CHAPTER 3

RESEARCH METHODS

3.1 Location of the study site

The study sites are located in Pak Ou District, Luang Prabang Province. It lies in the 19° 57' 39'' to 20° 02' 22'' N latitude and 102° 13' 01'' to 102° 19' 47'' E longitude with elevations ranging from 279 to 659 meters, this area is surrounded by high mountains and rough valleys. They are located and lie up along the Nam Seuang river bank the road lead to Pak Xeang District and the national road No.13th north the road lead to China border, far from the center point of Luang Prabang city to the north about 30 km (Figure 3.1). The study site consists of 3 villages: Sanghai, Thinchaleon and Phonsavang, the total area covered about 3,269.94 hectares, which has 319 families with the population of 1,609 people. The study site shared boundaries with the neighboring villages are as the following: Pak Ou, Houy Mard and Phonhome village to the north. Pak Seuang and Phon Ngam village to the south. Kokwanh and Tha Oui village to the east. The Northern Agriculture and Forestry College's area and Mekong River to the west. Most of the farmers in this region engage in agriculture in the slopes of the mountains and along the river bank; some farmers still practice slash and burn cultivation for food consumption. The agricultural and social cultural systems in the area are very diversified and complex.

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

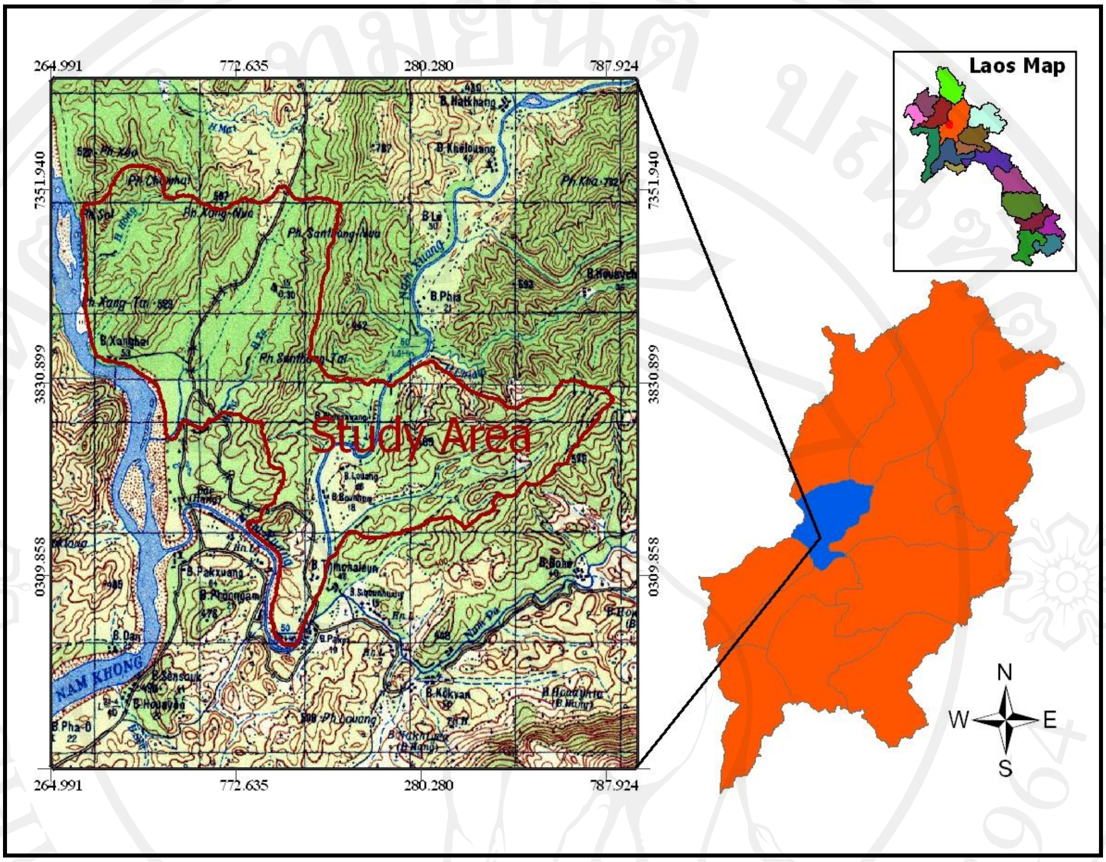

Figure 3.1 Location map of the study site

3.2 Field survey and data collection

3.2.1 Sampling design

