roads and rivers reflect the landscape accessibility. Serneels and Lambin [66], and Overmars and Verburg [68], demonstrated that land use and land cover change was strongly related to distance to roads and rivers, while Nagendra *et al.* [73] claimed that accessible areas featured a higher degree of deforestation and fragmentation. Therefore and thirdly, distance to roads and rivers were included as human disturbance factors in the current analysis (Table 4). Most of the digital line (rivers and roads) and point (villages and towns) data were derived from the topographic maps. Because the topographic maps date back to the 1970s, an update, especially for the road database, was required. This was accomplished using the traffic map of Yunnan published in 2007 (http://www.ynbsm.gov.cn/emapshow.asp?col=ynly). The roads were divided into two categories: wide (width ≥ 3 m) and narrow (width < 3 m) roads. The rivers were categorized into three levels based on the standards of China's river system (http://nfgis.nsdi.gov.cn/nfgis/english/default.htm) (Table 5).

Table 5. Human disturbance factors.

Code	Proximate Drivers	Unit
SDV	Surface density of villages	km^2
CDV	Cost distance to villages	m
CDT	Cost distance to towns	m
CDLC	Cost distance to Lancang River	m
CDMR	Cost distance to mid-class level river	m
CDS	Cost distance to streams	m
CDLR	Cost distance to large (wide) roads	m
CDSR	Cost distance to small (narrow) roads	m

Distance is a basic concept inherent to any geographical space. Euclidian distance, the simplest distance measurement, is widely used to calculate the distance of a location to the nearest destination of interest [66,68,74,75]. However, Euclidean distance is computed by using the Pythagorean Theorem, which is not very realistic when representing effective spatial accessibility [76]. Typically in a mountainous environment, the distance between two points over a sloping surface is significantly larger than over a flat path. Another critical issue is that of friction which impedes movement [77]. Several researchers have shown the utility of modeling cost distance that takes into account the spatial heterogeneity [76,78,79]. Hence, we measured the cost distance of a cell to the destinations of interest taking into account the existing friction based on slope angles (Table 5). There are a number of ways to assign friction coefficients to slopes. Here, an empirically binomial equation (Equation (1)), which was developed from sample data of walking times in mountains in the Middle Himalaya of Nepal, was used to model friction magnitudes for slope [77]:

$$Y = 0.031 \times X^2 - 0.025X + 1 \tag{1}$$

where *Y* is friction and *X* is the slope derived from DEM. Afterward, the cost distance is measured as the least cost (in terms of effort, expense, *etc.*) when moving over a friction surface [80].
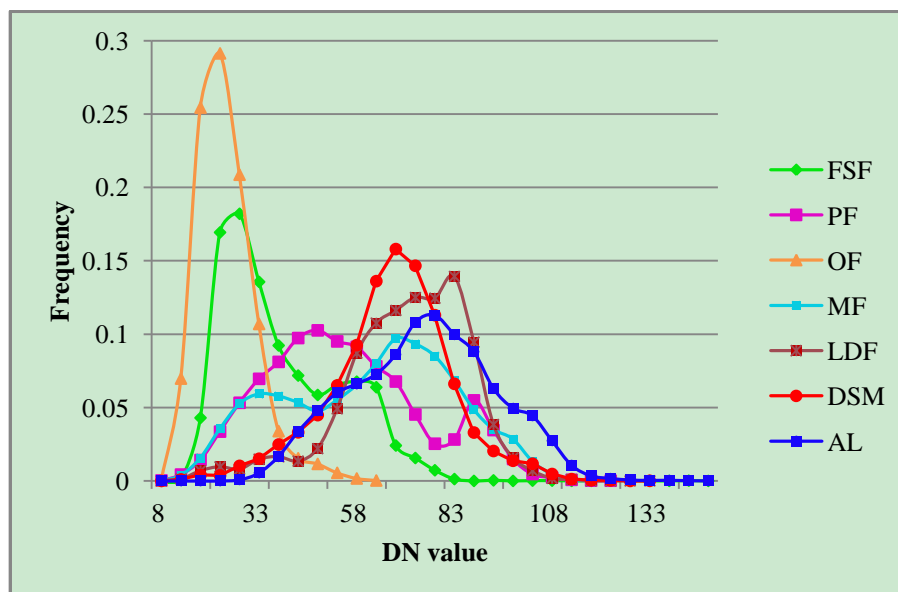
2.3.5. Training and Validation Data

Three field surveys (September to October 2004 in the northern study area; March 2004 and September to October 2006 in the southern study area) were carried out. During field surveys, at several spots that were thought to be of a relevant size with regard to the spatial resolution of the available images, the present vegetation type, as well as a GPS measurement of the co-ordinates and elevation, were recorded. For each point, we also recorded the dominant vegetation type (Tables 2 and 3), the cover percentage and estimated height of vegetation types. In addition, human impact, such as harvest residue, fire, logging or grazing was recorded. Other features that were recorded, include vegetation groups (e.g., coniferous or broad-leaved, evergreen or deciduous, *etc.*), vertical structure of plant communities (e.g., subcanopy forest regeneration), and succession stages (such as weeds, pioneers, or climax vegetation types). They were used to establish the main characteristics and variability of the vegetation cover, and to acquire reliable field data for both training the classifiers and evaluating the accuracy of the classification results. Furthermore, at each point several photographs were taken in different directions to ensure the possibility of future comparison after the field work was completed. The field data were supplemented with manually digitized reference pixels. As a specific vegetation class might consist of several spectral subclasses, training data for each different spectral subclass was assembled in a ground truth digitizing stage [81]. There exists, for example, "pine forest on shaded slopes", "pine forest on sunny slopes", "old rubber trees" and "young rubber trees". The number of spectral subclasses of the two images is listed in Table 1. The ground truth polygons were converted into a raster map with a cell size of 30 m × 30 m. Then, this map was randomly divided into training set, test set, and validation set be used in an artificial neural networks classification protocol. The test set was used to evaluate the algorithm's convergence while the validation data served for independent accuracy assessment. The number of pixels of the training sets and validation sets for each class is listed in Table 6. Moreover, in order to measure the topographic effects on spectral patterns of vegetation classes, histograms of digital number (DN) values of vegetation classes were examined (Figure 2).

Table 6. The number of pixels of the training and validation sets for each class in the both study areas.

The Northern Study Area			The Southern Study Area		
Code of Classes	Number of Training Pixels	Number of Validation Pixels	Code of Classes	Number of Training Pixels	Number of Validation Data
FSF	2,779	2,736	ORT	3,302	3,321
PF	9,540	9,594	YRT	1,727	1,695
OF	1,813	1,815	EF	23,866	23,981
MF	4,539	4,423	LDF	2,386	2,383
LDF	6,319	6,230	DF	2,472	2,439
DSM	6,898	6,933	SGL	4,838	4,710
AL	23,851	23,800	AL	9,310	9,408
SN	3,912	3,777	BL	1,715	1,623
WT	1,138	1,176	WT	1,217	1,250
CS	2,327	2,383	CS	241	267

Figure 2. The frequency distribution of digital number (DN) values of the training data in the Thematic Mapper near infrared (NIR) band (band 4). FSF: fir and spruce forest; PF: Pine forest; OF: oak forest; MF: mixed forest; LDF: low density forest and tall shrub; DSM: dwarf shrub and meadow; AL: agricultural land.

2.3.6. Classification Algorithm Selection

Since the beginning of the 1990s, artificial neural networks (ANN) have been applied in the analysis of remote sensing images [82]. Many authors have reported considerable advantages of ANN over conventional methods (e.g., Verbeke *et al.* [55], Mas and Flores [83]). One of the main advantages of ANN classifiers is that they are independent of the distribution of the class-specific data in feature space [84]. Figure 2 clearly shows that the DN distributions of the ground truth data of some classes (*i.e.*, fir and spruce forest, pine forest, and mixed forest) deviate from a Gaussian model. Considering the non-normality and heteroscedasticity of the ground truth data ANNs were used to perform image classification. After classification, the subclasses were merged into 10 broad land cover classes based upon ancillary information (GPS points, topographic maps, DEM data, and plant community data) and field ecological expertise (Tables 2 and 3). In addition, in order to build a CCA ordination model, a spatial database including vegetation data as well as topographic and human disturbance attributes was built.

2.4. Methods

2.4.1. Outline

The mountain vegetation mapping approach of this study contains three main stages: CCA ordination analysis (Stage 1), interpolation of CCA axes (Stage 2), and image classification incorporating ordination results or DEM data (Stage 3). In Stage 1, the relationship between topography and human disturbance factors and vegetation data was identified. In Stage 2, the CCA scores for unobserved locations were predicted based on the CCA scores at the sampled locations

using ordinary kriging. Next, in Stage 3 an ANN was applied classifying the TM image data incorporating ancillary information from the 4 CCA axes or from the original DEM. Finally, classification accuracy was assessed.

Stage 1: Ordination Analysis

Ordination methods can be divided in indirect ordination or unconstrained ordination (*i.e.*, detrended correspondence analysis, DCA and principal component analysis, PCA) and direct ordination or constrained ordination (*i.e.*, canonical correspondence analysis, CCA and redundancy analysis, RDA) [85]. The direct ordination methods are particularly suited to address relationships between species composition and environmental variables. Direct ordination methods have also been applied to study the relationships between vegetation SDP and environmental factors, especially topographic factors [61,64,86–88]. In this study, the spatial distribution of the known pixels was derived from our ground truth data which are homogeneous. In order to build CCA ordination model, a spatial database including vegetation data and topographic and human disturbance attributes was built. This spatial database included two matrices: vegetation type matrix (to each ground truth polygon a vegetation type was assigned) and topography and human disturbance attributes matrix (to each ground truth polygon, the average value of topography and human disturbance attributes listed in Tables 4 and 5 were assigned).

