The Use of GIS in Exploring Settlement Patterns of the Ethnic Groups in Nan, Thailand

Pannee Cheewinsiriwat

pannee.ch@chula.ac.th

Department of Geography, Faculty of Arts, Chulalongkorn University,

Bangkok 10330, Thailand.

Topic: Science, Environment and the Humanities

The Use of GIS in Exploring Settlement Patterns of the Ethnic Groups in Nan, Thailand

Pannee Cheewinsiriwat

pannee.ch@chula.ac.th

Department of Geography, Faculty of Arts, Chulalongkorn University,

Bangkok 10330, Thailand.

Abstract

Nan is a mountainous province located in the north of Thailand with diversity of ethnic groups and languages. Linguistics of the languages spoken by ethnic groups were studied by a previous geo-linguistics research, and maps to illustrate locations of the ethnic groups were made to visually express the spatial distribution of settlement patterns over Nan area. This paper continues this study by using Geographic Information System (GIS) to explore the settlement patterns in relation to geographic factors. Geographic data layers (i.e. village locations, a digital elevation model (DEM), roads and rivers) were input to GIS to carry out the research. Geostatistic functions were used to measure dispersion/clustered patterns of the settlement distribution, while proximity and overlay functions were used to find out spatial relationships between such patterns and geo-environmental factors. The general findings reveal that although people of the same ethnic group tend to live in clustered fashion, some groups tend to live in a more clustered fashion than others. In addition, ethnic groups located on an elevation of 200-500 meters above the mean sea level tend to live near both roads and rivers (within 1 kilometer), while those on higher elevations tend to live further away from roads and rivers. Furthermore, the unique characteristics of the settlement patterns of each ethnic group were also identified. The findings and the methodology introduced within this paper can be valuable as a knowledge-base for a variety of studies in ethnology-related areas as well as for the cultural tourism industry in Nan.

1. Introduction

Ethnic diversity has been a wide concern for countries all over the world such as the United States of America, Canada, United Kingdom, Indonesia, Malaysia, Singapore and Thailand. This is because while these countries have a number of native tribes which have lived there for several hundred years, they have also experienced immigration of people looking for better living conditions. As a result, multi-culture environments have grown within single societies such as a province, or a nation. In order to live within a multi-cultural society, knowing the characteristics of each other can lead to better understanding between ethnic groups.

Ethnic diversity has been studied in various aspects of social sciences and humanities such as demography, politics, and linguistics. However, one common purpose of these studies is to

identify the characteristics of ethnic groups such as population, spoken languages, beliefs, economic activities, and religions. Although the diversity distribution involves a spatial factor, this spatial dimension has rarely been included in ethnic diversity studies. Therefore, by including spatial dimension in an ethnic diversity study, geographers can relate the characteristics of ethnic groups, especially settlement patterns, to the spatial variation factor.

Instead of describing the settlement patterns by visual examination from paper maps, modern geography has introduced geographic information systems (GIS) to analyze geographical phenomena and produce quantified results to explain such phenomena. GIS has been widely known as a powerful toolset to analyse geo-spatial data, not only for the geography research area but also for other spatially-related research areas such as linguistics spatial diversity, epidemiology, and disaster management. The efficient and popular functionalities within GIS include overlay, 3D surface analysis, geostatistics and network analysis (Maguire and Batty 2005; Demers 2009).

This paper utilises GIS tools to investigate the characteristics of settlement patterns of ethnic groups and the relationship between these settlement patterns and the spatial dimension. The paper selected Nan province in Thailand as a study case for this research where at least thirteen ethnic groups are found (Cheewinsiriwat 2009). The investigation aims to achieve two aspects of the settlement characteristics: spatial distribution and spatial relationships. Additionally, it is also expected that GIS technology demonstrated in this paper can contribute to providing new insight for not only other ethnic diversity studies, but also other humanity studies.

In the next section, an introduction about GIS is provided with a focus on the understanding of the geographic technology used in this paper. The previous GIS-related work in settlements is then given in section 3. The procedure of using GIS technology, study area and datasets, and the results and analysis are then explained respectively in sections 4, 5, and 6. The main conclusions drawn from the paper are specified in section 7.

2. Geographic Information System

Geographic Information System (GIS) has emerged as a computer system which is capable of capturing, querying, analyzing and displaying geographic data (Clarke 1999; Demers 2009). Geographic data used in GIS consists of graphic or location data, and non-graphic or attribute data. For example, a village is represented by a pair of x,y coordinates (location data) and its village code (attribute data) as shown in Figure 1. Having the two types of data, GIS can thus either retrieve all attributes related to a number of specified locations of villages or highlight locations of villages which correspond to the given attributes. Map reproduction from digital GIS data also provides an easier and faster option for map makers to show spatial information on screen or papers. Since GIS data is normally stored in separate layers such as a road layer and a river layer, it is easy to select only required layers to display and symbolise features for appropriate visualisation. In terms of spatial analysis, GIS provides capability to analyse and produce valuable information for better decision making, which has a multitude of advantages over a wide range of GIS-related applications.

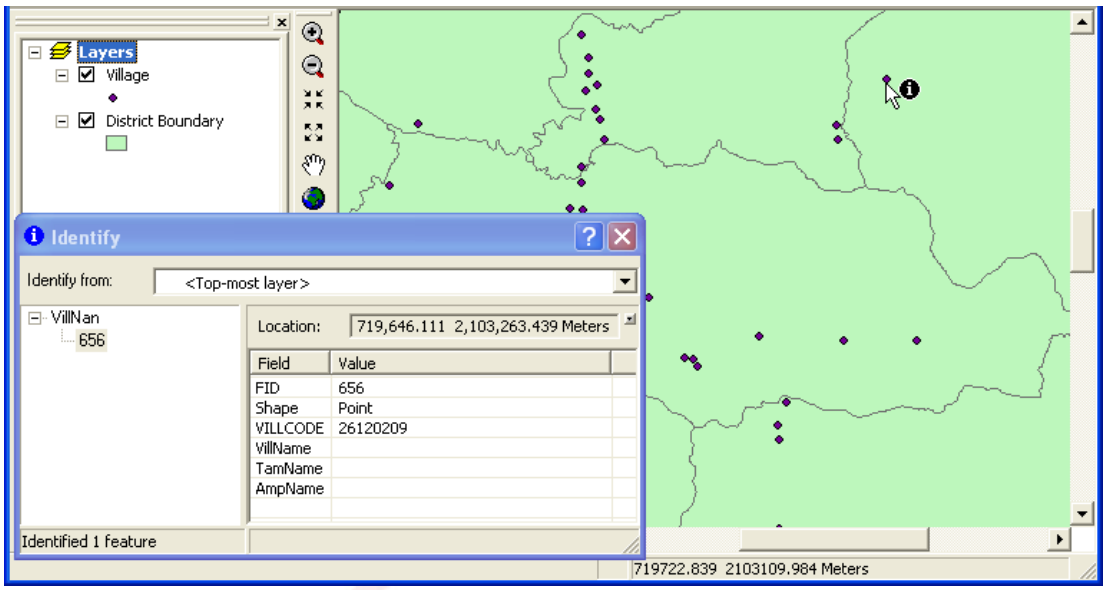
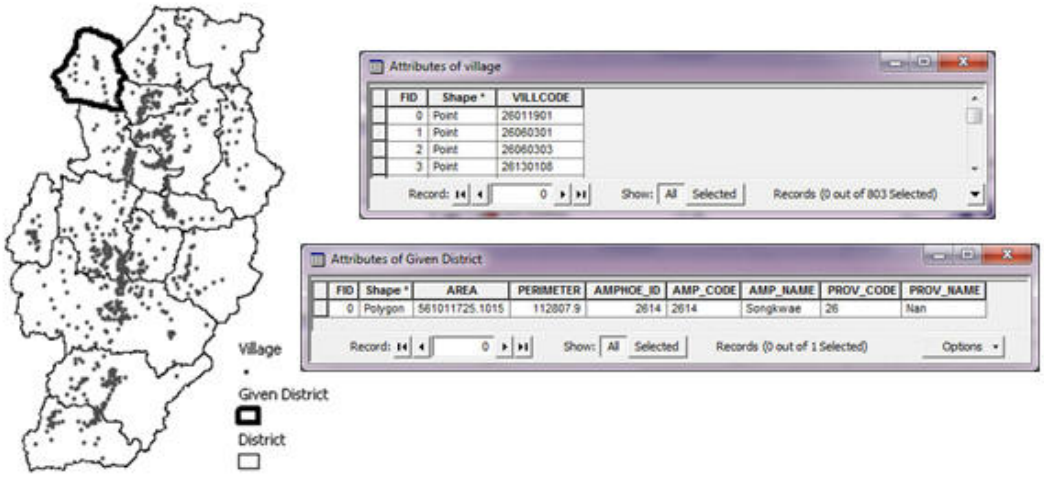


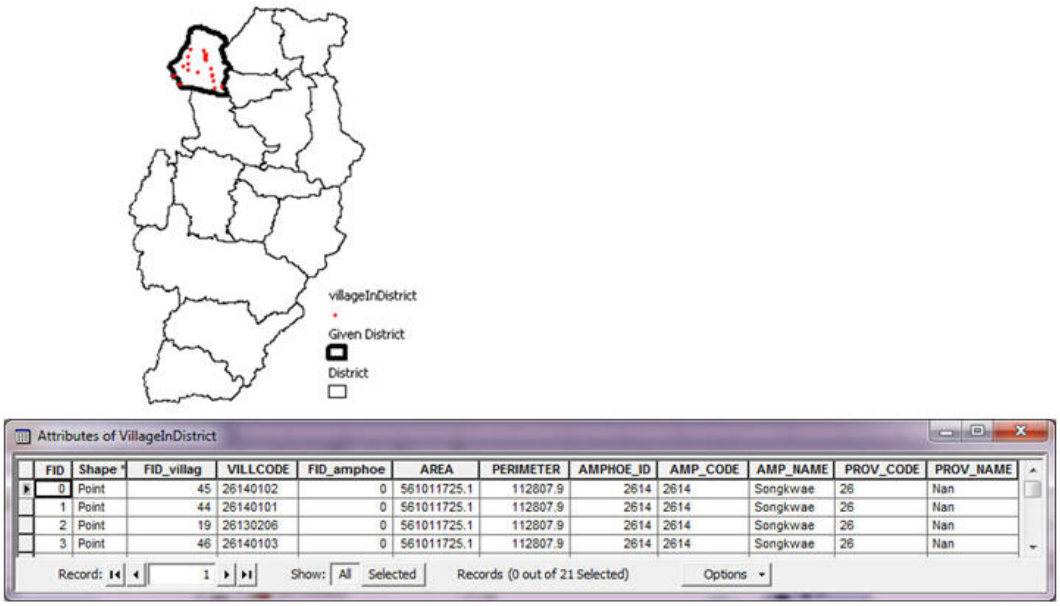
Figure 1 Locations of villages and the attributes of the pointed village.

In 1969, a manually valuable overlay technique was introduced by McHarg for evaluating social and environmental costs of land use change (McHarg 1992). The technique was to draw maps of each theme on separate transparencies where the darker shade means more important, and then overlay them to acquired synthesised information. Currently, the technique can be done easily and effectively by GIS. The overlay capability is now one of the popular GIS functions and considered as a basic powerful Boolean functionality which consists of intersection, union and difference. In GIS, data of the resultant layer from overlay functions is derived from both locations and attributes of the input two layers. For example, when intersecting villages in Nan province to a district of Nan province, the result is points of villages within the given district, with the attributes of the village and district layers (see Figure 2).

Displaying location data on maps is sometimes not adequate for decision making, therefore, statistical measurements on geographic data is required in order to perfectly explore and explain spatial phenomena. For example, propability test and multiple regression analysis have been used in population models to investigate the relationships between incidents, such as the number of people who died from cancer, which occurred in different areas (Jones 1990). Geostatistics functions have been included in GIS as an extended functionality to perform statistics on geographic data. Such functions have bridged statistics and geography to successfully analyse spatial phenomena. Statistical index of distribution patterns, spatial relationships prediction, probability prediction, and deterministic and geostatistical interpolation are examples of the geostatistics functions.



(a)



(b)

Figure 2 (a) Locations of villages and the given district with their attribute tables (b) The result of intersecting the village and the given district layers.

3. GIS-related Work in Settlement Exploration

Geographers study settlements in terms of their spatial distribution over the Earth's surface. The distribution means 1) dispersion type: compact, nucleated, and agglomerated, 2) density, and 3) patterns: rural, urban, and suburban. Both cultural and physical elements influence the

distribution of settlements as people basically choose favorable areas to suit their activities. In addition, relationships between settlement patterns and physical/cultural factors also have been studied.

The GIS overlay techniques and statistical redundancy analyses (RDA) were used to evaluate the environmental impacts caused by informal settlements (Zeilhofera and Topanotti 2008). The important impacts were then identified as the results. In addition to exploring settlements in the present time, GIS was also used to analyze and reconstruct settlement patterns in the past of North West America (Carlisle and Friedel 2009). The analysis started from using digital evaluation model (DEM) data and GIS weighted overlay cost surface technique to delineate the site exploitation territory.

In geography of Linguistics, settlements of the ethnic groups can be associated to their spoken dialects which attract geography-linguistics researchers to study their spatial variation (Heap 2006). Exploring spatial variation of dialects spoken by villagers in Nan province, Thailand was developed as an interactive GIS application for display, query and map reproduction (Cheewinsiriwat 2009). Maps showing settlement locations of ethnic groups were displayed to show the distribution of ethnic groups. Pie charts showing proportion of ethnic groups within a specified village, district, or province were also illustrated on the maps to show the variation of ethnic groups over the area (Cheewinsiriwat 2009). In an area of dialect geography, a GIS system was also developed to define boundaries of dialects spoken in Thailand by considering boundaries on where the same words are differently pronounced (Teerarojanarat and Tingsabadh 2008).

As mentioned, GIS has been used to analyse spatial relationships between thematic sorts of data as well as to express locations of interesting or useful data on maps. Valuable information can be thus derived from the analysis. In Nan province, spatial variation of spoken languages has been studied and resulted in maps showing where different languages are spoken using villages as the smallest unit of study (Cheewinsiriwat 2009). This paper conducted further study on the settlement preferences of ethnic groups in terms of relationships to physical factors as well as their behavior in tightly or loosely settlement distribution patterns over the area of Nan province.

4. Methodology

In order to reveal geographical characteristics of where the ethnic groups choose to live, analytical GIS functions were employed to derive spatial relationships between settlement locations and surrounding geographical factors, as well as the degree of spatial clustering of ethnic population. Four groups of GIS functions were used in this paper: surface interpolation, overlay, proximity and geostatistics. Combining the results from these functions provides the knowledge about the settlement patterns of ethnic groups in Nan, Thailand. Figure 3 presents the work flow of surface interpolation, overlay, and proximity functions.

4.1 Surface Interpolation

As Nan is a mountainous area, the elevation, slope, and aspect of where villages located are factors to be explored to observe the settlement preference related to Nan's topography. Elevation on maps is normally indicated by interval contour lines or regular grid points. Only

locations along the contour lines or on the grid points are locations with known elevation. Surface interpolation calculates the elevation for every location throughout the given area by estimating it from the elevations of the surrounding known locations. After the surface interpolation is completed, 3D surface throughout the area can be viewed and investigated. Several estimation techniques such as Spline and Kriging can be used to perform such interpolation. However, Kriging is one of the most popular techniques and was selected to carry out this task in this paper.

The process of deriving topology-related information from the contour lines is illustrated in Figure 3. The 100m-interval contour lines were used to create a Triangular Irregular Network (TIN) model from which elevation, slope, and aspect can be derived and 3D surface of Nan also can be viewed. However, the TIN model does not explicitly store the derived information. In order to use the derived information in finding out spatial relationships with villages' locations, the TIN model was therefore needed to convert to raster grids and then to vector polygon features. The results from this step were a 3D surface for visualization and the three layers of vector polygon features (elevation, slope, and aspect).

4.2 Overlay

Overlay functions are widely used among GIS applications. When two data layers are overlaid, the extracted information can be done by union, intersection or difference of the layers. To gain elevation, slope, and aspect of each village location, the intersection overlay function was applied on the villages' locations layer and the elevation, slope and aspect layers as shown in Figure 3. As the result, the attributes of elevation, slope and aspect were attached to the village locations layer.

4.3 Proximity

Proximity functions are involved in calculating the distance between different features within the same data layer or in other data layers. Since accessibility to transportation and water resources is one of the important factors of human settlements, distance from where they live to such resources can be used as accessibility indicators. In this paper, the nearest distances from villages' locations to rivers and roads were derived by the near function. The identifier (ID) of the nearest road line and river line together with the nearest distance were then attached to the village locations as shown in Figure 3.

4.4 GeoStatistics

Geographers are also interested in the distribution of human settlements. Geostatistics provide calculations to gain indices of the spatial distribution to indicate if they are clustered, random, or dispersed. The nearest neighborhood method was selected to perform the calculation in this paper. An index was calculated for each ethnic group to indicate whether their villages's locations are clustered, random, or dispersed.

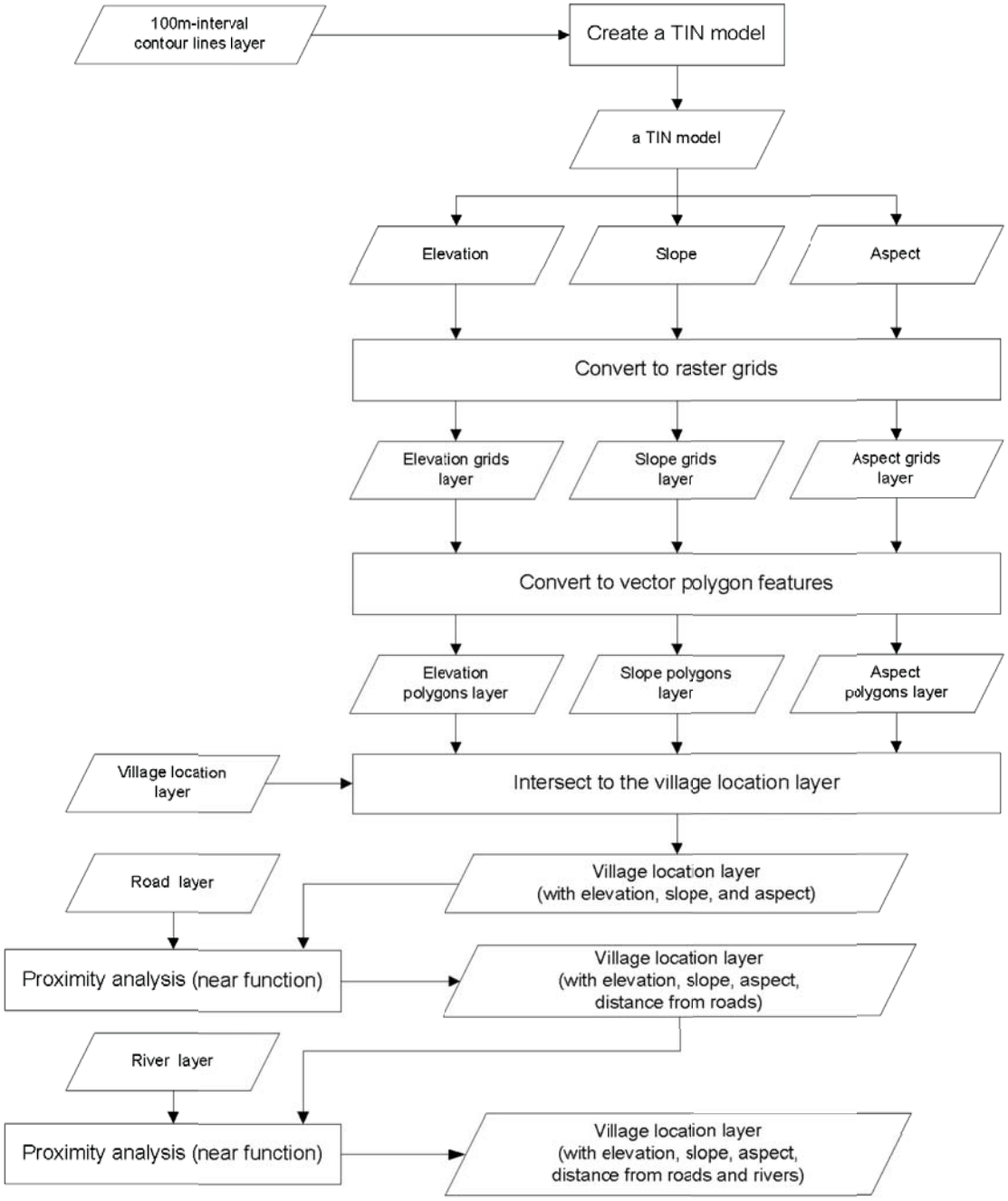


Figure 3 The process of deriving required information from contour lines, roads, and rivers in Nan study area.

5. Study Area and Dataset

The area of interest for applying GIS to analyse the settlement patterns covers the whole province of Nan, which is around 12,200 square kilometers. Nan is located in the northern region of Thailand, and its eastern mountainous border is adjacent to the Lao People's Democratic Republic (Lao LPR). Nan consists of fifteen districts which are divided into ninety-nine sub-districts, in which 902 villages are located (Cheewinsiriwat 2009). As the northern part of Thailand has borders with Burma and Lao PDR and is close to China, provinces along the border normally have a number of ethnic groups. However, Nan is one of the provinces that has the most number of ethnic groups, at least thirteen, living in its provincial area.

This paper concentrates on seven selected major ethnic groups: Maung, Tai Lue, Lao, Mong, Yao, Tin (Prai), and Tin (Mal). A village with people of more than one ethnic group is represented by one point for each ethnic group. Population and numbers of village locations of each ethnic group are listed in Table 1.

No.	Ethnic Group	Population	Number of Villages
1	Maung	296,758	608
2	Tai Lue	50,324	81
3	Lao	9,672	14
4	Mong	21,954	17
5	Yao	10,347	14
6	Tin (Prai)	10,884	26
7	Tin (Mal)	24,450	52

Table 1 Numbers of people and village locations of the selected seven ethnic groups in Nan.

The digital dataset used is obtained from the previous geo-linguistics research project, funded by the Thailand Research Fund (TRF). The dataset consists of five shapefile data layers consisting of village locations of the seven ethnic groups, contour lines, road centerlines, river centerlines, and administrative boundary lines. The village location layer was then separated into seven layers, one for each ethnic group. The village locations are represented as point features with an attribute of numbers of population of each ethnic group. Contour lines, road centerlines, and river centerlines are represented as line features. The administrative boundary lines are represented as polygon features with attributes of area codes and names**.**

6. Results and Analysis

In order to obtain the topography characteristics of each village, the contour line shapefile was firstly input into the interpolation process to produce the surface elevation, slope, and aspect data layers. The village layer was then overlaid on the three layers one at a time as shown in Figure 4, 6 and 8, and the attributes of elevation, slope, and aspect were transferred to village points by the intersection function. The elevation, slope and aspect of where ethnic groups are located, based on the villages' locations, were plotted as shown in Figure 5, 7 and 9 respectively.

It was found that most of the villages occupied by the local Thai ethnic group, called Maung, are scattered over Nan on areas with elevation between 200-500 meters from mean sea level (MSL).

Tai Lue and Lao ethnic groups also live on the areas below 500 meters from MSL but only in the upper part of Nan. Villages of Yao are linearly scattered along a north-south direction in the north-west of Nan mostly below 600 meters from MSL, while villages of Mong and Tin (Prai) are on the higher land up to 1,000-1,300 meters from MSL. Most villages of Tin (Mal) are on the areas higher than the others, which is up to 1,600 meters from MSL in the mountainous area in the north-east.

Although considered as a high land and mountainous province, most areas of Nan province have slopes less than 20 degrees as shown in Figure 6. The higher slopes are located in the areas along the provincial boundary especially along Thailand-Lao PDR border, while the lower slopes are located in the inner part of the province. As shown in Figure 7, most villages of Maung, Tai Lue, and Lao are located on flat areas, while more than half of the others; Mong, Yao, Tin (Prai), and Tin (Mal), are on the areas up to 30 degrees slopes. However, only four villages of the seven ethnic groups are located on more than 20 degrees slopes.

Apart from slope, the difference between the elevations of two or more locations also provides aspect. Aspect describes the direction to which the facets of surface are facing. The values of aspect are between 0 and 359 degrees in a clockwise direction, where 0 means facing north direction. The villages' locations were plotted on top of the aspect of Nan area as shown in Figure 8, while percentages of people of each ethnic group who live in the given classes of aspect are shown in Figure 9. The villages of Maung, Tai Lue, and Lao are mostly located on flat areas, so no aspect is given in these areas. Mong villages are on flat areas only 34%, while 31% and 21% are on the aspect of north-east and south-west respectively. About 38% of Yao villages are on flat areas, while 51% are located facing east. As Yao are mostly located on the west side of Nan, facing to the east means facing in-bound to Nan territory. Only about 30% of the villages of Tin (Prai) and Tin (Mal) are located on flat areas. Apart from this number, Tin (Prai) tend to live facing north-west direction more than other directions, while Tin (Mal) live in north-west and south-west directions. The similarity of Tin (Prai) and Tin (Mal) is that they live facing west much more than facing east. As Tin (Prai) and Tin (Mal) are located on the east side of Nan, facing west means facing in-bound to Nan territory.

Proximity analysis was performed to measure distances from the village locations to the nearest roads and rivers. This helps understanding their way of life in terms of transportation access and/or water resources for agriculture. As shown in Table 2, it is evidently seen that most people of Maung, Tai Lue, and Lao, who live in flat areas, are within a distance of 1 kilometer from both rivers and roads, while only 32.5% of Mong and 27.1% of Yao live within 1 kilometer from rivers and roads. Among the seven ethnic groups, Mong have the highest number of people (33.3%) who live far from both rivers and roads with a distance of more than 1 kilometer, while Yao have the second highest of 23.4%. About 85-90% of Tin (Prai) and Tin (Mal) people live near roads (within 1 kilometer), and only small numbers of them live further than 1 kilometer from both rivers and roads.

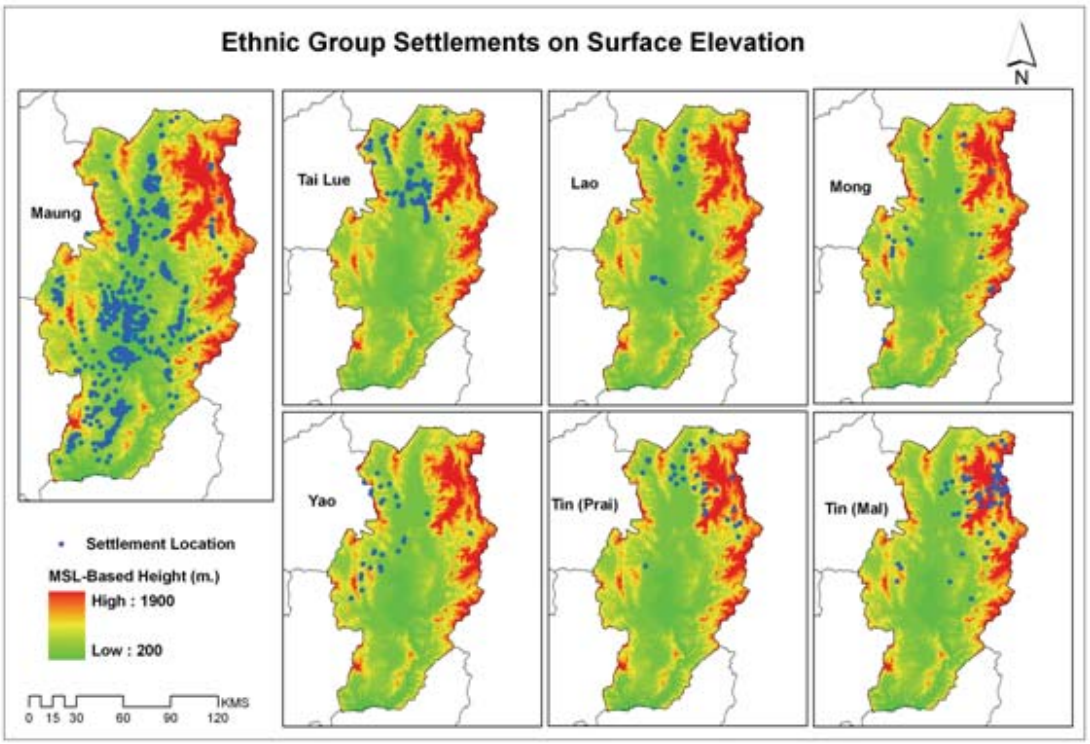


Figure 4 The villages of the seven ethnic groups overlay onto the surface of elevation.

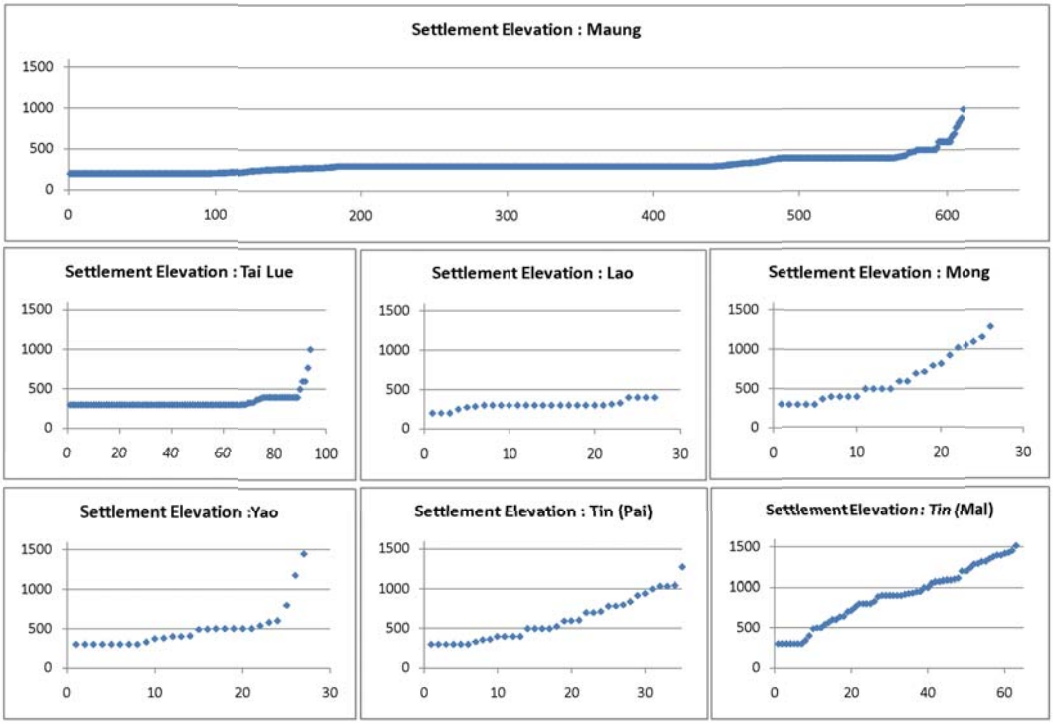


Figure 5 Elevation of villages by ethnic groups, where x-axis is the elevation in meters from MSL, and y-axis is number of individual villages.

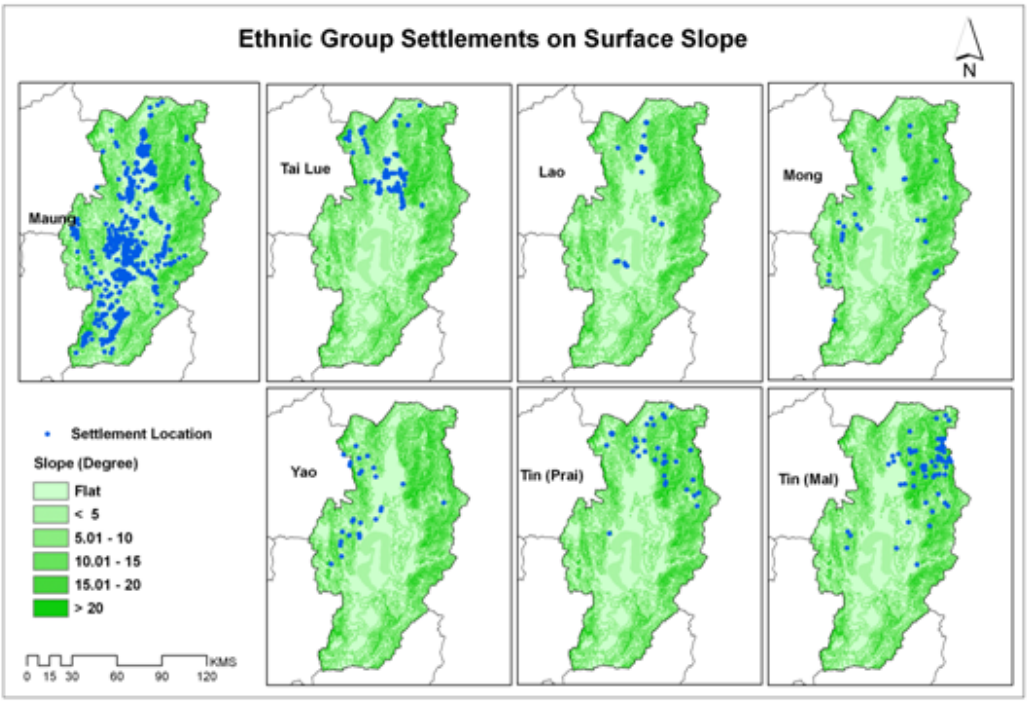


Figure 6 The villages of the seven ethnic groups overlay onto slope.

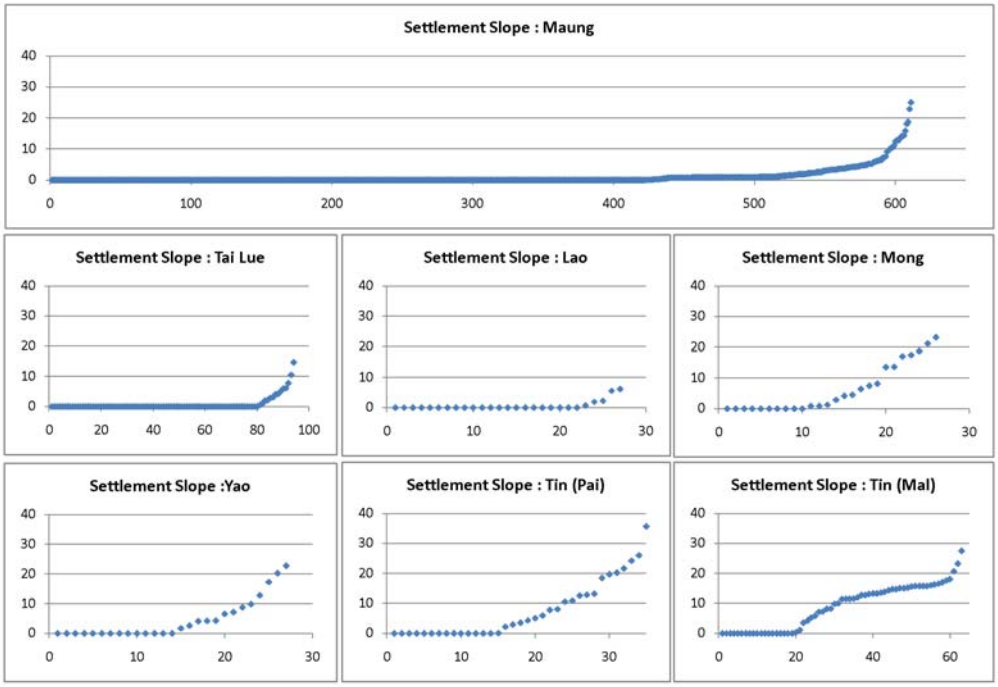


Figure 7 Slope of villages by ethnic groups, where x-axis is the degree of slope, and y-axis is number of individual villages.

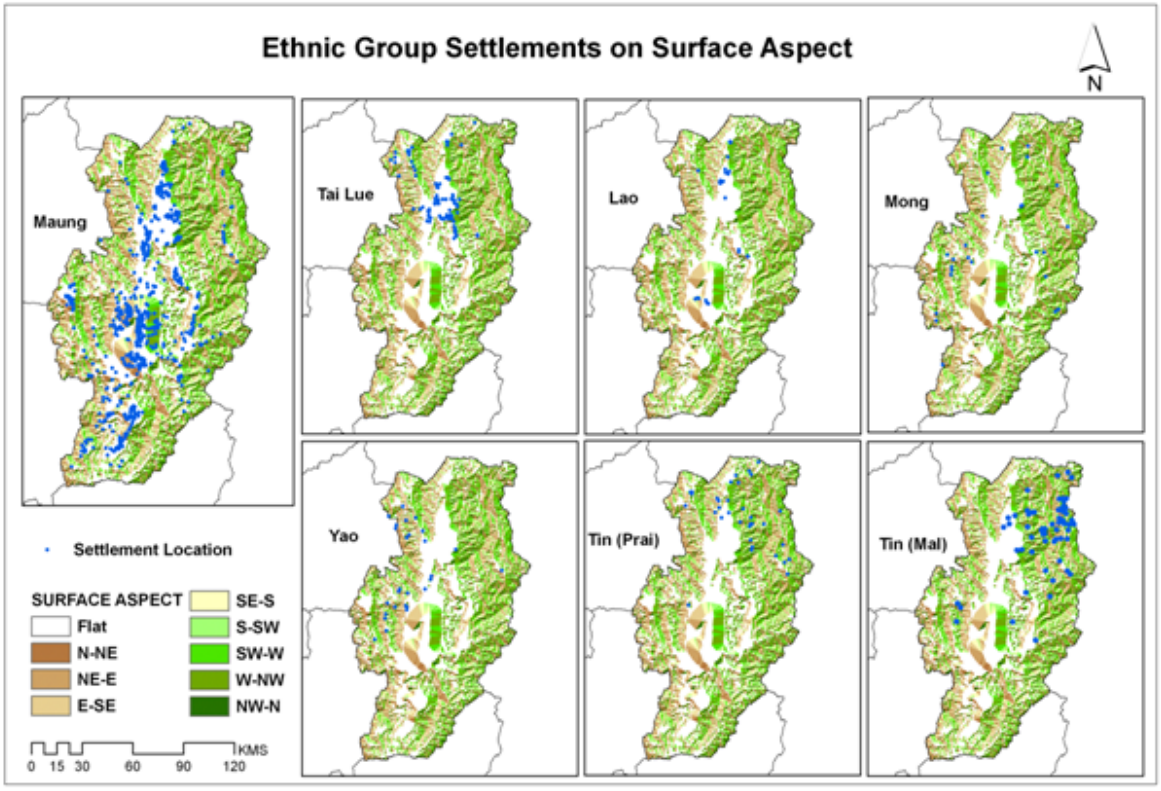


Figure 8 The villages of the seven ethnic groups overlay onto aspect.

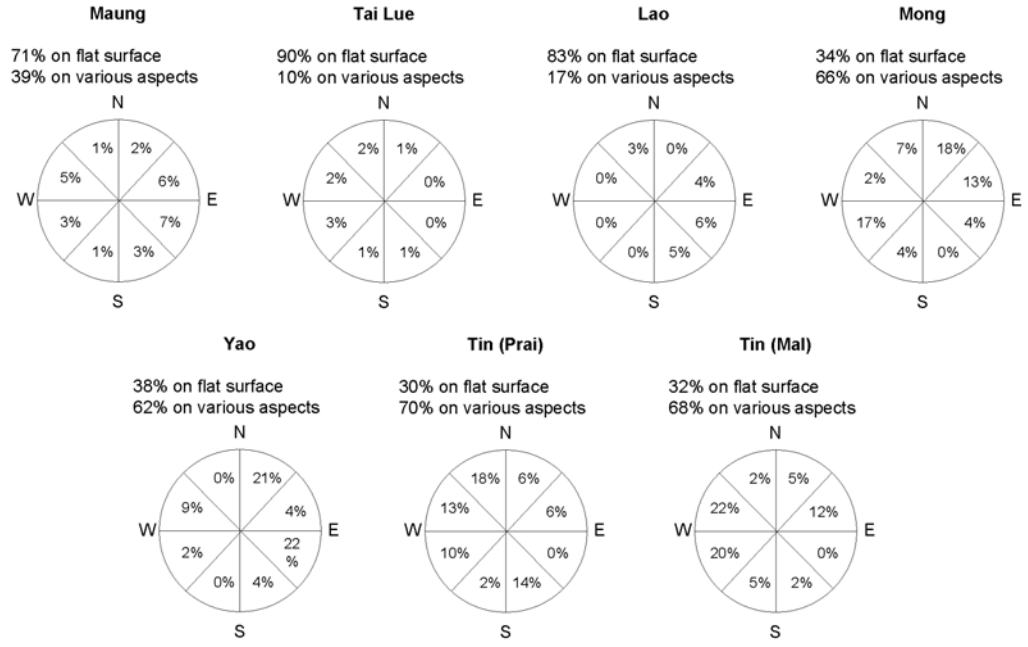


Figure 9 Aspect of villages in percentage by ethnic groups.

No	Ethnic Group	<= 1Km from Rivers				> 1 Km from Rivers			
		<= 1Km from Roads		> 1 Km from Roads		<= 1Km from Roads		> 1 Km from Roads	
		Population	%	Population	%	Population	%	Population	%
1	Maung	240,679	81.1	18,756	6.3	28,797	9.7	8,526	2.9
2	Tai Lue	44,859	89.2	2,239	4.5	2,016	4.0	1,164	2.3
3	Lao	9,642	99.7	0	0.0	30	0.3	0	0.0
4	Mong	7,145	32.5	3,629	16.5	3,874	17.6	7,306	33.3
5	Yao	2,805	27.1	4,807	46.5	316	3.1	2,419	23.4
6	Tin (Prai)	5,794	53.2	1,031	9.5	3,861	35.5	198	1.8
7	Tin (Mal)	14,533	59.4	2,253	9.2	6,003	24.6	1,661	6.8

Table 2 Numbers of population living classified by distance from their villages to rivers and roads.

Although the distribution of villages of each ethnic group can be seen on maps visually, the type of distribution, being equal or not, is difficult to be visually examined. However, Geostatistics help to measure how these villages distribute spatially, and then return indices of the distributions. The nearest neighbor method was used to compute the distribution indices of each ethnic group. The index of 1 means randomly distributed. The index of less than 1 exhibits clustering, while the index of more than 1 exhibits dispersion. After applying the nearest neighbor method on villages' locations of each group, the resultant distribution indices are measured and listed in Table 3. Lao and Tai Lue have the most clustered distribution among the seven ethnic groups in Nan, with an index of 0.19 and 0.28 respectively. Maung, Yao, Tin (Prai) and Tin (Mal) have a clustered distribution at the same level with indices of 0.5-0.6, while Mong have the least clustering level (0.72) among the others. The resultant indices show that every ethnic group tends to settle close to the same ethnic groups but with slightly different degrees of clustering.

No	Ethnic Group	No of Villages	Index	Patterns
1	Maung	608	0.54	Clustered
2	Tai Lue	81	0.28	Clustered
3	Lao	14	0.19	Clustered
4	Mong	17	0.72	Clustered
5	Yao	14	0.52	Clustered
6	Tin (Prai)	26	0.57	Clustered
7	Tin (Mal)	52	0.53	Clustered

Table 3 Geostatistical results of nearest neighborhood analysis showing degrees of the settlement distributions.

Combining the results explained above, it has been found that Maung, Tai Lue, and Lao ethnic groups normally live on flat plain areas near roads and rivers. On the contrary, only about one third of Yao, Mong, Tin (Prai), and Tin (Mal) live on flat plains, while the rest spread over areas of up to 30 degree of slopes and also higher elevation. Yao prefer living near rivers compared to roads, while Tin (Prai) and Tin (Mal) prefer roads to rivers. It is shown that one-third of Mong live further from both roads and rivers. This implies that Mong have the least accessibility to both transportation and water resources. Therefore, Mong might be considered as one of the most

conservative ethnic groups in Nan. In terms of Aspect, most Yao villages are located facing to the east, while most Tin (Prai) and Tin (Mal)'s are facing to the west. Mong villages are facing north-east and south-west. In terms of the village locations in Nan, the aspect of the four ethnic groups are all facing to the flat plain of inner Nan.

7. Conclusion

Diversity of ethnic groups in Nan has attracted cultural tourists to visit as well as researchers to study various aspects of how different ethnic groups live within a single province. How they settle themselves over the area in relation to the topography of Nan is an aspect that interests geographers. This paper has demonstrated GIS, a modern geography technology, in studying and explaining the settlement patterns of the selected seven ethnic groups in Nan. Apart from introducing GIS technology, knowledge about the settlement patterns of ethnic groups in Nan was also disclosed. As described in the previous section, the seven ethnic groups have different preferences in elevation, slope, aspect, accessibility to transportation and water resource. Besides, people of the same ethnic group tend to live close to each other but with slightly different degrees of clustering. The findings and methodology introduced within this paper can be valuable as a knowledge-base for a variety of studies in ethnology-related areas as well as for the cultural tourism industry in Nan. In addition, this paper has shown that geographic technology can support not only scientific studies but also humanities studies. As currently there are quite a few humanities applications using GIS, it is expected that the methodology introduced in this paper will be applied to more humanities applications in the future.

Acknowledgements

The author would like to express her appreciation to Faculty of Arts, Chulalongkorn University for financial support of this research project. The author gratefully acknowledges the support of the "Linguistic Diversity in Nan Province: A Foundation for Tourism Development" project, which is sponsored by the Thailand Research Fund (TRF), and Professor Dr. Theraphan L-Thongkum, the research project director, for the valuable digital maps and associated data. The author wishes to express personal appreciation to Associate Professor Pongsri Chanhow and Associate Professor Narote Palakawongsa for their valuable consultancy in geography concepts. Finally, the author also wishes to kindly thank Ajarn Captain William S. Whorton for the proof-reading throughout this paper.

References

Carlisle, K. and R. Friedel (2009). Reconstructing the Past: A GIS Settlement Pattern Analysis for Corths est Corth America. Spring 2009.

Chees insiris at, P. (2009). "GIS Application for the h aps of Tourist Attractions and Ethnic Groups of Can Proeince, Thailand." h anusya a2(2): a9CC0.

Clari e, K. C. (a999). Getting Started s ith Geographic Information Systems. r .S.A., PrenticeCall, Inc.

Cemers, h . C. (2009). Fundamentals of Geographic Information Systems. r .S.A., KohnCC iley & Sons.

Ceap, C. (2006). Spatial Variation (Geolinguistics). Encyclopedia of Language & Linguistics: 6a2C6a6.
Kones, C. R. (a990). Population geography London, Paul Chapman Pu, .
h aguire, C. K and h . Batty (2009). GIS, spatial analysis, and modeling Redlands, California., ESRI Press.
h cCarg, I. L. (a992). Cesign s ith nature Ces i ori , Kohn C iley & Sons.
Teeraroeanarat, S. and K. Tingsa, adh (2008). "Ceeelopment of Geographical Cata, ase of LeFical items in Thai Cialects r sing Geographic Information System." Kournal of Language and Linguistics 26(2): aC20.
Ceilhofera, P. and V. P. Topanotti (2008). "GIS and ordination techniques for eealuation of eneironmental impacts in informal settlements: A case study from Cuia, t, central Brazil " Applied Geography 28(a): aCa9.