There are many villages in Pak Ou district, three villages were selected for the study area with their different situations regarding land use change. Sanghai is a village which is changing in term of small industrial growth, Thinchaleon is a village which is changing in term of horticulture production and Phonsavang is a village which is changing in term of resettlement and multi ethnic group relocation. The samples were randomized according to the number of households in the village at least 50% of households from each village. Data collection was conducted covering

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

information on spatial data or geographical information that present the land use change as well as forest land, agricultural land and its' impacts on households' socio-economic context of the study site.

The conceptual framework for this study draws heavily from two types of data collection (Figure 3.2). Landsat5 imagery taken in 1990 and 2010 were used for derive the spatial data, while topographic map used for geo reference in selection of sample site. Field survey and data collection made use of global positioning system to collect ground control points and tracking study boundary. Socio-economic data were collected using interviews of each household with questionnaires and group discussion with key informants to obtain socio-economic context variables such as non-farm income, on-farm income, ethnicity, labor, household size, rice output, capital, livestock and household's education. Population data in 1990 and 2010 were used to explain the causes of land use change in the study area.

Landsat-5 imagery of 1990 and 2010 were interpreted and the spatial data were analyzed to obtain land use change of the study area during the last two decades and to obtain spatial variables such as distance from road, distance from water, distance from urban, slope and elevation. The spatial and socio-economic variables were analyzed by using logistic regression model to analyze the relationship between the factors that influencing on land use change in the study area.