DCA was used in a preliminary test to assess the length of the environmental gradients of the axes [85]. If that value was larger than 4.0, unimodal methods (DCA, CA, or CCA) was used. On the other hand, if the longest gradient was shorter than 3.0, the linear method (PCA or RDA) was considered to be a better choice [89]. In the range between 3 and 4, both types of ordination methods are assumed to be applicable.

In both study areas, the lengths of the environmental gradients are larger than 4 (Table 7); the CCA ordination procedure was carried out to investigate relationships between topographic attributes and vegetation spatial distribution pattern at both study areas. The core of direct constrained ordination by CCA developed by Ter Braak [90] is a weighted multivariate linear regression of community gradients on environmental variables [91]. First, the main variation in the species (or vegetation) data is analyzed by indirect gradient ordination (e.g., correspondence analysis, CA); Secondly, the indirect gradient ordination axis is used to relate to the environmental variables by multiple regression of the site scores on the environmental variables [90]. A joint plot (biplot) showed the general vegetation types and their correlations with topographic and human disturbance variables. The relationships between vegetation data and topographic/human disturbance variables can be estimated by perpendicularly projecting vegetation points (triangles) onto the arrows (topographic and human disturbance variables). The lengths of the arrows are used to compare the size of such an effect across the topographic and human disturbance variables. The approximated correlation between topographic and human disturbance variables is equal to the cosine of the angle between the corresponding arrows (Leps and Smilauer, 2003) [85]. The significance of the relationships between vegetation variables and topographic/human disturbance variables was determined by Monte Carlo permutation tests (Legendre and Legendre, 1998) [92]. In addition, the intraset correlations between topography and human disturbance variables and canonical axes in CCA were used as indicators of which variables were more important in structuring

the ordination (Ter Braak, 1986) [90]. Therefore, a multiple linear regression was used to address those intraset correlations.

Table 7. Summary table of DCA applied to vegetation training data.

Study Areas		DCA1	DCA2	DCA3	DCA4	*Dispersion of All EV*
Northern study area	Eigenvalues (EV)	1.000	1.000	0.826	0.067	9.000
	Lengths of gradient	0.000	0.000	4.877	4.151	
Southern study area	Eigenvalues (EV)	1.000	0.842	0.444	0.077	9.000
	Lengths of gradient	0.000	6.501	6.087	5.619	

Stage 2: Spatial Interpolation of CCA Axes Scores

CCA analysis reduced the vegetation data and environmental (topography and human disturbance) variables into four independent ordination axes. The four CCA axes explained the vegetation and environment data by recording the dispersion of vegetation and environmental (topography and human disturbance) variables scores. The spatial interpolation of 4 CCA environment axes scores at the sampled locations was carried out by using kriging [35,91,93,94]. This geostatistical approach assumes that the spatial correlation structure between points (samples), which are closer together have more similar values than those that are further apart [95]. Firstly, semivariogram analysis was used to explore the spatial correlation of the four CCA scores (Figure 3); Secondly, to assess spatial correlation, a model was fitted to the semivariogram with spherical, exponential, or Gaussian functions as candidate models resulting in estimated prediction standard errors. If the prediction standard errors are valid the root-mean-square standardized errors (RMSSE) should be close to 1 [95]. From the calculation of the RMSSE a spherical model was used to fit the semivariogram of axis1, axis3, and axis4, while the exponential model was used to fit the semivariogram of axis2 [90].

Afterward, ordinary kriging was used to predict CCA scores at unobserved locations (Burrough and McDonnel 1998) [93]. It is the most robust and often used method because it minimizes the variance error between the model and the estimate [96]. The kriging interpolation is a weighted average of the observed values Z_{x_i} which is used to estimate the value of Z_{x_0}, identified at a specific location x_0 where there are no measured values [93,95,97]. By using the spatial structure defined by the semivariogram, a kriging system of linear equations combining neighboring information can be defined [98].

Stage 3: Image Classification

There are many different types of neural networks for image classification [99]. In this study, a three layer neural network was constructed, consisting of one input layer, one hidden and one output layer [100]. The number of input neurons corresponded to the number of input layers (image bands complemented with DEM or CCA axes). In order to be consistent with the DN value range of the TM image data, all of the extra channels (DEM or CCA axes) were normalized on a 0 to 255 value scale. The number of neurons in the output layer equaled the number of subclasses while the number of the hidden neurons was arbitrary set to twenty. A sigmoid activation function was used with a learning rate of 0.001 and momentum of 0.2.

Figure 3. Fitted variogram for each extracted canonical correspondence analysis (CCA) ordination axis in the northern study area. (**a**) Semivariance diagram of CCA scores for axis 1; (**b**) Semivariance diagram of CCA scores for axis 2; (**c**) Semivariance diagram of CCA scores for axis 3; (**d**) Semivariance diagram of CCA scores for axis 4.

(**a**)

(**b**)

(**c**)

(**d**)
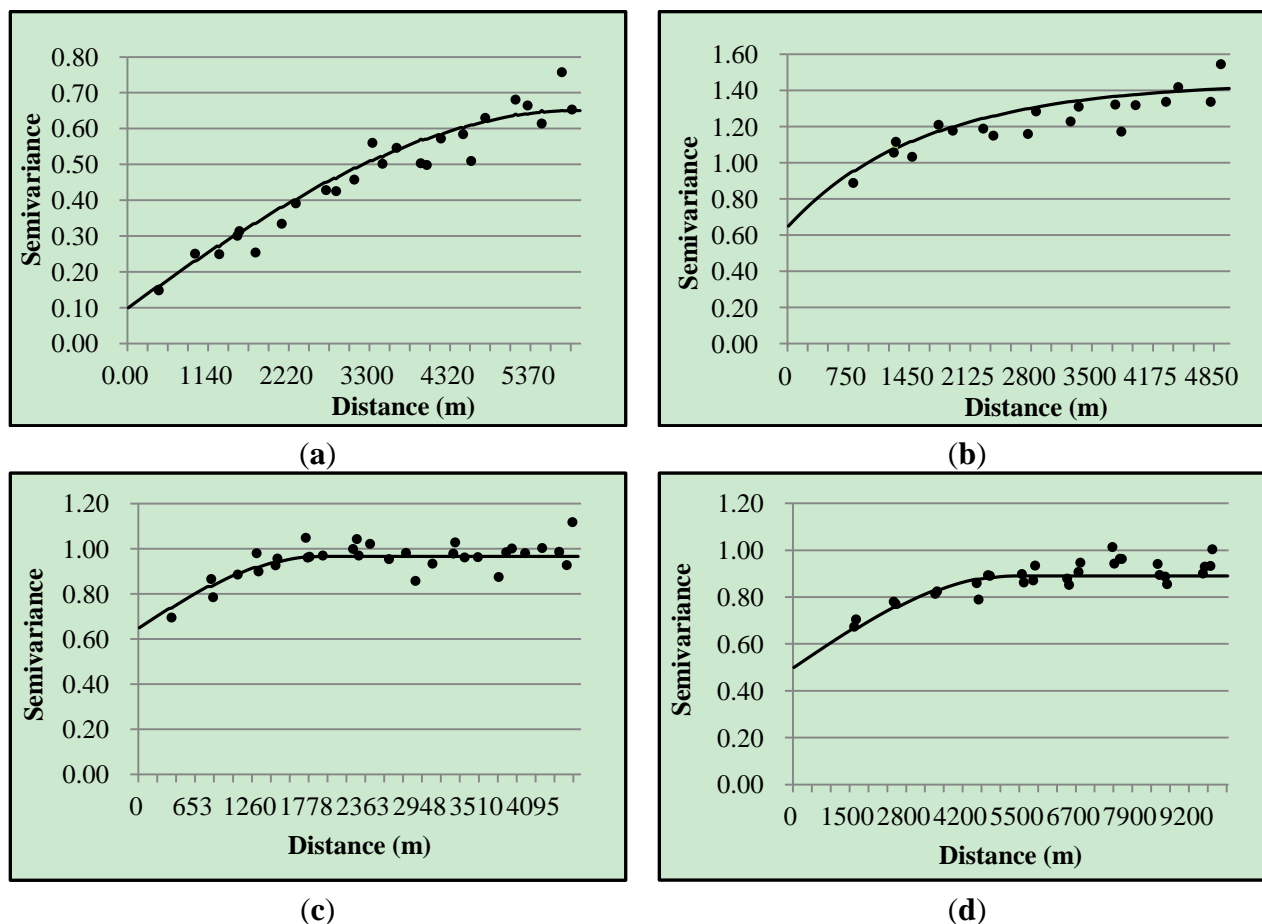
2.4.2. Accuracy Assessment

The accuracy of the classifications was assessed using confusion or error matrices and Kappa statistics (KHAT) [101]. In addition, in order to test the null hypothesis of no differences between different images classification strategies, the pair-wise Z-statistic was calculated [102–104].

3. Results

3.1. Ordination Results

From the Monte Carlo permutation test, except for planiform curvature (PLF) and cost distance to streams (CDS), all topographic and human disturbance variables were significantly correlated with the vegetation and non-vegetation spatial distribution pattern data in the northern study areas (Table 8). In the southern study area, topographic position index (TPI), Slope planiform curvature (PLF), and slope profile curvature (PRF) were not significantly correlated with the vegetation data (Table 8). The results of CCA are shown in two diagrams for the two study areas (Figures 4 and 5). Vegetation classes are represented by triangles, and topography and human disturbance factors are represented by

arrows. The length of arrows indicate that elevation (ELEV) and surface density of villages (SDV) are the two most important driving factors determining vegetation spatial distribution pattern in both study areas (Figures 4 and 5). Profile (PRF) and planiform (PLF) curvature have relatively small effects on the vegetation spatial distribution pattern.