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

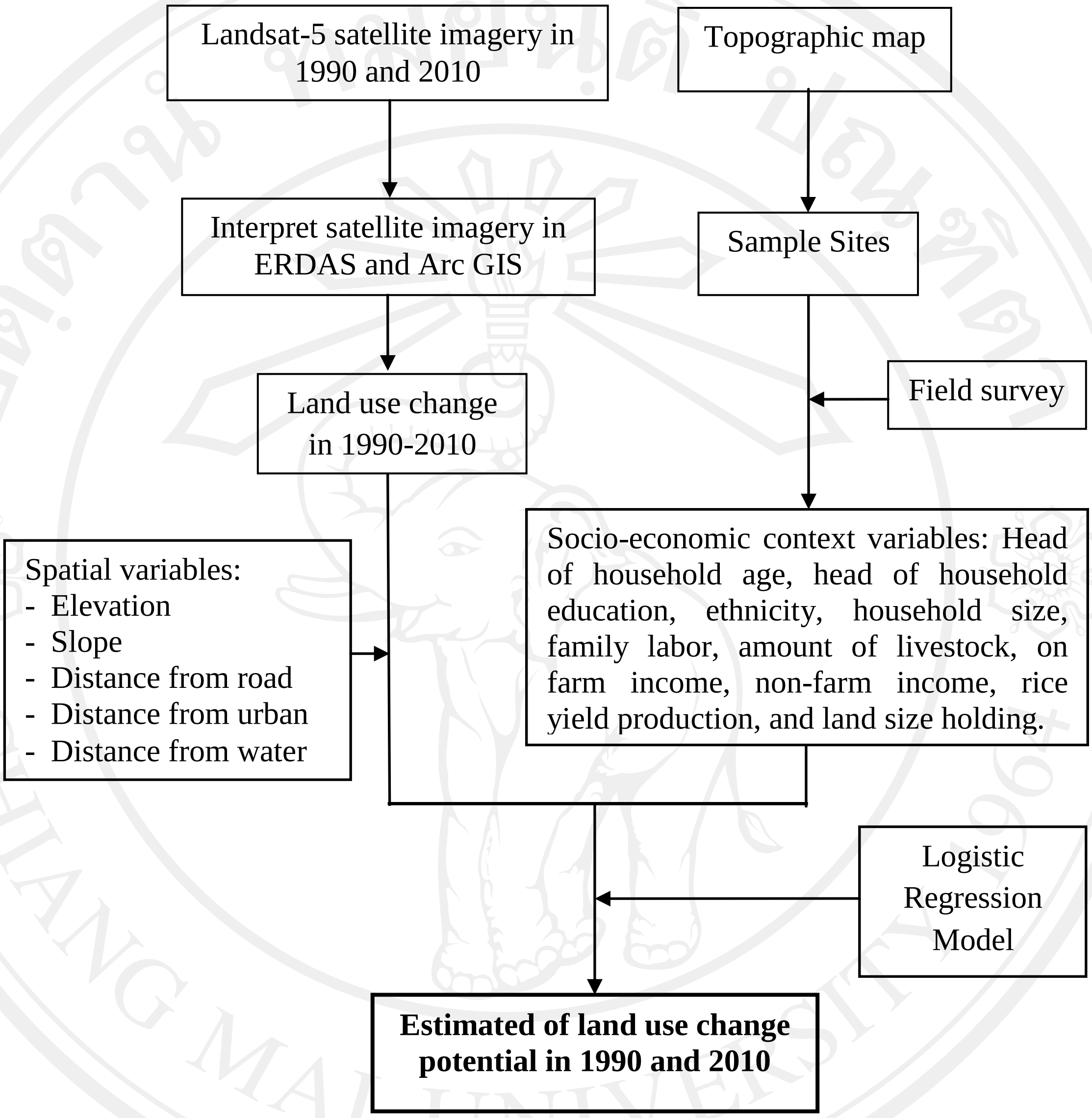

Figure 3.2 The overall conceptual framework of the study

3.2.2 Land use change characteristics

Changing in land use was derived from Landsat-5 imagery and field survey using spatial data in 1990 and 2010. Each image was taken at the dry season which has no heavy cloud between in January to March. Hence, they stem from the same cropping season and from comparable climatic conditions. This enhances the

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

interpretation as the spectral reflection of land cover is easier to be compared. In addition recently updated 1:50,000 UTM map is used for geo-referenced and merged to obtain a consistent set of base information. This map is allowed verification of land cover delineation using additional point information and linear features such as contours, roads, and rivers. The interpretation resulted in maps of land use change, which indicated physical land cover of the respective years for the major land use classes. Overlays will produce quantitative estimates in land use changes over the last two decades.

Remote sensing (RS) data from Landsat-5 imagery 2 dates taken on 15th January 1990 and on 7th February 2010 with resolution of 25 meters were used for this research (Figure 3.3). It is widely used in land use classification and can combine different attributes onto maps to provide an overview of the villages, their land use changes and management habitats, especially in land use change in the period of time for the last 20 years. It is necessary to identify the location and characteristics of natural resources by using Global Positioning System (GPS) applications accompanies with Geographical Information Systems (GIS) to identify the locations, ground control points and derived spatial data from satellite imagery. Locations can be visualized and linked with other related information such as geographical conditions, agricultural potential inherent in a specific plots, height slope and elevation from the sea level.