Table 8. Statistical significance of topography and human disturbance variables for vegetation spatial patterns in the two study areas.

Code	Variables	Northern Study Area (n = 2,085)		Southern Study Area (n = 2,986)	
		p-Level	F Value	*p*-Level	F Value
ELEV	Elevation	0.002	21.88	0.002	13.05
SDV	Surface density of villages	0.002	3.52	0.002	12.84
CTI	Compound topographic index	0.002	18.09	0.002	8.29
CDLC	Cost distance to Lancang River	0.002	15.23	0.018	2.49
PADIR	Potential annual direct incident radiation	0.002	17.16	0.002	4.95
SLO	Slope angle	0.002	4.93	0.002	4.41
CDLR	Cost distance to large roads	0.002	5.05	0.002	12.60
CDMR	Cost distance to mid-class level river	0.002	7.21	0.002	4.81
CDV	Cost distance to villages	0.018	2.51	0.002	5.52
CDT	Cost distance to towns	0.002	6.49	0.002	5.80
CDSR	Cost distance to small roads	0.002	2.71	0.002	4.22
TPI	Topographic position index	0.01	2.81	0.108	1.59
CDS	Cost distance to streams	0.036	2.11	0.002	3.97
PRF	Slope profile curvature	0.094	1.63	0.302	1.15
ASP	Slope aspect	0.002	3.92	0.002	6.66
PLF	Slope planiform curvature	0.846	0.49	0.762	0.62

Note: The underlined numbers are cases with *p* > 0.05.

CCA diagrams also illustrate the relationships between each vegetation class and the set of environmental variables. Compared to the other land cover classes, snow (SN) and fir and spruce forest (FSF) are predicted to occur at higher elevation and are located relatively far from villages, towns, the Lancang River, and roads (Figure 4). Conversely, these two classes have lower surface density of village (SDV) values implying that, in the northern study area, snow (SN) and fir and spruce forest (FSF) occur at higher altitude with less human disturbance. Water (WT), pine forest (PF) and agricultural land (AL) all feature higher surface density of village (SDV) values and lower elevation and proximate values. Based on field knowledge, water (WT) (Lancang River), agricultural land (AL) and pine forest (PF) always occur at lower elevations where most of villages and roads are present. Water (WT) also has the highest compound topographic index (CTI) value and lowest slope degree. In contrast, the cast shadow class (CS) has the highest slope value. Figure 4 shows that dwarf shrub and meadow (DSM) is close to the center. It means that the dwarf shrub and meadow is weakly correlated to elevation and other factors. It can be found at any elevation and anywhere in the study area. It includes both natural vegetation (alpine and subalpine shrub/meadow), and vegetation that is featuring anthropogenic disturbance (grazing land and abandoned farm land). Mixed forest is also close to the center. It also includes both natural vegetation (oak and spruce mixed forest),

and vegetation that is featuring anthropogenic disturbance (secondary mixed forest such as pine and poplar mixed forest).

Figure 4. CCA ordination of vegetation classes' spatial pattern in relation to environmental variables in the northern study area. The symbols of vegetation classes are listed in Table 4. The symbols of environmental variables are described in Tables 2 and 3. By projecting the points of the vegetation classes (triangles) onto the arrows of quantitative environmental variable, an approximate ordering of those vegetation classes with respect to that environmental variable can be obtained.
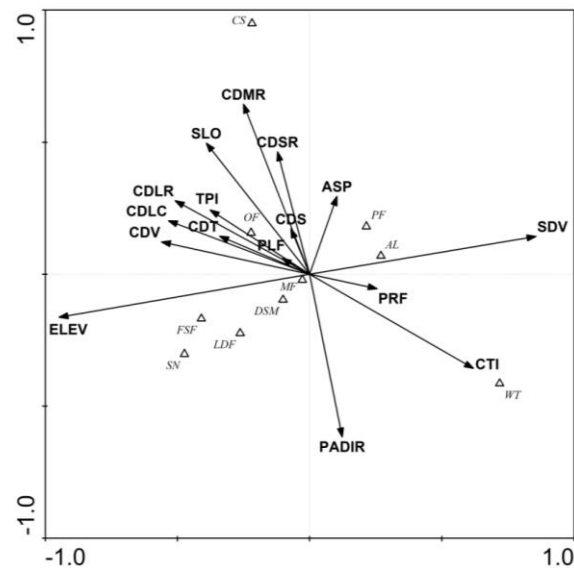
Figure 5. CCA ordination of vegetation classes' spatial pattern in relation to environmental variables in southern study area. The symbols of vegetation classes are listed in Table 4. The symbols of environmental variables are described in Tables 2 and 3. By projecting the points of the vegetation classes (triangles) onto the arrows of quantitative environmental variable, an approximate ordering of those vegetation classes with respect to that environmental variable can be obtained.
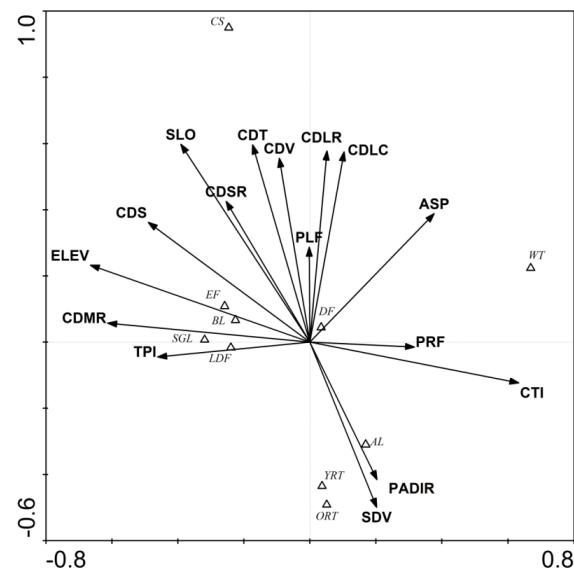
In the southern study area, rubber plantation (YRT and ORT) and agricultural land (AL) are predicted to have a lower elevation and distance to towns (CDT), villages (CDV), roads (CDSR and CDLR), and rivers (CDS, CDMR, and CDLC) (Figure 5). However, these classes have higher values for surface density of villages (SDV) and potential annual direct incident radiation (PADIR). In contrast, evergreen forest (EF), low density forest (LDF), shrub and grass land (SGL), and burned land (BL) have higher elevation (ELEV) values (Figure 4). From field expertise it is known that the two anthropogenic classes, rubber plantation (YRT and ORT) and agricultural land (AL), always occurs at lower altitudes and close to towns and villages.

The importance of each ordination axis with respect to the original vegetation pattern is indicated by the eigenvalues. For the northern study area, all summed CCA axes explained about 41% of the vegetation data variability and about 32% of the variation of the vegetation data was explained by the four extracted axes. On the other hand, for the southern study area, about 37% of vegetation data variability was assessed by all CCA axes, and the four extracted axes explained about 28%. The eigenvalues showed that the explained variability in the vegetation data is for the larger part expressed by the first two dimensions.

We compared the partial CCA ordination results for both study areas. In the northern zone the topographic attributes explained about 28% (17.76% + 10.82%) of the total variability of the vegetation and non-vegetation data, while human disturbance factors accounted for considerably less (23.65% = 12.83% + 10.82%) of the variance (Table 9). However, in the southern study area, the variance explained by human disturbance factors (22.86% = 15.63% + 7.23%) was almost equal to the variance explained by topographic attributes (21.94% = 14.71% + 7.23%).

Table 9. Decomposed explained variance in vegetation data.

Study Areas	Variance Explained	Partially Explained by Topographic Attributes	Partially Explained by Human Disturbance Attributes	The Shared Explained Variance
Northern study area	41.41%	17.76%	12.83%	10.82%
Southern study area	37.57%	14.71%	15.63%	7.23%

3.2. Image Classification

The poorest ANN classification accuracy was obtained using the 7 TM image bands training 10 broad classes (Tables 10 and 11). For both study areas, the highest classification accuracies were obtained when classifying all spectral subclasses (41 and 46 for the northern and southern study area respectively) incorporating the 4 CCA axes as extra channels (Tables 10 and 11) (Figure 6).

Pair-wise Z test values are listed in Tables 12 and 13. These values show that classification accuracy can be significantly improved by training spectral classes or subclasses as compared to training broad classes [81,105].

Table 10. Summary of the TM image (2 December 2003) classification results in the northern study area.

Classification Number	Classification and Training Strategies	Data Used	Overall Accuracy (OA) ($n = 62,867$)	Kappa ($n = 62,867$)
N1	ANN by training 10 classes	7 bands	83.69%	0.7996
N2	ANN Training 41 subclasses	7 bands	91.09%	0.8878
N3	ANN by training 41 subclasses	7 bands, DEM	95.3%	0.9412
N4	ANN by training 41 subclasses	7 bands, 4 CCA interpolated axes	96.49%	0.9561

Note: *n* is the number of pixels. ANN, Artificial Neural Network. DEM, Digital Elevation Model.

Table 11. Summary of the TM image (1 March 2004) classification results in the southern study area.

Classification Number	Classification and Training Strategies	Data Used	Overall Accuracy (OA) ($n = 62,867$)	Kappa ($n = 62,867$)
S1	ANN by training 10 classes	7 bands	85.97%	0.8021
S2	ANN by training 46 subclasses	7 bands	90.1%	0.8631
S3	ANN by training 46 subclasses	7 bands, DEM	93.49%	0.9103
S4	ANN by training 46 subclasses	7 bands, 4 CCA interpolated axes	96.45%	0.9500

Note: *n* is the number of pixels.

Table 12. Pair-wise Z statistic test of the comparisons of the different classification strategies in the northern study area.

| Pair-Wise	N1	N2	N3	N4
N1				
N2	−5.64 ***			
N3	−7.04 ***	−2.42 *		
N4	−6.87 ***	−2.79 **	−0.54	

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Table 13. Pair-wise Z statistic test of the comparisons of the different classification strategies in the southern study area.