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

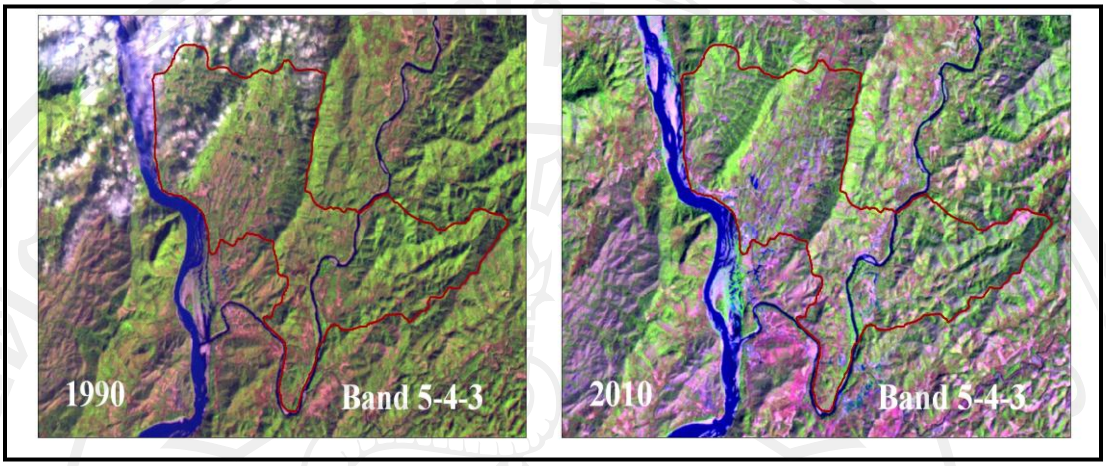

Figure 3.3 Landsat5 TM acquired in 1990 and 2010 covered study area

3.2.3 Primary and secondary data of socio-economic

The target survey was conducted in 3 selected villages out of a total at least 50% of households in each village of the study site were surveyed. The study used primary and secondary data. Primary data was collected through direct interview with respondents by using questionnaires and focused group discussion with key informants such as village headmen and oldest persons in the village. The survey will provide primary data on the current status and changes in land use practices within the last two decades and causes of these changes and effects on their livelihoods.

Secondary data of socio-economic are demographics; agricultural production and household's socio-economic statistics were obtained from official government statistics such as Provincial Agriculture and Forestry Office (PAFO) or District Administrative Bureau (DAB). Data analysis for socio-economic status was implemented in logistic regression model to analyze the relationship between influencing factors and land use change in the study area.

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

3.3 Image classification and data analysis

To answer the research questions and fulfill the first objective, to assess the spatial land use change pattern in the last two decades from 1990 to 2010 of the study area, spatial data from satellite images acquisition originating from different dates were used. Hence, Landsat-5 Thematic Mapper (TM) from January-March of 1990 and 2010 were used and analyzed to assess the spatial land use change of the study area.

To achieve the second objective, to determine factors affecting on land use change and their effects on farmers' livelihoods. Spatial and non-spatial data of land use were analyzed by using socio-economic context variables such as non-farm income, on-farm income, ethnicity, labor, household size, rice output, capital, livestock and household's education. Population in 1990 and 2010 were used to explain the relationship on land use change as well. While spatial variables comprised of distance from road, distance from water, distance from settlement and elevation. Both spatial and socio-economic variables were used as independent variables for logistic regression model to analyze the relationship between the factors that influencing on land use change in the study area.

3.3.1 Layer stacking

During layer stacking, the Universal Traverse Mercator (UTM) system with WGS84 as a datum was assigned as a preference as well as projection is concerned. All seven bands of TM were considered for Layers stacking in order to process satellite image. The nature of these different bands had to be considered to

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

make a decision as to which three band combination would be most helpful for classification and visual interpretation. Since band 4 reflective infrared wavelength (0.76-0.90 µm) is absorbed by water (appearing dark) and reflected by vegetation (appearing bright), while mid-infrared bands 5 (1.55-1.75 µm) and 7 (2.08-2.35 µm) contrast well, revealing differences in types and conditions of vegetation and soil. So a false color composite were created by using band combinations of band 4 in the green, band 5 in the red and band 7 (McHugh, 2006). After layer stacking, all the scenes were re-projected to UTM Zone 48 North using WGS 84 as a datum.