Pair-Wise	S1	S2	S3	S4
S1				
S2	−3.53 ***			
S3	−5.40 ***	−2.20 *		
S4	−8.23 ***	−4.45 ***	−2.03 *	

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Figure 6. Land cover maps derived from ANN image classification incorporating the four CCA axes as extra channels in both study areas. (**a**) Land cover image classification in the northern study area; (**b**) Land cover image classification in the southern study area.

(**a**)

(**b**)
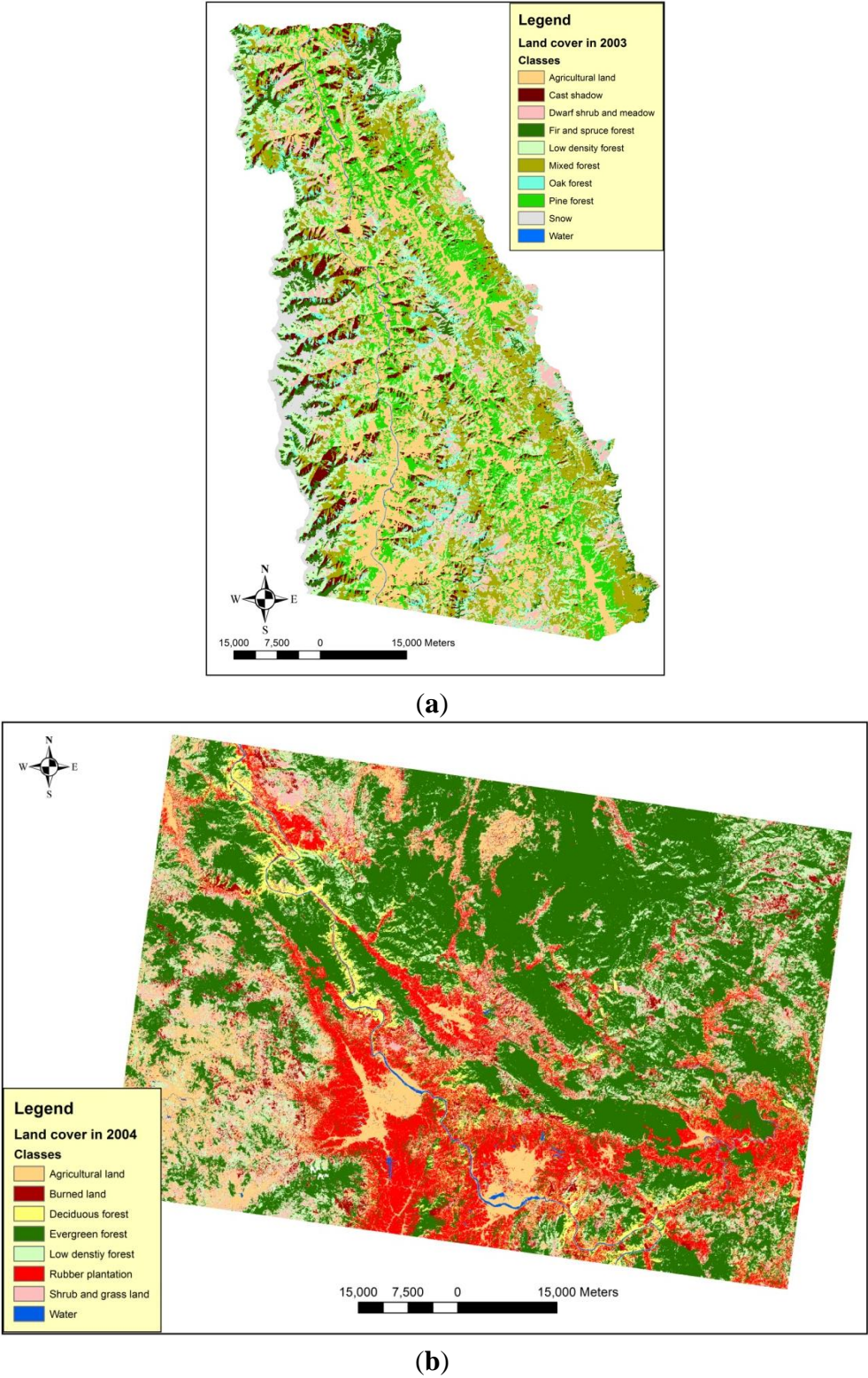
4. Discussion

In steep mountainous terrain, topographic effects may cause the same vegetation type on opposite facing slopes to feature different spectral values and different vegetation types to have similar spectral values [106]. Figure 7 shows a bimodal distribution of the pine forest (PF) spectral response patterns (training data) indicating the presence of two pine subclasses, pine on sunny slopes and pine on shady slopes.

Figure 7. Distribution of DN values for pine forest training data in Landsat 5 TM NIR (band 4).
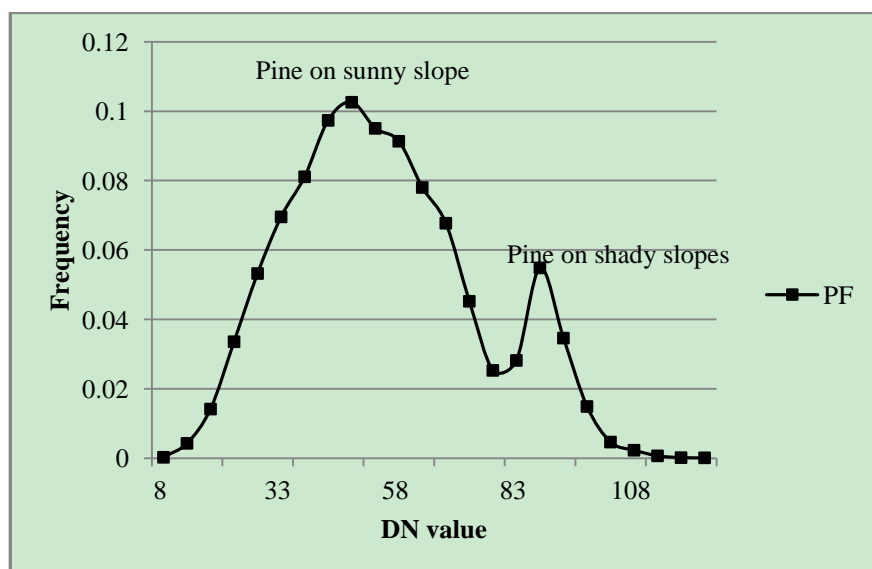
Table 14. Pearson correlations of CCA axes with topographic and human disturbance attributes.
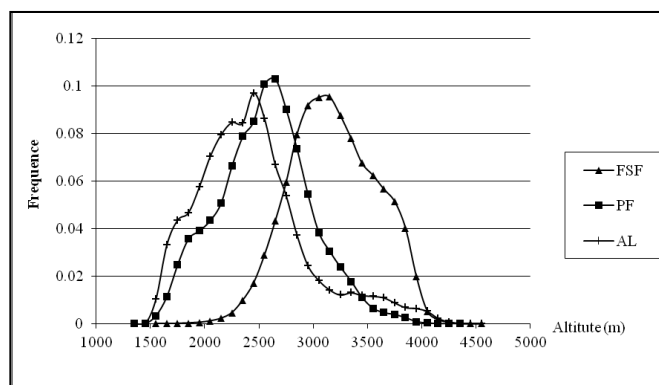
Code	Northern Study Area				Southern Study Area			
	Axis1	Axis2	Axis3	Axis4	Axis1	Axis2	Axis3	Axis4
ELEV	−0.95 ***	−0.16	0.14	0.14	−0.67 ***	0.23 *	−0.52 ***	0.14
SLO	−0.39 ***	0.50 ***	−0.47 ***	0.10	−0.39 ***	0.60 ***	0.49 ***	0.19
ASP	0.10	0.29 **	0.09	0.01	0.38 ***	0.39 ***	−0.05	−0.17
TPI	−0.38 ***	0.24 *	−0.23 *	−0.12	0.63 ***	−0.12	−0.28 **	−0.03
PRF	0.25 *	−0.06	0.20 *	0.18	−0.46 ***	−0.05	0.16	−0.23 *
PLF	−0.10	0.06	−0.02	−0.09	0.00	0.29 **	0.12	−0.34 ***
CTI	0.62 ***	−0.36 ***	0.64 ***	0.08	0.32 **	−0.01	−0.16	0.35 ***
PADIR	0.12	−0.62 ***	−0.53 ***	0.02	0.20 *	−0.42 ***	−0.09	0.70 ***
CDT	−0.34 ***	0.14	−0.18	0.51 ***	−0.17	0.60 ***	0.01	0.36 ***
CDV	−0.56 ***	0.12	0.01	0.40 ***	−0.09	0.56 ***	0.21 *	0.36 ***
CDSR	−0.12	0.46 ***	−0.09	0.28 **	−0.25 *	0.43 ***	0.34 ***	0.21 *
CDS	−0.07	0.17	−0.33 ***	−0.09	−0.49 ***	0.36 ***	0.38 ***	0.13
CDMR	−0.25 *	0.64 ***	0.09	0.20 *	−0.61 ***	0.06	0.05	−0.03
CDLC	−0.53 ***	0.20 *	−0.01	0.71 ***	0.10	0.57 ***	−0.50 ***	0.07
CDLR	−0.51 ***	0.28 **	0.05	0.55 ***	0.05	0.58 ***	0.07	0.39 ***
SDV	0.86 ***	0.14	−0.23 *	−0.35 ***	0.20 *	−0.50 ***	−0.21 *	−0.39 ***

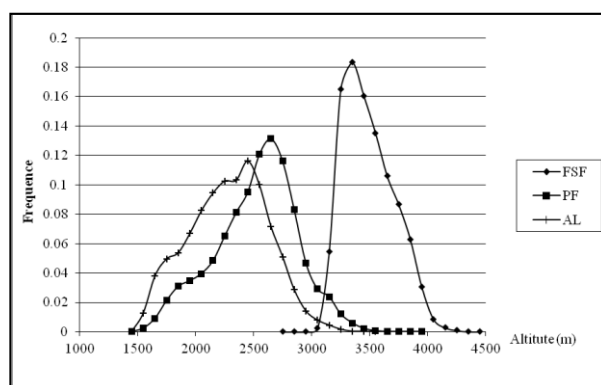*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

For both the northern and southern study areas, classification accuracy could be significantly improved by incorporating DEM or 4 CCA axes as additional channels (Tables 12 and 13). In the northern steep mountainous terrain, the importance of the DEM or CCA axes for vegetation mapping can be explained by the fact that the distribution of forest types is mainly determined by elevation (*i.e.*, CCA axis1 is strongly correlated to elevation (Table 14). Likewise, the spatial distribution of fir and spruce forest (FSF), pine forest (PF), agricultural land (AL) and water are strongly affected by elevation (Figure 4). In the southern study area, elevation is also the main determining factor for the distribution of rubber trees, agricultural land, deciduous forest, and evergreen forest.

The classification performed without the DEM resulted in unrealistic distributions of some of the vegetation classes: e.g., pine forest (PF) and agricultural land (AL) occurring at altitude levels above 3,300 m, and even above 4,000 m; and fir and spruce forest (FSF) occurring below 2,500 m (Figure 8). Without DEM agricultural land (AL) was confused with alpine and subalpine shrubs and meadows (DSM), and pine forest (PF) was always mistakenly classified as fir and spruce forest (FSF) (Figure 2). Hence, integrating DEM data in the classification process resulted in a more realistic spatial distribution pattern of vegetation classes whose occurrence is determined by altitude, such as fir and spruce forest (FSF), pine forest (PF) and agricultural land (AL) (Figure 9). Both pine forest (PF) and agricultural land (AL) are no longer found in areas with an elevation above 3,500 m, while the fir and spruce forest (FSF) no longer occurs at elevations lower than 3,000 m.

Figure 8. Frequency distribution of fir and spruce forest (FSF), pine forest (PF), and agricultural land (AL) obtained by ANN classification without DEM as an extra channel.
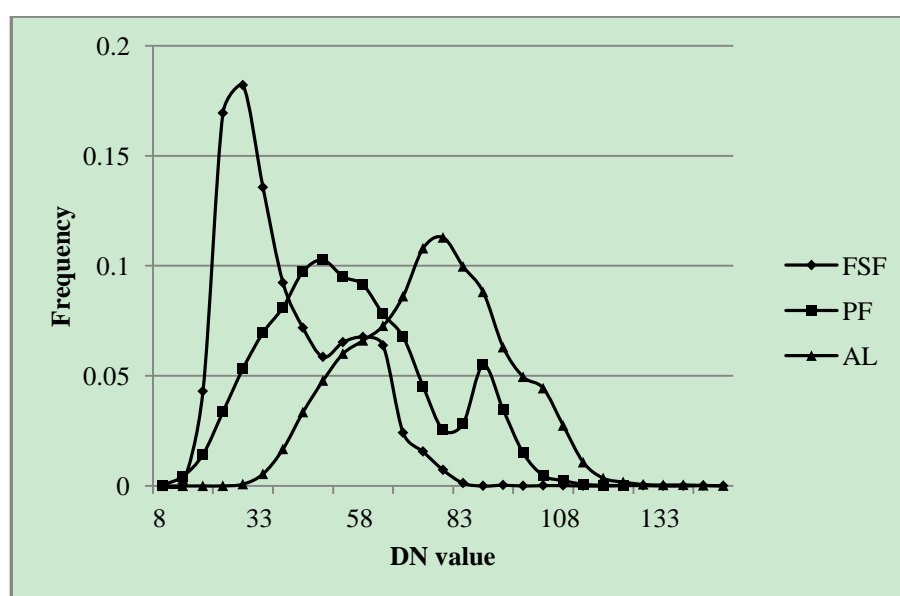

Figure 9. Frequency distribution of fir and spruce forest (FSF), pine forest (PF), and agricultural land (AL) obtained by ANN classification using DEM as an extra channel.
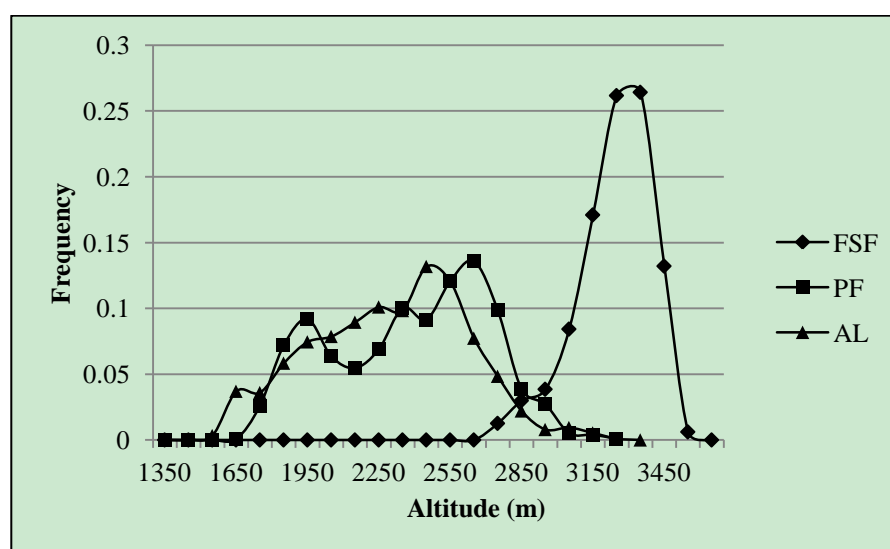
The distributions of the training data for three classes (fir and spruce forest (FSF), pine forest (PF), and agricultural land (AL)) in the Landsat TM near infrared (NIR) band are displayed in Figure 10. Pine forest (PF) overlaps both with the patterns of fir and spruce forest (FSF) and agricultural land (AL) indicating that pine forest is difficult to separate from both classes. Patterns of overlap in the DEM values are different: the DEM values of fir and spruce forest (FSF) almost have no overlap with those of pine forest (PF) (Figure 11). On the other hand, agricultural land and pine forest overlap completely when their altitude distributions are compared. These graphs indicate that incorporating DEM values in the image classification process is recommended for vegetation types whose distributions are determined by altitude. This corresponds with research work reported by Elummoh *et al.* [23], Liu *et al.* [26], Ren *et al.* [28], Zhang *et al.* [34].

Figure 10. Frequency distribution of DN value for FSF, PF and AL in Landsat 5 TM NIR band.
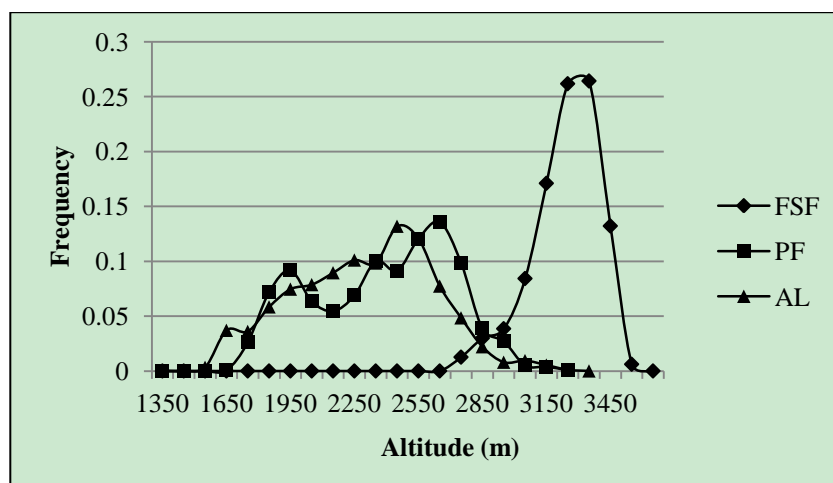

**Figure 11.** Frequency distribution of DEM values for FSF, PF and AL.

In the northern study area, the improvements to classification accuracy when adding the DEM or CCAs were very similar (Table 12). This is quite logical as the CCA ordination results showed that the CCA axes are strongly correlated to the computed topographic attributes, particularly for CCA ax1 and elevation ($r = 0.95$) (Table 10), suggesting that CCA axes and DEM data are redundant.

In the southern study area, the accuracy is significantly higher using 4 CCA axes compared to incorporating a DEM as an extra band (Table 13). As mentioned above, in the southern study area, the variance explained by human disturbance factors is higher than the variance explained by topographic attributes. As vegetation and spatial distribution pattern's are strongly influenced by human disturbance (e.g., rubber plantations, burned land, and agricultural land), classification accuracy can be significantly improved by adding ancillary data relating to human influence factors. Figure 12 shows a large overlap between the spectral response patterns of young rubber trees and agricultural land in the NIR band (Band 4). This overlap can be reduced by using CCA axis2, which is significantly correlated with the human influence factors, except for cost distance to middle rivers (Table 14) as an extra channel (Figure 13).

Figure 12. Frequency distribution of DN values for YRT and AL in Landsat TM band 4.

Figure 13. Frequency distribution of CCA scores for YRT and AL in axis2.
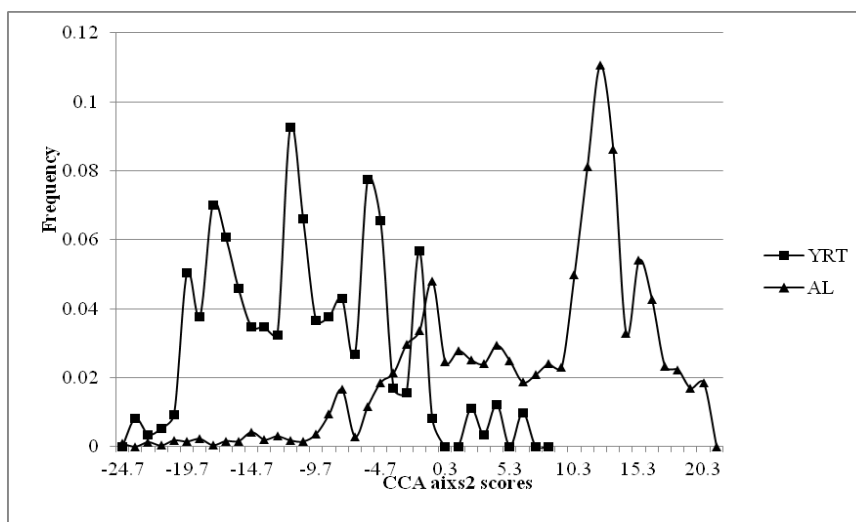
A large part of the variation in the vegetation data remains unexplained. The explained variability depends on the number of vegetation classes and samples [81]. Mountain vegetation is the result of a complex interaction between natural processes and human intervention, and it is impossible to completely explain the response of the vegetation to a limited set of attributes [4,32]. Additionally, some species have a relatively wide ecological amplitude, so there will never be a perfect species-environment correlation. Soil is another important biophysical factor for vegetation spatial distribution pattern, but this parameter was not included in the research work. Furthermore, the Kriging interpolation scheme for CCA axes introduced errors in feature space, which may explain why the CCA scores do add little to the accuracy of the image classification.

Compared with the southern study area, the northern study area is much more rugged: The effects of topographic factors on vegetation spatial distribution, hence stronger. However, in the southern study area, with its higher population density, the human disturbance is stronger than in the northern study area [107]. Anthropogenous effects on the vegetation spatial pattern are stronger than the effects of topography in this area. Since the establishment of rubber plantations in 1956 in the southern study area (Xishuangbanna prefecture, Yunnan, China), the expansion of rubber plantation has accelerated, particularly after 1978 [108–110]. In the southern study area, another important human disturbance factor is fire (burned land). Xishuangbanna is the traditional homeland of upland minority people ("hill tribes") including Dai, Hani (called Akha in Thailand), and Bulang [110]. Hani (Akha) and Bulang people traditionally practice shifting cultivation in the uplands and cultivated rice in the lowlands [110–112]. From field observations it has been established that the local people commonly use fire to clear forest and shrub land for rubber plantation.

5. Conclusions

The purpose of this study was to present a novel method for mapping mountain vegetation in Lancang Mekong, China, whereby the added value of Canonical Correspondence Analysis (CCA) to an Artificial Neural Network (ANN) was investigated. All the topographic attributes were significantly correlated with vegetation spatial distribution patterns in both study areas. Elevation and slope were established as the two most important influencing factors. In addition, CCA ordination reduced the vegetation data and environmental (topography and human disturbance) factors into four independent ordination axes. Inclusion of these additional variables improved the classifications up to a maximum of 6%. This was particularly the case for vegetation classes strongly affected by topography, such as example pine forest, fir and spruce forest in the north, and rubber plantation in the south. Also, with an overall accuracy of more than 95% and kappa value of more than 0.95, it was demonstrated that inclusion of the CCA ordination model into image analysis holds considerable promise for an automated approach of mountain vegetation mapping. It was found important to separate the vegetation classes into subclasses based on the spectral response patterns, to implement the training stage eventually resulting in improved classification accuracy. In the case of the more rugged terrain in the northern study area, elevation (DEM) and CCA axes were equally important since the most important factor driving the distribution of vegetation is altitude. For the southern area, the CCA axes were more important than the DEM for improving the classification. Accurate mapping of rubber

plantations in the south will be very useful for ecosystem conservation, sustainable planning, and management of ecologically fragile mountain area.

As mentioned in the introduction, one of the big challenges for mapping mountain vegetation is the scale issue of the integration between plant community data, human disturbance factors and geo-spatial data (*i.e.*, DEM and image data). Further work will therefore focus on the question of scale interactions and scale gaps concerning the data used for dominant vegetation community mapping, such as plant community data, geographic spatial data and multi-resolution image data. Furthermore, it is acknowledged that the Kriging interpolation scheme for CCA axes introduces errors in feature space, which may partly explain why the CCA scores attribute little to the accuracy of the final image classification. Further development of the interpolation scheme may need to be developed in future work as well.

Acknowledgments

Author Contributions

Zhiming Zhang, Robert De Wulf, and Xiaokun Ou designed research; Zhiming Zhang performed research and analyzed data; Zhiming Zhang, Robert De Wulf, and Frieke Van Coillie wrote the paper.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Itten, K.I.; Meyer, P. Geometric and radiometric correction of TM-data of mountainous forested areas. *IEEE Trans. Geosci. Remote Sens.* **1992**, *31*, 764–770.
2. Gu, D.; Gillespie, A.R.; Adams, J.B.; Week, R. A statistical approach for topographic correction of satellite images by using spatial context information. *IEEE Trans. Geosci. Remote Sens.* **1999**, *37*, 236–246.
3. Dymond, C.C.; Johnson, E.A. Mapping vegetation spatial patterns from modeled water, temperature and solar radiation gradients. *ISPRS J. Photogramm. Remote Sens.* **2002**, *57*, 69–85.
4. Cantón, Y.; Barrio, G.D.; Solé-Benet, A.; Lázaro, R. Topographic controls on the spatial distribution of ground cover in the Tabernas badlands of SE Spain. *Catena* **2004**, *55*, 341–365.
5. Deng, Y.X.; Chen, X.F.; Chuvieco, E.; Warner, T.; Wilson, J.P. Multi-scale linkages between topographic attributes and vegetation indices in a mountainous landscape. *Remote Sens. Environ.* **2007**, *111*, 122–134.

6. Brown, D.G. Predicting vegetation types at treeline using topographic and biophysical disturbance variables. *J. Veg. Sci.* **1994**, *5*, 641–656.

7. Crave, E.; Gascuel-Odoux, C. The influence of topography on time and space distribution of soil surface water content. *Hydrol. Process.* **1997**, *11*, 203–210.

8. Gómez-Plaza, A.; Martíñez-Mena, M.; Alvadalejo, J.; Castillo, V.M. Factors regulating spatial distribution of soil water content in small semiarid catchments. *J. Hydrol.* **2001**, *253*, 211–226.

9. Western, A.W.; Grayson, R.B.; Blöschl, G.; Willgoose, G.R.; McMahon, T.A. Observed spatial organization of soil moisture and its relation to terrain indices. *Water Resour. Res.* **1999**, *35*, 797–810.

10. Frank, T.D. Mapping dominant vegetation communities in the Colorado Rocky Mountain front range with Landsat Thematic Mapper and digital terrain data. *Photogramm. Eng. Remote Sens.* **1988**, *54*, 1727–1734.

11. Smith, J.A.; Lin, T.L.; Ranson, K.J. The Lambertian assumption and Landsat data. *Photogramm. Eng. Remote Sens.* **1980**, *46*, 1183–1189.

12. Civco, D.L. Topographic normalization of Landsat Thematic Mapper digital imagery. *Photogramm. Eng. Remote Sens.* **1989**, *55*, 1303–1309.

13. Gu, D.; Gillespie, A. Topographic normalization of Landsat TM images of forest based on subpixel sun-canopy-sensor geometry. *Remote Sens. Environ.* **1998**, *64*, 166–175.

14. Dymond, J.R; Shepherd, J.D. Correction of the topographic effect in remote sensing. *IEEE Trans. Geosci. Remote Sens.* **1999**, *37*, 2618–2610.

15. Gao, Y.; Zhang, W. A simple empirical topographic correction method for ETM+ imagery. *Int. J. Remote Sens.* **2009**, *30*, 2259–2275.

16. Soenen, S.A.; Peddle, D.R.; Cobure, C.A. SCS + C: A modified sun-canopysensor topographic correction in forested terrain. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 2148–2159.

17. Soenen, S.A.; Peddle, D.R.; Cobure, C.A.; Hall, R.J.; Hall, F.G. Improved topographic correction of forest image data using a 3-D canopy reflectance model in multiple forward mode. *Int. J. Remote Sens.* **2008**, *29*, 1007–1027.

18. Huang, H.; Gong, P.; Clinton, N.; Hui, F. Reduction of atmospheric and topographic effect on Landsat TM data for forest classification. *Int. J. Remote Sens.* **2008**, *29*, 5623–5642.

19. Cuo, L.; Vogler, J.B.; Fox, J.M. Topographic normalization for improving vegetation classification in a mountainous watershed in Northern Thailand. *Int. J. Remote Sens.* **2010**, *31*, 3037–3050.

20. Reese, H.; Olsson, H. C-correction of optical satellite data over alpine vegetation areas: A comparison of sampling strategies for determining the empirical c-parameter. *Remote Sens. Environ.* **2011**, *115*, 1387–1400.

21. Teillet, P.M.; Guindon, B.; Goodenough, D.G. On the slope-aspect correction of multispectral scanner data. *Can. J. Remote Sens.* **1982**, *8*, 84–106.

22. Hutchinson, C.F. Techniques for combining Landsat and ancillary data for digital classification improvement. *Photogramm. Eng. Remote Sens.* **1982**, *48*, 123–130.

23. Elummoh, A.; Shrestha, R.P. Application of DEM data to Landsat image classification: Evaluation in a tropical wet-dry landscape of Thailand. *Photogramm. Eng. Remote Sens.* **2000**, *66*, 297–304.

24. Dorren, L.K.A.; Maier, B.; Sejimonsbergen, A.C. Improved Landsat-based forest mapping in steep mountainous terrain using object-based classification. *For. Ecol. Manag.* **2003**, *183*, 31–46.

25. Wulder, M.A.; Franklin, S.E.; White, J.C.; Cranny, M.C.; Dechka, J.A. Inclusion of topographic variables in an unsupervised classification of satellite imagery. *Can. J. Remote Sens.* **2004**, *30*, 137–149.

26. Liu, Q.J.; Takamura, T.; Takeuchi, N. Mapping of boreal vegetation of a temperate mountain in China by multitemporal Landsat TM imagery. *Int. J. Remote Sens.* **2002**, *23*, 3385–3405.

27. Brown, D.C.E.C.; Story, M.H.; Thompson, C.C.K.; Timothy, G.S.; James, R.I. National park vegetation mapping using multitemporal Landsat 7 data and a decision tree classifier. *Remote Sens. Environ.* **2003**, *85*, 316–327.

28. Ren, G.; Zhu, A.; Wang, W.; Xiao, W.; Huang, Y.; Li, G.; Li, D.; Zhu, J. A hierarchical approach coupled with coarse DEM information for improving the efficiency and accuracy of forest mapping over very rugged terrains. *For. Ecol. Manag.* **2009**, *258*, 26–34.

29. Mather, P.M. A computationally-efficient maximum-likelihood classifier employing prior probabilities for remotely-sensed data. *Int. J. Remote Sens.* **1985**, *6*, 369–376.

30. Maselli, F.; Conese, C.; Petkov, L.; Resti, R. Inclusion of prior probabilities derived from a nonparametric process into the maximum likelihood classifier. *Photogramm. Eng. Remote Sens.* **1992**, *58*, 833–839.

31. Maselli, F.; Conese, C.; de Tiziana, F.; Maurizio, R. Integration of ancillary data into a maximum-likelihood classifier with nonparametric priors. *ISPRS J. Photogramm. Remote Sens.* **1995**, *50*, 2–11.

32. Foody, G.M.; Campbell, N.A.; Trodd, N.M.; Wood, T.F. Derivation and applications of probabilistic measures of class membership from the maximum-likelihood classification. *Photogramm. Eng. Remote Sens.* **1992**, *58*, 1335–1341.

33. McIver, D.K.; Friedl, M.A. Using prior probabilities in decision-tree classification of remotely sensed data. *Remote Sens. Environ.* **2002**, *81*, 253–261.

34. Zhang, Z.; de Clercq, E.M.; Ou, X.; de Wulf, R.R.; Verbeke, L.P.C. Mapping dominant communities vegetation in Meili Snow Mountain, Yunnan Province, China using satellite imagery and plant community data. *Geocarto Int.* **2008**, *23*, 135–153.

35. Pfeffer, K.; Pebesma, E.J.; Burrough, P.A. Mapping alpine vegetation using vegetation observation and topographic attributes. *Landsc. Ecol.* **2003**, *18*, 759–776.

36. Turner, M.G.; Gardner, R.H.; O'Neill, R.V. *Landscape Ecology in Theory and Practice: Pattern and Process*; Springer: New York, NY, USA, 2001.

37. Hörsch, B. Modeling the spatial distribution of montane and subalpine forests in the central Alps using digital elevation models. *Ecol. Model.* **2003**, *168*, 267–282.

38. Farina, A. Emerging Processes in the Landscape. In *Principles and Methods in Landscape Ecology*, 3rd ed.; Springer: Dordrecht, The Netherlands, 2006; pp. 110–128.

39. Sherman, R.; Mullen, R.; Li, H.; Fang, Z.D.; Wang, Y. Spatial patterns of plant diversity and communities in alpine ecosystems of the Hengduan Mountains, northwest Yunnan, China. *Plant Ecol.* **2008**, *1*, 117–136.

40. Ohmann, J.; Spies, T.A. Regional gradient analysis and spatial pattern of woody plant communities of Oregon forests. *Ecol. Monogr.* **1998**, *68*, 151–182.

41. Begon, M.; Townsend, C.R.; Happer, J.L. The Nature of the Community: Patterns in Space and Time. In *Ecology—from Individuals to Ecosystems*, 4th ed.; Blackwell: Malden, MA, USA, 2006; pp. 469–498.

42. Gauch, H.G.J. *Multivariate Analysis in Community Ecology*; Cambridge University Press: Cambridge, UK, 1982.

43. Thomas, V.; Treitz, P.; Dennis, J.; John, M.; Peter, L.; McCaughey, H.J. Image classification of a northern peatland complex using spectral and plant community data. *Remote Sens. Environ.* **2002**, *84*, 83–89.

44. Bio, A.M.F. Does Vegetation Suit Our Models? Data and Model Assumptions and the Assessment of Species Distribution in Space. Ph.D. Thesis, Faculty of Geographic Science, Utrecht, The Netherlands, 2000; p. 195.

45. Dirnböck, T.; Dullinger, S.; Gottfried, M.; Ginzler, C.; Grabherr, G. Mapping alpine vegetation based on image analysis, topographic variables and Canonical correspondence analysis. *Appl. Veg. Sci.* **2003**, *6*, 85–96.

46. Franklin, J.; Woodcock, C.E. Multiscale Vegetation Data for the Mountains of Southern California: Spatial and Categorical Resolution. In *Scale in Remote Sensing and GIS*; Quattorchi, D.A., Goodchild, M.F., Eds.; CRC-Lewis Publishers: Boca Raton, FL, USA, 1997; pp. 141–171.

47. Dirnböck, T.; Hobbs, R.J.; Caccetta, P.A. Vegetation distribution in relation to topographically driven process in south-western Australia. *Appl. Veg. Sci.* **2002**, *5*, 147–158.

48. Guisan, A.; Zimmermann, N.E. Predictive habitat distribution models in ecology. *Ecol. Model.* **2000**, *135*, 147–168.

49. Horning, N.; Robinson, J.A.; Sterling, E.J.; Turner, W.; Spector, S. Measuring and Monitoring Land Cover, Land Use, and Vegetation Characteristics. In *Remote Sensing for Ecology and Conservation*; Oxford University Press: Oxford, UK, 2010; pp. 81–119.

50. Pu, Y.S; Zhang, Z.Y. A strategic study on biodiversity conservation in Xishuangbanna. *J. For. Res.* **2001**, *12*, 25–30.

51. Anderson, D.M.; Salick, J.; Moseley, R.K.; Ou, X. Conserving the sacred medicine mountains: A vegetation analysis of Tibetan sacred sites in Northern Yunnan. *Biodivers. Conserv.* **2005**, *14*, 3065–3091.

52. Mittermeier, R.A.; Myers, N.; Thomsen, J.B.; Fonseca, G.A.B.D.; Olivieri, S. Biodiversity hotspots and major tropical wilderness areas. *Conserv. Biol.* **1998**, *12*, 516–520.

53. Salick, J.; Amend, A.; Anderson, D.; Hoffmeister, K.; Gunn, B.; Fang, Z. Tibetan sacred sites conserve old growth trees and cover in the eastern Himalayas. *Biodivers. Conserv.* **2007**, *16*, 693–706.

54. Deng, M.; Zhang, P.; Zhao, B.; Wang, Y. *The Forest in Lanping Pumi Autonomous County*; Yuannan Ethnic Press: Kunming, China, 1997 (In Chinese).

55. Verbeke, L.P.C.; Vancoillie, F.M.B.; de Wulf, R.R. Reusing back-propagation artificial neural networks for land cover classification in tropical savannahs. *Int. J. Remote Sens.* **2004**, *25*, 2747–2771.

56. Gitas, Z.; Devereux, B.J. The role of topographic correction in mapping recently burned Mediterranean forest areas from Landsat TM images. *Int. J. Remote Sens.* **2006**, *27*, 41–54.

57. Zhang, Z.; de Wulf, R.R.; van Coillie, F.M.B.; Verbeke, L.P.C.; de Clercq, E.M.; Ou, X. The influence of different topographic correction strategies on vegetation classification accuracy in the Lancang watershed, China. *J. Appl. Remote Sens.* **2011**, *5*, doi:10.1117/1.3569124.

58. Wilson, J.P.; Gallant, J.C. *Terrain Analysis Principles and Applications*; John Wiley and Sons: New York, NY, USA, 2000.

59. Pinder, J.E.; Kroh, G.C.; White, J.D.; Basham May, A.M. The relationship between vegetation types and topography in Lassen Volcanic National Park. *Plant Ecol.* **1997**, *131*, 17–29.

60. Zhang, Z.; Wang, W.; Ou, X.; Wu, W. The correlation between vegetation spatial pattern and environmental factors in Meili Snow Mountain, China. *J. Yunnan Univ.* **2009**, *31*, 311–315 (In Chinese).

61. Shen, Z.; Zhang, X. The spatial pattern and topographic interpretation of the forest vegetation at Dalaoling Region in the Three Gorges. *Acta Bot. Sin.* **2000**, *42*, 1089–1095 (In Chinese).

62. Moore, I.D.; Grayson, R.B.; Ladson, A.R. Digital terrain modeling: A review of hydrological, geomorphological, and biological application. *Hydrol. Process.* **1991**, *5*, 3–30.

63. Jenness, J. Topographic Position Index (tpi_jen.avx) Extension for ArcView 3.x, v. 1.3a. Jenness Enterprises, 2006. Available online: http://www.jennessent.com/arcview/tpi.htm (accessed on 18 July 2007).

64. Shen, Z.; Zhang, X.; Jin, X. Gradient analysis of the influence of mountain topography on vegetation pattern. *Acta Phytoecol. Sin.* **2000**, *24*, 430–435 (In Chinese).

65. McCune, B.; Keon, D. Equations for potential annual direct incident radiation and heat load. *J. Veg. Sci.* **2002**, *13*, 603–606.

66. Serneels, S.; Lambin, E.F. Proximate causes of land-use change in Narok District Kenya: A spatial statistical model. *Agric. Ecosyst. Environ.* **2001**, *85*, 65–81.

67. Geist, H.J.; Lambin, E.F. Proximate causes and underlying driving forces of tropical deforestation. *BioScience* **2002**, *52*, 143–150.

68. Overmars, K.; Verburg, P.H. Analysis of land use drivers at the watershed and household level: Linking two paradigms at the Philippine forest fringe. *Int. J. Geogr. Inf. Sci.* **2005**, *19*, 125–152.

69. Mottet, A.; Ladet, S.; Coque, N.; Gibon, A. Agricultural land-use change and its drivers in mountain landscapes: A case study in the Pyrenees. *Agric. Ecosyst. Environ.* **2006**, *114*, 296–310.

70. McCoy, J.; Johnston, K. *Using ArcGIS' Spatial Analyst*; ESRI Press: Redlands, CA, USA, 2001.

71. Long, H.L.; Liu, Y.S.; Wu, X.Q.; Dong, G.H. Spatial-temporal dynamic patterns of farmland and rural settlements in Su-Xi-Chang region: Implications for building a new countryside in coastal China. *Land Use Policy* **2009**, *26*, 322–333.

72. Deffontaines, J.P.; Thenail, C.; Baudry, J. Agricultural systems and landscape patterns: How can we build a relationship. *Landsc. Urban Plann.* **1995**, *31*, 3–10.

73. Nagendra, H.; Southworth, J.; Tucker, C. Accessibility as a determinant of landscape transformation in western Honduras: Linking pattern and process. *Landsc. Ecol.* **2003**, *18*, 141–158.

74. Aspinall, R. Modelling land use change with generalized linear models—A multi-model analysis of change between 1860 and 2000 in Gallatin Valley, Montana. *J. Environ. Manag.* **2004**, *72*, 91–103.

75. Rutherford, G.N.; Bebi, P.; Edwards, P.J.; Zimmermann, N.E. Assessing land-use statistics to model land cover change in a mountainous landscape in the European Alps. *Ecol. Model.* **2008**, *212*, 460–471.

76. Foltête, J.C.; Berthier, K.; Cosson, J.F. Cost distance defined by a topological function of landscape. *Ecol. Model.* **2008**, *210*, 104–114.

77. Schneider, K.; Robbins, P. Exercise 3: Modeling Access to Health Care: Anisotropic Cost Distance. Exploration in Geographic Information Systems Technology. In *GIS and Mountain Environments*; United Nations Institute of Training and Research, Clark University: Worcester, MA, USA, 1995; Volume 5.

78. Rees, W.G. Least-cost paths in mountainous terrain. *Comput. Geosci.* **2004**, *30*, 203–209.

79. LaRue, M.A.; Nielsen, C.K. Modelling potential dispersal corridors for cougars in midwestern North America using least-cost path methods. *Ecol. Model.* **2008**, *212*, 372–381.

80. Eastman, J.R. Pushbroom algorithms for calculating distances in raster grids. *Proc. Autocarto* **1989**, *9*, 288–297.

81. Lillesand, T.M.; Kiefer, R.W.; Chipman, J.W. Digital Image Processing. In *Remote Sensing and Image Interpretation*, 5th ed.; John Wiley & Sons: New York, NY, USA, 2004; pp. 482–557.

82. Atkinson, P.M.; Tatnall, A.R.L. Neural networks in remote sensing. *Int. J. Remote Sens.* **1997**, *18*, 699–709.

83. Mas, J.F.; Flores, J.J. The application of artificial neural networks to the analysis of remotely sensed data. *Int. J. Remote Sens.* **2008**, *29*, 617–663.

84. Benediktsson, J.A.; Swain, P.H.; Ersoy, O.K. Neural network approaches *versus* statistical methods in classification of mutisource remote sensing data. *IEEE Trans. Geosci. Remote Sens.* **1990**, *28*, 540–551.

85. Leps, J.; Smilauer, P. *Multivariate Analysis of Ecological Data Using of CANOCO*; Cambridge University Press: Cambridge, UK, 2003.

86. Parker, A.J. Comparative gradient structure and forest cover types in Lassen Volcanic and Yosemite National Parks, California. *Bull. Torrey Bot. Soc.* **1995**, *122*, 58–68.

87. Parker, A.J. Forest/environment relationship in Yosemite National Park, Califonia. *Vegetatio* **1989**, *82*, 41–54.

88. Shen, Z. A multi-scale study on the vegetation-environment relationship of a mountain forest transect. *Acta Ecol. Sin.* **2002**, *22*, 461–470 (In Chinese).

89. Ter Braak, C.J.F.; Šmilauer, P. *CANOCO Reference Manual and CanoDraw for Windows User's Guide: Software for Canonical Community Ordination (Version 4.5)*; Microcomputer Power: Ithaca, NY, USA, 2002; p. 500.

90. Ter Braak, C.J.F. Canonical correspondence analysis: A new eigenvector technique for multivariate direct gradient analysis. *Ecology* **1986**, *67*, 1167–79.

91. Wagner, H.H. Direct multi-scale ordination with canonical correspondence analysis. *Ecology* **2004**, *85*, 342–351.

92. Legendre, P.; Legendre, L. *Numerical Ecology*, 2nd ed.; Elsevier: Amsterdam, The Netherlands, 1998.

93. Burrough, P.A.; McDonnell, R.A. *Principles of Geographical Information Systems*; Oxford University Press: Oxford, UK, 1998.

94. Wagner, H.H. Spatial covariance in plant communities: Integrating ordination, geostatistics, and variance testing. *Ecology* **2003**, *84*, 1045–1057.

95. Johnston, K.; ver Hoef, J.M.; Krivoruchko, K.; Lucas, N. *Using ArcGIS' Geostatistical Analyst*; ESRI Press: Redlands, CA, USA, 2001.

96. Webster, R.; Oliver, M.A. *Geostatistics for Environmental Scientists*, 2nd ed.; John Wiley & Sons: New York, NY, USA, 2007.

97. Cressie, N. *Statistics for Spatial Data*, Revised ed.; John Wiley & Sons: New York, NY, USA, 1993; p. 900.

98. Guillemette, N.; St-Hilaire, A.; Ouarda, T.B.M.J.; Bergeron, N.; Robichaud, É.; Bilodeau, L. Feasibility study of a geostatistical modelling of monthly maximum stream temperatures in a multivariate space. *J. Hydrol.* **2009**, *364*, 1–12.

99. Dayhoof, J.E. *Neural Network Architectures*; Van Nostrand Reinhold: New York, NY, USA, 1990.

100. Zhang, Z.; Verbeke, L.P.C.; de Clercq, E.M.; Ou, X.; de Wulf, R.R. Vegetation change detection using artificial neural networks with ancillary data in Xishuangbanna, Yunnan Province, China. *Chin. Sci. Bull.* **2007**, *52*, 232–243.

101. Congalton, R.G.; Green, K. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*; Lewis Publisher: Boca Raton, FL, USA, 1999.

102. Cohen, J. A coefficient of agreement for nominal scales. *Educ. Psychol. Meas.* **1960**, *1*, 37–40.

103. Congalton, R.G.; Oderwald, R.G.; Mead, R.A. Assessing Landsat classification accuracy using discrete multivariate analysis statistical techniques. *Photogramm. Eng. Remote Sens.* **1983**, *46*, 1671–1678.

104. Van Coillie, F. Design and Application of Artificial Neural Networks for Digital Image Classification of Tropical Savanna Vegetation. Ph.D. Thesis, Ghent University, Ghent, Belgium, 2003.

105. Schowengerdt, R.A. Thematic Classification. In *Remote Sensing, Models and Methods for Image Processing*, 3rd ed.; Academic Press: New York, NY, USA, 2007; pp. 387–455.

106. Fahsi, A.; Tsegaye, T.; Tadesse, W.; Coleman, T. Incorporation of digital elevation models with Landsat-TM data to improve land cover classification accuracy. *For. Ecol. Manag.* **2000**, *128*, 57–64.

107. Yunnan Statistic Bureau. *Yunnan Statistical Yearbook 2004*; Yunnan Statistical Press: Kunming, China, 2004 (In Chinese).

108. Li, H.M.; Aide, T.M.; Ma, Y.X.; Liu, W.J.; Cao, M. Demand for rubber is causing the loss of high diversity rain forest in SW China. *Biodivers. Conserv.* **2007**, *16*, 1731–1745.

109. Liu, W.J.; Hu, H.B.; Ma, Y.X.; Li, H.M. Environmental and socioeconomic impacts of increasing rubber plantations in Menglun Township, southwest China. *Mt. Res. Dev.* **2006**, *26*, 245–253.

110. Xu, J.C.; Fox, J.; Vogler, J.B.; Zhang, P.F.; Fu, Y.S.; Yang, L.X.; Qian, J.; Leisz, S. Land-use and land-cover change and farmer vulnerability in Xishuangbanna Prefecture in Southwestern China. *Environ. Manag.* **2005**, *36*, 404–413.

111. Xu, J.C.; Fox, J.; Lu, X.; Podger, N.; Leisz, S.; Ai, X. Effects of swidden cultivation, state polices and customary institutions on land cover in Hani Village, Yunnan. *Mt. Res. Dev.* **1999**, *19*, 123–132.

112. Rerkasem, K.; Yimyam, K.; Rerkasem, B. Land use transformation in the mountainous mainland Southeast Asia region and the role of indigenous knowledge and skills in forest management. *For. Ecol. Manag.* **2009**, *257*, 2035–2043.