3.3.2 Geo-referencing of images

When the ground control points (GCPs) were overlaid on the color composite of the Lansat-5 images, it was observed that the points did not fit exactly with the corresponding ground features represented on the image. This triggered the need to undertake geo-referencing of satellite images, which was accomplished through the image-to-map geo-referencing operation of Environment for Visualization of Images (ENVI) that requires use of GCPs (satellite images to topographic map) which geographical accuracy is checked with the GCPs served as a base map and Landsat-5 as a warp image. Warp image refers to the one to be corrected using a geometrically corrected map, which will serve as a base map. At least four points are required for defining a warp polynomial so as to predict the corresponding locations of the selected GCPs in the warp image.

Both the base map and warp images were displayed side to side and effort has been made to minimize the overall registration error by relocating the

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

position of each GCPs within the particular ground feature selected as ENVI provides the flexibility, positional accuracy of both corrected Landsat (warped) images were checked visually by linking it to the corresponding topographic map that served as base map and using Google earth as well.

3.3.3 Satellite image classification

Lansat-5 images were earlier and very recent images available for study areas. Hence, it was possible to undertake field visit and collect GCPs. Supervised and unsupervised classification were preferred. The two landsat-5 TM images were also included to meet the preferred time horizon of the study, but due to lack of fine details for 1990, so unsupervised classification was selected to classify Landsat-5 TM image 1990. Meanwhile, it must be noted that effort has been made to integrate few historical information acquired from surveys to minimize complete reliance on spectral information and solve mystery of spectral similarity of different land use and land cover classes in order to improve classification accuracy. Unsupervised classification is more computer-automated. It allows user to specify parameters that the computer uses as guidelines to uncover statistical patterns in the data (ERDAS IMAGINE Tour Guides™, 1999).

For 2010, Lansat-5 image supervised classification was used. Supervised classification is more closely controlled by user than unsupervised classification. In this process, user select pixels that represent patterns recognize or that user can identify with the help from other sources. Knowledge of the data, the classes desired and the algorithm to be used is required before begin selecting training samples. By

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

identifying patterns in the imagery, users can train the computer system to identify pixels with similar characteristics. By setting priorities to these classes, user supervise the classification of pixels as they are assigned to a class value. If the classification is accurate then each resulting class corresponds to a pattern that user originally identified (ERDAS IMAGINE Tour Guides™, 1999).

3.3.4 Accuracy assessment of image classification

Accuracy of classification and Kappa coefficient error matrix were also determined based on classification result of images. Finally, the classified images were exported to ArcGIS for map preparation and to explain the situation of the spatial land use change of the study area. Accuracy assessment is a general term for comparing the classification to geographical data that are assumed to be true, in order to determine the accuracy of the classification process. Usually, the assumed-true data are derived from ground truth data. It is usually not practical to ground truth or otherwise test every pixel of a classified image. Therefore, a set of reference pixels is usually used. Reference pixels are points on the classified image for which actual data are (or will be) known. The reference pixels are randomly selected (Congalton 1991).

3.4 Logistic regression model for land use change analysis

Logistic regression is useful for situations where the dependent variable has a binary output, e.g. the presence or absence of a characteristic or outcome. The method is useful to predict the probability that a case will be classified into one as opposed to the other of the two categories of the dependent variable. Several transformations are

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

made to adequately deal with the binary structure of the dependent variable (Menard 2001).

The technique used in this study is logistic regression model, following Serneels and Lambin (2001). It is designed to estimate parameters of a multivariate explanatory model in situations where the dependent variable is dichotomous and the independent variables are continuous or categorical. The logistic technique yields coefficients for each independent variable based on a sample of data. These coefficients are interpreted as weights in an algorithm that generates a map depicting the probability of a specific category of land use change for all sampling units.

Logistic regression identifies the role and intensity of explanatory variables X_n in the prediction of the probability of one state of the dependent variable, which is defined as a categorical variable Y. Suppose X is a vector of explanatory variables and p is the response probability to be modeled with, in the case of a dichotomous dependent variable, p = Pr(Y = 1|X), with Y = 0 meaning the absence of change and Y = 1 meaning the presence of change. The linear logistic model has the form:

$$\text{Logit}(P) = \log\left(\frac{P}{1-P}\right) = \boldsymbol{\alpha} + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \ldots\ \beta_iX_i \quad \ldots\ldots\ldots\ldots\ldots\ldots \textbf{(1)}$$

Where α is the intercept. β_i are slope parameters and X_i are independent variables. The probability values can thus be quantitatively expressed in terms of explanatory variables by

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

$$P = \frac{\exp(\alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots \beta_i X_i)}{1 + \exp(\alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots \beta_i X_i)} \quad \ldots\ldots \textbf{(2)}$$

Odds ratios can be used to facilitate model interpretation (Stokes et al., 1995; Menard, 1995). The odds ratio ($\psi = e^{\beta i}$) is a measure of association which approximates how much more likely (or unlikely) it is for the outcome to be present for a set of values of independent variables (Hosmer and Lemeshow, 1989).

The odds ratio can be interpreted as the change in the odds for an increase of one unit in the corresponding risk factor (Menard, 1995).

$$\text{Odds (p)} = \exp(\alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots \beta_n X_n) \ldots\ldots \textbf{(3)}$$

In this case of the study, spatial logistic regression model was performed using the logistic function in IDRISI GIS software. The predictive ability of a logistic regression model is evaluated from the table of maximum likelihood estimates (MLE), which contains the MLE of the parameters, the estimated standard errors of the parameter estimates, the Wald χ^2 statistics and the significance probabilities for the parameter estimates. Positive values of the parameter estimate indicate that larger values of the explanatory variable will increase the likelihood of the occurrence of the event. Likewise, negative values of the parameter estimates indicate that larger values of the explanatory variable will decrease the likelihood of the occurrence of the event. The χ^2 statistic indicates the relative weight of each explanatory variable in the model and allows us to assess the role of each variable in the prediction of an event (Menard, 1995).

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

For the majority of the variables of land use change classes were organized. These classes will be aggregated into more general classes to allow an analysis of the main land use change processes over the time period (Pena et al, 2007), which allows identifying the transitions between land use classes. These variables were used in the logistic regression model to associate land use change with demographic, econometric and biophysical driving forces and to generate land use change probability map.

Dependent variable: P, 1 = land use change and 0 = no land use change.

Independent variables consist of two parts: spatial variables X_1= Elevation, X_2= Slope, X_3= Distance from road, X_4 = Distance from urban, X_5 = Distance from water and socio-economic variables X_1= HHA (Head of household age), X_2= ETH (Ethnicity), X_3= HHS (Household size), X_4= ED_1 (Secondary School), X_5= ED_2 (High School), X_6= ED_3 (Vocational or University), X_7= FL (Family labor), X_8= OL (Ownership of Livestock), X_9= OFIC (On Farm Income), X_{10}= NFIC (Non-Farm Income), X_{11}= RP (Rice Production) and X_{12} = LSH (Land size holding).

These independent variables were used to predict the probability of the land use change in the study area that cover of 3 villages which has 150 samples.

The affects of the government of Lao PDR's policies will be explained such as the policy to eliminate the cultivation of upland rice by means of 'slash-and-burn' cultivation and to replace it with more ecologically stable systems based on sustainable land use at the village and household level, policy to alleviate poverty and to introduce more sustainable management of agricultural resources.

The government has initiated a program of relocation from the upland to "focal areas" in the lowland which marketing, distribution and other services can be

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

supplied. Moreover, population density data for 1990 and 2010 for the study area were used to explain factors influencing land use change of the village and community level.

Well-being of the farmers in two periods was assessed by comparing farm and non-farm income as well as rice output for the two periods. Explanations of the changes in livelihood related to the changes of land use are made.

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved