

doi: 10.14456/jiskku.2021.19

Grouping of Land Use Data based on Mapping
Clusters' Techniques in The Case of Muang
Nakhonratchasima District, Thailand

การจัดกลุ่มข้อมูลการใช้ที่ดินด้วยเทคนิคการจัดกลุ่มเชิงพื้นที่
ในกรณีศึกษา อำเภอเมืองนครราชสีมา ประเทศไทย

Yaowaret Jantakat^{1*}, Pongpun Juntakut², Pradeep Kumar Shrestha³

เยาวเรศ จันทะคัต^{1*}, พงศ์พันธุ์ จันทะคัต², Pradeep Kumar Shrestha³

*Corresponding author email: yaowaret.ja@rmuti.ac.th

Received: August 18, 2021

Revised: October 11, 2021

Accepted: October 15, 2021

Abstract

Purpose of the study: The objective of this study is to group statistical significantly the land using grouping technique for mapping cluster technique in Muang Nakhonratchasima District.

Methodology: The investigation focuses on the land use grouping using two techniques for cluster mapping in ArcMap program: 1) Cluster and Outlier Analysis (Anselin Local Moran's) and 2) hot spot analysis (Getis-Ord-Gi)

Main Findings: The findings reveal that the cluster analysis is appropriate for grouping urban and housing areas while the outlier analysis is good for agricultural and dwelling ground. Only contiguity edges and corners with Euclidian distance are recommended for clustering since it can group as many large areas as possible. Furthermore, hot spot analysis is suitable at various confidence levels for examining urban and dwelling areas and agricultural lands. Inverse distance (square)-based Euclidian and Manhattan distance is suggested for hot spot clustering

¹ Department of Information and Communication Technology, Faculty of Sciences and Liberal Arts, Rajamangala University of Technology Isan; สาขาเทคโนโลยีสารสนเทศและการสื่อสาร คณะวิทยาศาสตร์และศิลปศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน

² Department of Water Resource, Faculty of Civil Engineering, Academic Division of Chulachomkhalo Royal Military Academic; สาขาแหล่งน้ำ คณะวิศวกรรมโยธา ส่วนวิชาการของโรงเรียนนายพระจุลจอมเกล้า

³ Department of Civil Engineering, Pulchowk Campus, Tribhuvan University, Kathmandu, Nepal

at high confidence level because it can group as many areas as possible at the greatest confidence level too.

Applications of this study: Findings of this study can be utilized for the spatio-temporal analysis and land use planning of the city of Nakhonratchasima to enable the land use operation more systematically and at the maximum benefit to the public.

Keyword: Land use, Spatial clustering analysis, Technique of cluster and outlier analysis, Technique of hot spot analysis

บทคัดย่อ

วัตถุประสงค์: งานวิจัยนี้มีวัตถุประสงค์ เพื่อจัดกลุ่มการใช้ที่ดินทางสถิติอย่างมีนัยสำคัญ โดยใช้เทคนิคการจัดกลุ่มสำหรับการทำแผนที่ในพื้นที่อำเภอเมืองนครราชสีมา

วิธีการศึกษา: การศึกษานี้เน้นการจัดกลุ่มการใช้ที่ดินด้วยสองเทคนิคของการจัดกลุ่มสำหรับการทำแผนที่ในโปรแกรม ArcMap คือ (1) เทคนิคการวิเคราะห์แบบการจัดกลุ่มและที่ไม่สามารถจัดกลุ่มได้ (หลักการของ Anselin Local Moran's) และ (2) เทคนิคการวิเคราะห์หาพื้นที่การใช้ที่ดินที่มีการจัดกลุ่มมากที่สุด (หลักการของ Getis-Ord G_i^*)

ข้อค้นพบ: ผลการวิเคราะห์การใช้เทคนิคการวิเคราะห์แบบการจัดกลุ่มและที่ไม่สามารถจัดกลุ่มได้ พบว่าการวิเคราะห์แบบการจัดกลุ่ม เหมาะสำหรับพื้นที่ชุมชนและสิ่งปลูกสร้าง และการวิเคราะห์ที่ไม่สามารถจัดกลุ่มได้ เหมาะสำหรับพื้นที่เกษตรและหมู่บ้าน ซึ่งวิธีของขอบต่อเนื่องกันเท่านั้นและมุมที่มีระยะห่างแบบยุคลิดได้ถูกแนะนำเพราะสามารถจัดกลุ่มพื้นที่การใช้ที่ดินได้มากที่สุดเท่าที่เป็นไปได้ สำหรับการวิเคราะห์จุดที่รวมกลุ่มมากที่สุดที่ระดับความเชื่อมั่นต่างๆ พบว่ามีความเหมาะสมกับพื้นที่ชุมชนและสิ่งปลูกสร้างและเกษตร ซึ่งวิธีของระยะทางผกผันทั่วไปหรือแบบกำลังสองที่มีระยะห่างแบบยุคลิดและแมนฮัตตันได้ถูกแนะนำสำหรับการจัดกลุ่มที่รวมกันมากที่สุดหรือการหาจุดร้อนที่ระดับความเชื่อมั่นที่สูงเพราะวิธีนี้สามารถจัดกลุ่มพื้นที่การใช้ที่ดินได้มากที่สุดเท่าที่เป็นไปได้ที่ระดับความเชื่อมั่นที่สูงมากที่สุดด้วย

การประยุกต์ใช้จากการศึกษา: ผลของการจัดกลุ่มการใช้ที่ดินนี้จะนำไปใช้ในการวิเคราะห์เชิงพื้นที่และเวลาและการวางแผนการใช้ที่ดินในเขตเมืองนครราชสีมาให้เกิดประโยชน์สูงสุดอย่างเป็นระบบต่อไป

คำสำคัญ: การใช้ที่ดิน การวิเคราะห์การจัดกลุ่มเชิงพื้นที่ การวิเคราะห์แบบการจัดกลุ่มและที่ไม่สามารถจัดกลุ่มได้ การวิเคราะห์หาพื้นที่การใช้ที่ดินที่มีการจัดกลุ่มมากที่สุดหรือจุดร้อน

Introduction

From the report of land use in Thailand year 2017-2018 (Land Development Department, 2019), there are the first highest percent of an agricultural area (55.73%) and other land use areas such as forest area (32.63%), followed by urban and built-up area (5.85%), water bodies (2.92%) and miscellaneous area (2.86). Especially, forest area is still not meeting the goal set in the 20-year national strategy (2018-2037) (Office of the National Economic and Social Development Council, 2020) that has determined preserving and increasing green areas that are environmentally friendly to cover at least 55% with creating and developing green spaces for recreation and nature, restoring river and canal, participating and promoting community forests and mapping and database of green areas. The study of land use and land cover helps to develop land system science (Verburg et al. 2015) that constitutes the terrestrial component of the Earth system and encompass all processes and activities related to the human use of land, including socioeconomic, technological and organizational investments and arrangements, as well as the benefits gained from land and the unintended social and ecological outcomes of societal activities (Verburg et al., 2013). In addition, land is an essential natural resource, both for the survival and prosperity of humanity, and for the maintenance of all terrestrial ecosystems (Food and Agriculture Organization of the United Nations, 1992). In the present time, Geospatial technology (GT) has been widely used in the study of land resources. GT plays a key role in realizing the sustainable management of land resources through their core potential, which includes mapping through the collection of thematic and quantitative baseline resources data, measuring through more rigorous mapping process by quantifying and documenting the historical in a geographic format, and monitoring through regular assessment of the conditions by recording the changes in natural phenomena and human-induced activities (Obi Reddy, 2018). There are various related papers that have done research on the application of geospatial technology for studying land resources and planning land use such as a landscape classification map and land use monitoring (Carlier et al., 2021) and spatial cluster analysis of agricultural land conversion (Susilo, 2016) etc. Land use clustering refers to land use patterns in which related activities are located close together, usually within convenient walking distance (Victoria Transport Policy Institute, 2011). Land use was analyzed by techniques of spatial clustering and distance measures that help planning land use development (Smith, 2011). There are various related papers that show spatial clustering analysis, e.g., using hot spot analysis and cluster and outlier analysis for defining the

groupings of rural accommodation (Sanchez-Martin et al., 2019) using Moran's I and hot spot analysis for identifying clusters of clear-cuts and partial harvests (Oregon State University, 2016), etc.

This paper aims to propose grouping of land use data-based mapping clusters' techniques in the case of Muang Nakhonratchasima (MN) district, Thailand. This study focuses on land use clustering with two techniques of mapping clusters in ArcMap: (1) cluster and outlier analysis (Anselin Local Moran's I) and (2) hot spot analysis (Getis-Ord G_i^*). Such two techniques were analyzed together with conceptualization of spatial relationships and distance methods. Moreover, this study attempted to use the land use data of the Department of Land Development (LDD) which is a national agency in Thailand to take and produce useful spatial information in terms of use as a basis for planning and development of local land use further, especially rehabilitation and increasing of green areas or forest land. These green areas would be able to reduce stress on society due to the COVID-19 pandemic because the reports of WHO (2016; 2017) have confirmed that living near the green or natural areas such as parks or playgrounds can significantly reduce the incidence of mental health problems.

Literature review

Spatial cluster analysis played an important role in quantifying geographic variation patterns that were commonly used in disease surveillance, spatial epidemiology, population genetics, landscape ecology, crime analysis, and many other fields, but the underlying principles were the same (Jacquez, 2008). However, these literature reviews were focused on the related papers that used a Geographical Information System (GIS) and spatial statistics as Table 1. Two specific mapping clusters' techniques were used: outlier analysis (Anselin Local Moran's I) and hot spot analysis (Getis-Ord G_i^*). These two-GIS techniques are selected because there are technique of spatial relationships (inverse distance, inverse distance squared, fixed distance band, zone of indifference, contiguity edges only and contiguity edges corners) and two distance methods (Euclidean and Manhattan) as same in Arc Toolbox of ArcGIS program. Importantly, such techniques and methods can give answers to the objectives of this work.

Table 1 The related researches for this paper (Continued)

No.	Titles	Important details	Authors (years)
1.	A quantitative framework for analyzing long term spatial clustering and vegetation fragmentation in an urban landscape using multi-temporal landsat data	<ul style="list-style-type: none"> - This study aimed to show Local the utility of Indicators of Spatial Autocorrelation (LISA) indices-based multi-temporal landsat data and in identifying key hot spot and cold spots for urban landscape. The obtained results were used for study urban vegetation patterns. - Getis-Ord Gi* statistic in hot spot analysis was for identifying key hot spot and cold spots for urban landscape. 	Kowe et al. (2020)
2.	Hot Spot Analysis versus Cluster and Outlier Analysis: An Enquiry into the Grouping of Rural Accommodation in Extremadura (Spain)	<ul style="list-style-type: none"> - The purpose of this paper was to define the grouping of rural accommodation-based GIS and spatial statistics in Extremadura, Spain, especially if these were conceptualized by dint of their lodging capacity. The obtained results were used for devising tourist policies and strategic plans in order to improve the management and efficiency of each business. - Two specific techniques were used: hotspot analysis and outlier analysis, which yield results that proved the existence of homogeneous and heterogeneous groups of accommodation businesses, based not only on their spatial proximity but also on their lodging capacity. 	Sanchez-Martin et al. (2019)
3.	Spatiotemporal change for agricultural distribution from local administrative to provincial scales-based spatial clustering analysis	<ul style="list-style-type: none"> - This study aimed to spatial-temporal clustering analysis (2007-2015) of agricultural areas and types in 37 locally administrative and province scales. The obtained results were used for planning connection between produced urban and provincial food-based agriculture types. 	Jantakat (2019)

Table 1 The related researches for this paper (Continued)

No.	Titles	Important details	Authors (years)
		<ul style="list-style-type: none"> - Two mapping clusters techniques (cluster and outlier analysis (Anselin Local Moran's I) and hot spot analysis (Getis-Ord Gi*)) were used based on fixed distance band for analyzing spatial relationships and Euclidean distance for analyzing distance relationships 	
4.	Map analysis and spatio statistic: assessment of spatial variability of agriculture land conversion at urban fringe area of Yogyakarta	<ul style="list-style-type: none"> - The aim of this study was assessment of the spatial variability of agricultural land conversion at urban fringe area of Yogyakarta. The obtained results were used for planning agricultural land and its related factors. - Using Getis-Ord Gi* statistic in hot spot analysis was for identifying area of agricultural land conversion in a village and its surrounding. 	Susilo (2016)
5.	Cuts and Partial Harvests using Moran's I and Hot Spot Analysis.	<ul style="list-style-type: none"> - The aim of this study was grouping of clear-cuts and partial harvests using spatial clustering analysis in 4 cities (Detroit, Sweet home, McKenzie River and Middle Fork). The obtained results were used for managing 4 land use types (non-forest land, matrix land, adaptive management area, and congressionally withdrawn and late successional reserve). - Moran's I and hot spot analysis in ArcGIS were used for distribution of high and low magnitude disturbances is interesting, especially when overlaid on the land use designations. 	Oregon State University (2016).

Study area and method

Study area and land use

We have selected Muang Nakhonrthasima (MN) district as the study area which is one of 32 districts in Nakhon Ratchasima province of Thailand. For reasons, MN district is the center point or link of districts in the economy, transportation, important government offices, education, tourism, public health, etc. Thus, this study requires understanding land use data that is spatially learnt by GIS-based spatial clustering analysis. In addition, MN is the location of the city hall and various government agencies and is the center of prosperity of the province. MN is in northern east of Thailand, named “Isan region.” Geographically, MN has an area of 765.49 sq.km. and locates from 14° - 17°N to 101° - 103°E as shown in Figure 1 and land use data of year 2019 from LDD is shown in Table 2. Agricultural land covered by the highest proportion (54.99%) while the other lands were urban and built-up land (3.60%), miscellaneous land (8.30%), water body (2.06%), and forest land (1.05%) respectively.

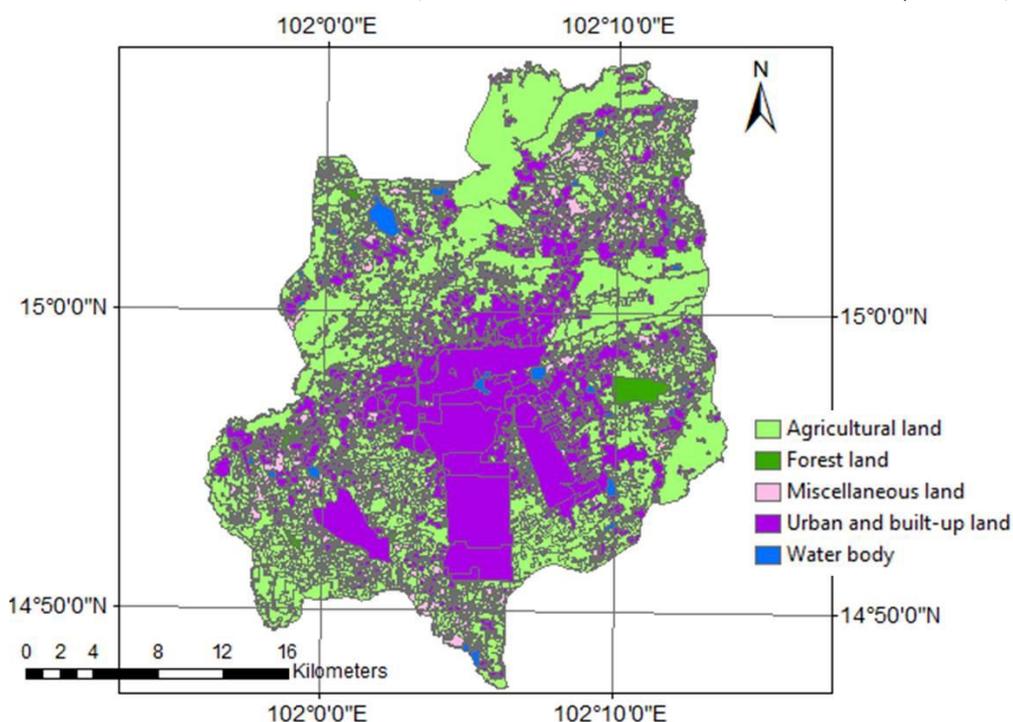


Figure 1 Study area and land use

Table 2 Land use year 2019 for MN district

No.	LU types	Area (sq.km.)	Percent
1.	Urban and built-up land (U)	257.21	33.60
2.	Agricultural land (A)	420.93	54.99
3.	Forest land (F)	8.05	1.05

Table 2 Land use year 2019 for MN district (Continued)

No.	LU types	Area (sq.km.)	Percent
4.	Miscellaneous land (M)	63.50	8.30
5.	Water body (W)	15.80	2.06
Total		765.49	100.00

Materials and methods

This study included three steps (as Figure 2) and more details as follows:

Step 1: Land use data of MN district year 2019 which was in the form of shapefile and was downloaded from web of LDD (<http://dinonline.ldd.go.th/>). This shapefile of land use was checked by using Editor Toolbar in ArcMap. On the other word, polygons in land use layer were checked errors before it was input in mapping clusters tool based on ArcMap or using in step 2.

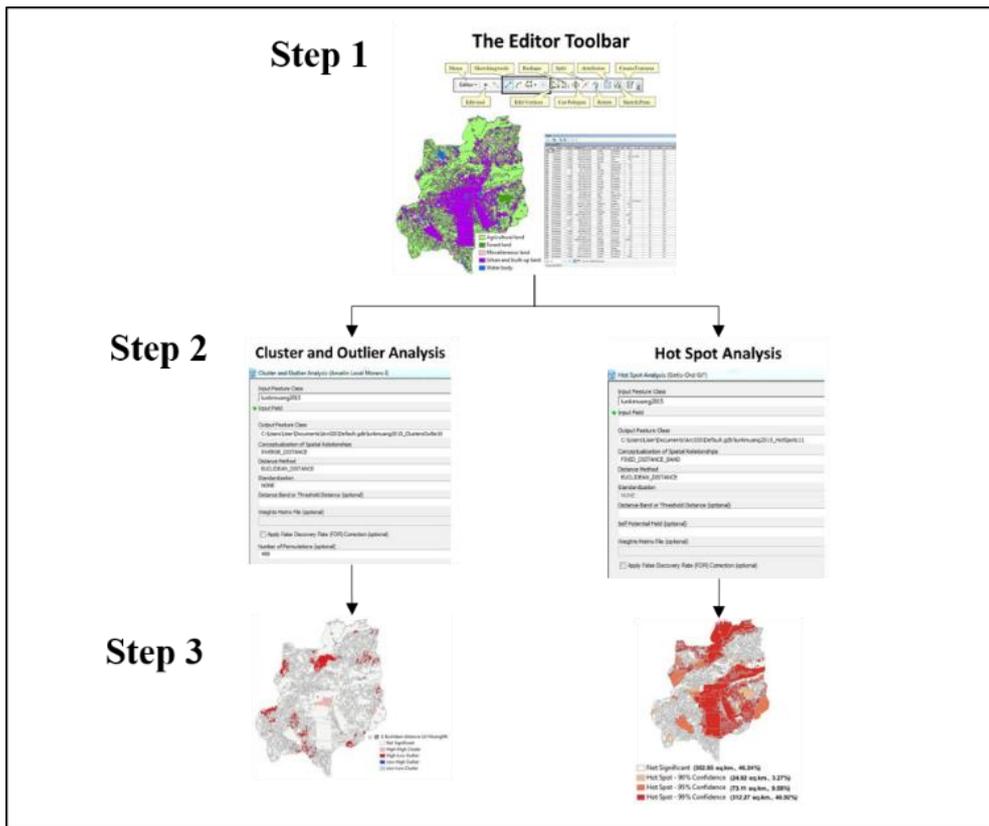


Figure 2 Steps for grouping of land use data based on mapping clusters' techniques in ArcMap

Step 2: This study focuses on land use clustering with two techniques of mapping clusters in ArcMap: (1) cluster and outlier analysis (Anselin Local Moran's I) and (2) hot spot analysis (Getis-Ord G_i^*) as follows:

1. Cluster and outlier analysis shows Local Moran's I index, z-score, p-value, and cluster/outlier type (COType) that indicate the apparent similarity (a spatial clustering of either high or low values) or dissimilarity (a spatial outlier) is more pronounced than one would expect in a random distribution (ESRI, 2018). These outputs of cluster and outlier analysis have benefits such as (1) the index of local Moran's I is a useful tool for identifying and classifying spatial clusters and outliers (Zhang et al., 2008); (2) the distributional characteristics of the Morn's I are more desirable and the indicator has greater general stability and flexibility (Zhang and Lin, 2016; Mitra, 2019); (3) Morn's I can qualitatively detect whether the spatial distribution is dispersed, random, or clustered in the entire space with respect to their attribute values (Zhang et al., 2019).

2. Hot spot analysis identifies statistically significant spatial clusters of high values (hot spots) and low values (cold spots) with a z-score, p-value, and confidence level bin (G_i _Bin) for each feature (ESRI, 2018). The benefits of hotspots analysis include the following factors (Barthel et al., 2015): the rapid assimilation and analysis of multiple evidence threads leading to accessible outputs and a clearer understanding of the actions required to eliminate, reduce or mitigate identified hotspots. For example, the hot spot analysis tool assesses whether high or low values (the number of crimes, accident severity, or dollars spent on sporting goods (ESRI, 2020).

In mapping clusters tool-based ArcMap, cluster and outlier analysis and hot spot analysis were set to be used six techniques of spatial relationships (inverse distance, inverse distance squared, fixed distance band, zone of indifference, contiguity edges only and contiguity edges corners) and two distance methods (Euclidean and Manhattan). Each technique of spatial relationship was analyzed together with such two distance methods. In addition, this study used guidelines of ESRI (ESRI, 2018a; ESRI, 2018b) for studying mapping clusters tool in ArcMap and using result interpretation and explanation in step 3.

Step 3: the results of spatial relationships and distance methods-based cluster and outlier analysis, and hot spot analysis are reported by interpreting their outputs feature class (such as shapefiles and attribute data) as follows:

1. Cluster and outlier analysis was reported from a new output feature class with the following attributes for each feature in the input feature class: Local Moran's I index, z-score, p-value, and cluster/outlier type (COType).

2. Hot spot analysis was reported with a z-score, p-value, and confidence level bin (Gi_Bin) for each feature in the input feature class.

Results and discussions

Area of MN district in year 2019 as reported by LDD-based Geographical Information System (GIS) data shows the highest percentage of agricultural land (54.99%), the second-highest of urban and built-up land (33.60%), and other areas (11.14%) e.g., forest land, water body etc. This GIS-land use data was analyzed by conceptualization of spatial relationships and distance methods-based cluster and outlier analysis and hot spot analysis in ArcMap. There are results as follows:

Land use clustering-based cluster and outlier analysis

Land use of NM district was grouped by cluster and outlier analysis that can present four clusters: High-High (HH as the light red), High-Low (HL as the dark red), Low-High (LH as the dark blue) and Low-Low (LL as the light blue) as shown in Figure 3 and Table 3. The statistically significant clusters correspond to the clusters and outliers with 95% confidence level and p-value smaller than 0.05. This study set cluster as HH and LL and outlier as HL and LH. We found that most groups of urban and built-up land that occupies 515.75 sq.km (67.58%) represented statistically significant HH cluster of high values using contiguity edges only and corners-based Euclidean distance. The distribution of HH clusters on city, town and commercial area were interesting, especially when overlaid on the land use designations. As shown in the study area (Figure 3) on the middle location (lightly red color), HH cluster was associated with a section of adjacent private lands. For outlier, it was mostly represented by statistically significant HL cluster (high values surrounded by low values) using fixed distance band and zone of indifference-based Euclidean distance. This HL cluster showed high value agricultural land occupied 177.23 sq.km (23.22%) which was surrounded by low value of village area. The distribution of the HL cluster on agricultural land and village area is interesting, when overlaid on the land use designations. As shown in the study area (Figure 3) on the northern location (red color), HL cluster was associated with a section of adjacent private lands as same as HH cluster. In addition, urban and built-up land was not significantly grouped

and covered, which was found to be varied from 419.71 sq.km (55.00%) to 704.16 sq.km (92.27%) in the study area.

When we required statistically significant clusters of high values or HH cluster, we should select technique of spatial relations such as contiguity edges only-based Euclidean distance and contiguity edges corners-based Euclidean distance.

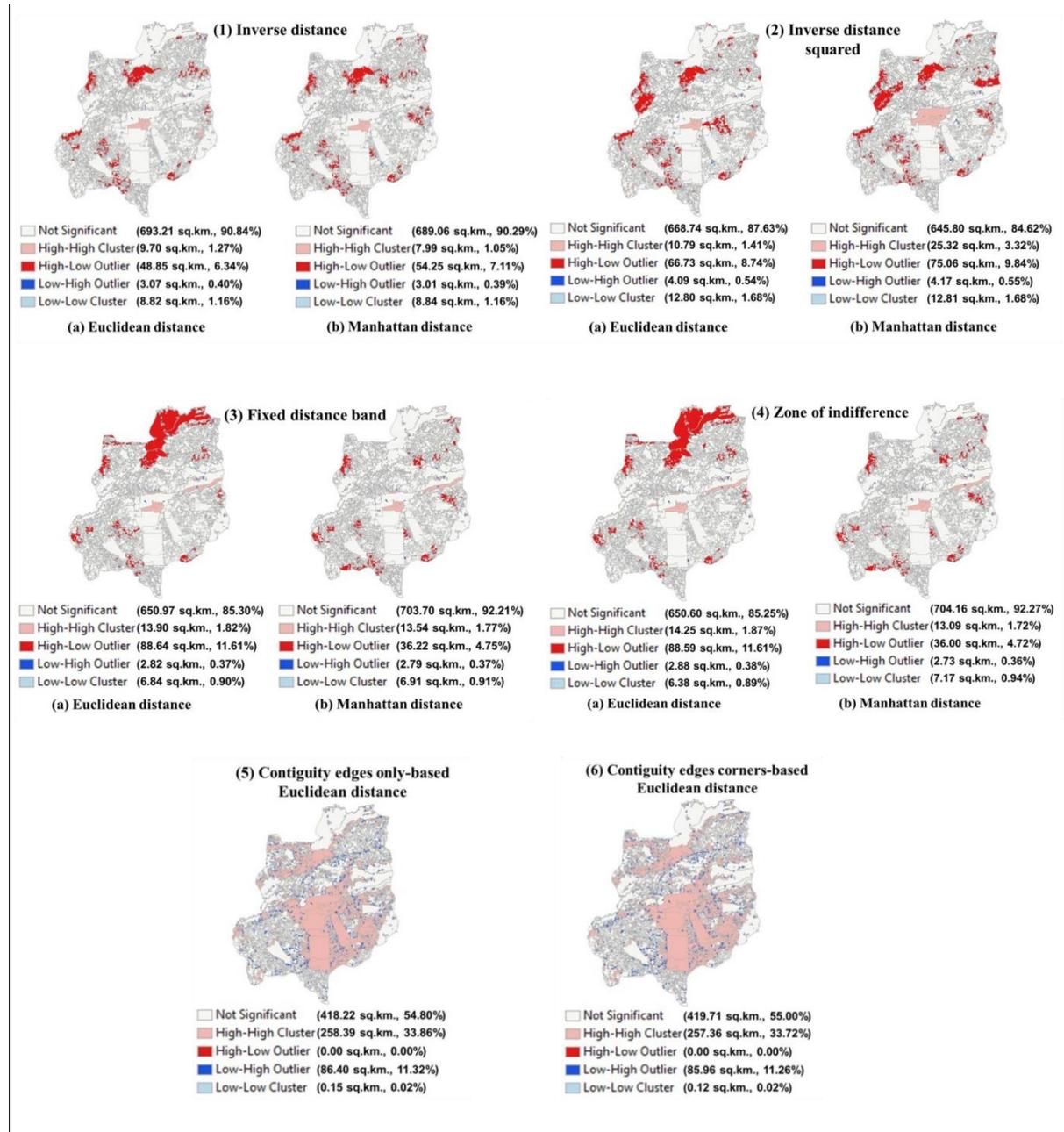


Figure 3 Results for analyzing spatial relationships and distance methods with cluster and outlier analysis

Table 3 Results for analyzing spatial Relationships and Distance Methods-based cluster and outlier analysis

Types of spatial relationships and distance methods	Cluster and outlier analysis, sq.km (%)														
	Euclidean distance						Manhattan distance								
	NS	HH	HL	LH	LL	NS	HH	HL	LH	LL	NS	HH	HL	LH	LL
1. Inverse distance	693.21 (90.84%)	9.70 (1.27%)	48.85 (6.34%)	3.07 (0.40%)	8.82 (1.16%)	689.06 (90.29%)	7.99 (1.05%)	54.25 (7.11%)	3.01 (0.39%)	8.84 (1.16%)					
2. Inverse distance square	668.74 (87.63%)	10.79 (1.41%)	66.73 (8.74%)	4.09 (0.54%)	12.80 (1.68%)	654.80 (84.62%)	25.32 (3.32%)	75.06 (9.84%)	4.17 (0.55%)	12.81 (1.68%)					
3. Fixed distance band	650.97 (85.30%)	13.90 (1.82%)	88.64 (11.61%)	2.82 (0.37%)	6.84 (0.90%)	703.70 (92.21%)	13.54 (1.71%)	36.22 (4.75%)	2.79 (0.37%)	6.91 (0.91%)					
4. Zone of indifference	650.60 (85.25%)	14.25 (1.87%)	88.59 (11.61%)	2.88 (0.38%)	6.38 (0.895)	704.16 (92.27%)	13.09 (1.72%)	36.00 (4.72%)	2.73 (0.36%)	7.17 (0.94%)					
5. Contiguity edges only-based Euclidean distance	418.22 (54.80%)	258.39 (33.86%)	0.00 (0.00%)	86.40 (11.32%)	0.15 (0.02%)	-	-	-	-	-					
6. Contiguity edges corners-based Euclidean distance	419.71 (55.00%)	257.36 (33.72%)	0.00 (0.00%)	85.96 (11.26%)	0.12 (0.02%)	-	-	-	-	-					

Remark: NS=Not Significant, HH=High-High Cluster, High-Low Outlier, LH=Low-High Outlier, LL=Low-Low cluster and (-) = Manhattan distance does not have in types of spatial relationships no.5 and 6.

Land use clustering-based hot spot analysis

Land use of the NM district was grouped by hot spot analysis that shows similar patterns of significant clustering by year 2019. The output was found in the highest (i.e. hot spot or HS) and the lowest (i.e. cold spot or CS) specific locations. These HS and CS clusters corresponded to p-value smaller than 0.10 (90% confidence as the lightest red), 0.05 (95% confidence as the middle red), and 0.01 (99% confidence as the darkest red), which are considered statistically significant (Figure 4). We found that land use data of NM district can be classed in the direction of HS cluster as CS cluster has not appeared. In HS classes, they were separated by three clusters: HS-90% confidence, HS-95% confidence, and HS-99% confidence as shown in Figure 4 and Table 4. The areas of land use 2019 in the NM district were mostly seen at HS-99% confidence that covered and varied from 122.48 sq.km. (16.65%) to 318.59 sq.km. (41.75%) using inverse distance-based Euclidean and Manhattan distance or inverse distance square-based Euclidean and Manhattan distance. In this land use grouping at 99% confidence, it was an expected increase in clustering of urban and built-up land and agricultural land on the study area. At the same time, the grouping of land use at HS 90% confidence was the expected decrease in clustering of water bodies and forest areas on the study area, covered, and varied from 0.77 sq.km (0.10%) to 24.92 sq.km (3.27%). In addition, the most agricultural land in the study area, which varied from 352.47 sq. km (46.19%) to 634.50 sq. km (83.14%), was not significantly grouped, and covered,

Therefore, if we required various land use types, we should select technique of spatial relations such as inverse distance, inverse distance squared, contiguity edges only-based Euclidean distance and contiguity edges corners-based Euclidean distance.

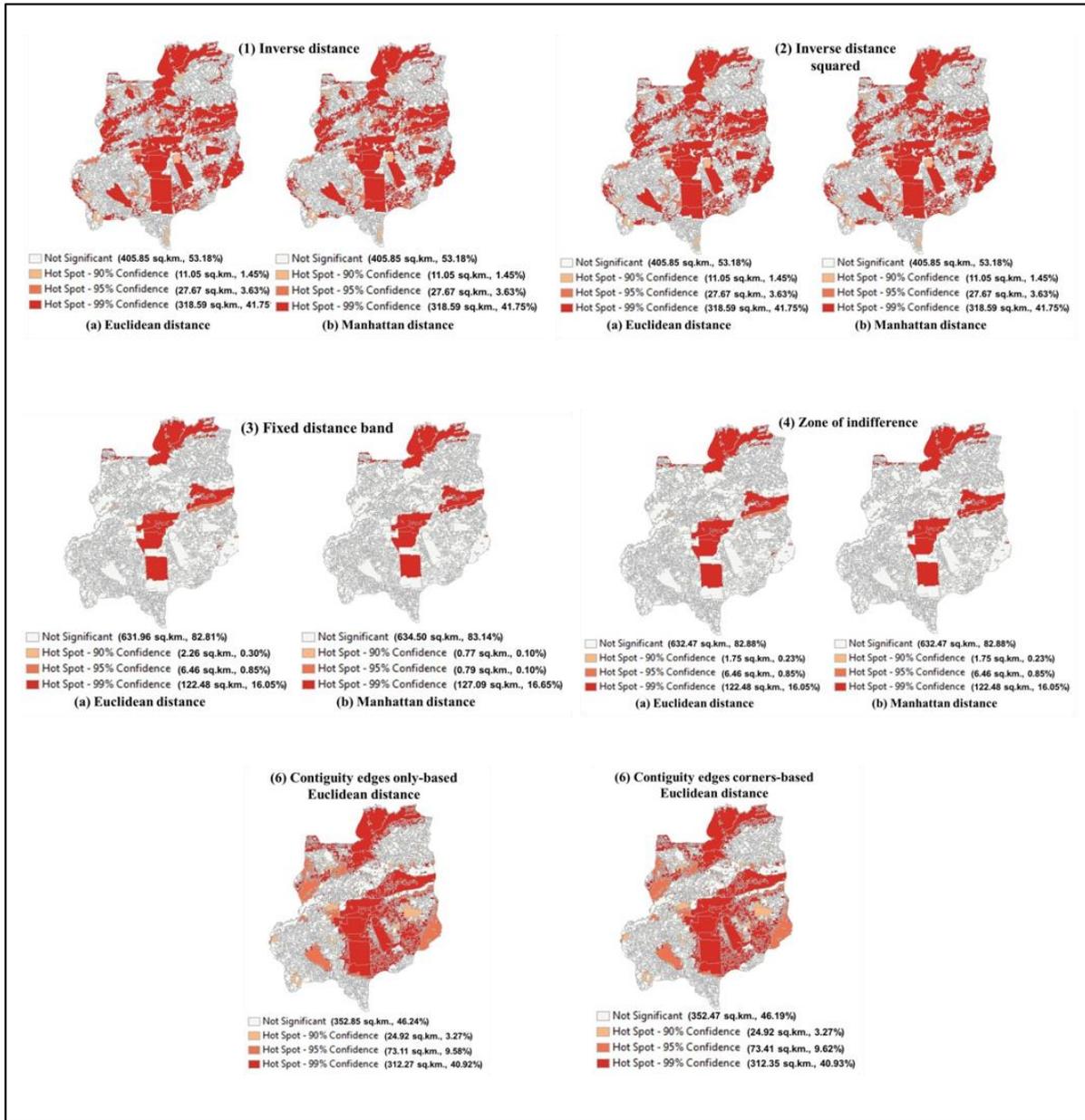


Figure 4 Results for analyzing spatial relationships and distance methods with hot spot analysis

Table 3 Results for analyzing spatial relationships and distance methods with hot spot analysis

Types of spatial relationships and distance methods	Hot spot analysis, sq.km. (%)									
	Euclidean distance					Manhattan distance				
	NS	HS-90%	HS-95%	HS-99%	NS	HS-90%	HS-95%	HS-99%		
1. Inverse distance	405.85 (53.18%)	11.05 (1.45%)	27.67 (3.63%)	318.59 (41.75%)	405.85 (53.18%)	11.05 (1.45%)	27.67 (3.63%)	318.59 (41.75%)		
2. Inverse distance square	405.85 (53.18%)	11.05 (1.45%)	27.67 (3.63%)	318.59 (41.75%)	405.85 (53.18%)	11.05 (1.45%)	27.67 (3.63%)	318.59 (41.75%)		
3. Fixed distance band	631.96 (82.81%)	2.26 (0.30%)	6.46 (0.85%)	122.48 (16.05%)	634.50 (83.14%)	0.77 (0.10%)	0.79 (0.10%)	127.09 (16.65%)		
4. Zone of indifference	632.47 (82.88%)	1.75 (0.23%)	6.46 (0.85%)	122.48 (16.05%)	632.47 (82.88%)	1.75 (0.23%)	6.46 (0.85%)	122.48 (16.05%)		
5. Contiguity edges only-based Euclidean distance	352.85 (46.24%)	24.92 (3.27%)	73.11 (9.58%)	312.27 (40.92%)	-	-	-	-		
6. Contiguity edges corners-based Euclidean distance	352.47 (46.19%)	24.92 (3.27%)	73.41 (9.62%)	312.35 (40.93%)	-	-	-	-		

Remark: NS=Not Significant, HS-90%=Hot Spot-90% Confidence, HS-95%=Hot Spot-95% Confidence, HS-99%=Hot Spot-99% Confidence and (-) = Manhattan distance does not have in types of spatial relationships no.5 and 6.

Conclusions

This paper proposed land use clustering of the MN district with a tool of mapping clusters in ArcMap. Land use data in 2019 from LDD was used for cluster and outlier and hot spot analysis. In cluster and outlier analysis, the output of cluster analysis indicates to study urban and built-up land, especially city, town, and commercial areas while outlier analysis indicates to study agriculture land and village area. This study suggests that, if we require studying urban area on land use in the year 2019, we should select the technique of contiguity edges only and corners with Euclidean distance because such two techniques have grouped as many areas as highly possible.

For hot spot analysis, the output of the expected increase (at 99% confidence) indicates clustering of urban and built-up land, and agricultural land while the expected decrease (at 90% confidence) indicates clustering of a water body and forest land. This study suggests that, if we require studying of urban and built-up land, and agricultural land at high confidence on land use in the year 2019, we should select the technique of inverse distance-based Euclidean and Manhattan distance or inverse distance square-based Euclidean and Manhattan distance as such technique have grouped as many areas as highly possible and at a significant confident level.

This land use clustering on land use data 2019 of MN district provides useful information on the spatial distribution of land use characteristics and patterns (homogenous and heterogeneous data) for planning land use or using analysis of spatio-temporal change. This study presents more rational understanding of choosing various techniques of spatial relation and distance methods. For example, the technique of contiguity edges only and corners with Euclidean distance is mostly suitable for HH clusters because they have the related concept of algorithm. Polygons are analyzed by contiguity edges only, and corners that share an edge and corner (that have coincident boundaries). For Euclidean distance, the shortest distance to a sourced polygon is determined, and if it is less than the specified maximum distance, the value is assigned or classed to the same group or location on the output vector.

Acknowledgement

The researchers would like to thank the Department of Land Development (LDD) for providing the land use information online and thank you to the Faculty of Sciences and liberal arts, Rajamangala University of Technology Isan (RMUTI) for supporting the writing and presentation of this article.

References

- Barthel, M., Fava, J. A., Harnanan, C. A. Strothmann, P., Khan, S. & Miller, S. (2015). **Hotspots analysis: providing the focus for action**. In G. Sonnemann & M. Margni (Eds.), *Life Cycle Management* (149-167). Retrieved from <https://link.springer.com/content/pdf/10.1007%2F978-94-017-7221-1.pdf>
- Carlier, J., Doyle, M., Finn, J.A., Uallachain, D.O. & Moran, J. (2021). A landscape classification map of Ireland and its potential use in national land use monitoring. **Journal of Environmental Management**, *289*, 112498.
- ESRI. (2018a). **Cluster and outlier analysis (Anselin local moran's i)**. Retrieved from <https://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-statistics-toolbox/cluster-and-outlier-analysis-anselin-local-moran-s.htm>
- ESRI, (2018b). **Hot spot analysis (Getis-Ord Gi*)**. Retrieved from <https://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-statistics-toolbox/hot-spot-analysis.htm>
- ESRI. (2020). **How hot spot analysis (Getis-Ord Gi*) works**. Retrieved from <https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/h-how-hot-spot-analysis-getis-ord-gi-spatial-stati.htm>
- Food and Agriculture Organization of the United Nations. (1992). **The future of our land**. Retrieved from <http://www.fao.org/3/x3810e/x3810e00.htm#TopOfPage>
- Jacquez, G. M. (2008). **Spatial clustering analysis**. Retrieved from https://www.biomedware.com/files/jacquez_ch22_preprint.pdf
- Kowe, P., Mutanga, O., Odindi, J. & Dube, T. (2020). A quantitative framework for analysing long term spatial clustering and vegetation fragmentation in an urban landscape using multi-temporal landsat data. **International Journal of Applied Earth Observation and Geoinformation**, *88*.

- Land Development Department. (2019). **Land use in Thailand 2017-2018**. (In Thai). Retrieved from http://www1.ldd.go.th/web_OLP/result/luse_result60-61.htm
- Mitra, S. (2009). **Spatial autocorrelation and bayesian spatial statistical method for analyzing intersections prone to injury crashes**. *Transportation Research Record*, 2136, 92-100.
- Obi Reddy, G. P. (2018). **Geospatial technologies in land resources mapping, monitoring, and management**. Retrieved from https://doi.org/10.1007/978-3-319-78711-4_1
- Office of the National Economic and Social Development Council. (2020). **20 year national strategy (2018-2037)**. (In Thai). Retrieved from [https://www.moac.go.th/pyp-dwl-files-402791791893#:~:text=The%20National%20Strategy%20\(2018%2D2037,developed%20pursuant%20to%20the%20Constitution.&text=Therefore%2C%20Thailand%20needs%20a%20comprehensive,towards%20security%2C%20prosperity%20and%20sustainability.](https://www.moac.go.th/pyp-dwl-files-402791791893#:~:text=The%20National%20Strategy%20(2018%2D2037,developed%20pursuant%20to%20the%20Constitution.&text=Therefore%2C%20Thailand%20needs%20a%20comprehensive,towards%20security%2C%20prosperity%20and%20sustainability.)
- Oregon State University. (2016). **Identifying clusters of clear-cuts and partial harvests using moran's i and hot spot analysis**. Retrieved from <https://blogs.oregonstate.edu/geo599spatialstatistics/2016/06/08/identifying-clusters-clear-cuts-partial-harvests-using-morans-hotspot-analysis/>
- Sanchez-Martin, J.-M., Rengifo-Gallego, J.-I. & Blas-Morato, R. (2019). Hot spot analysis versus cluster and outlier analysis: an enquiry into the grouping of rural accommodation in extremadura (Spain). *International Journal of Geo-Information*, 8(4), 176.
- Smith, J. (2011). **Creating land-use scenarios by cluster analysis for regional land-use and transportation sketch planning**. Retrieved from https://www.bts.gov/archive/publications/journal_of_transportation_and_statistics/volume_04_number_01/paper_03/index
- Susilo, B. (2016). **Map analysis and spatial statistic: assessment of spatial variability of agriculture land conversion at urban fringe area of Yogyakarta**. Retrieved from <https://iopscience.iop.org/article/10.1088/1755-1315/47/1/012001/pdf>
- Verburg, P. H., Crossman, N., Ellis, E. C., Heinimann A., Hostert, P., Mertz, O., Nagendra, H., Sikor, T., Erb, K.-H., Golubiewski, N., Grau, R., Grove, M., Konate, S., Meyfroidt, P., Parker, D. C., Chowdhury, R. R., Shibata, H., Thomson, A. & L. Zhen. (2015). Land system science and sustainable development of the earth system: A global land project perspective. *Anthropocene*, 12, 29-41.

- Verburg,P. H., Erb, K.-H., Mertz, O. & Espindola,G. (2013). Land system science: between global challenges and local realities. **Current Opinion in Environmental Sustainability**, **5**, 433-437.
- Victoria Transport Policy Institute. (2019). **Land use density and clustering**. Retrieved from <https://www.vtpi.org/tadm/tadm81.htm>
- WHO. (2016). **Urban green spaces and health – a review of evidence**. Retrieved from <https://www.euro.who.int/en/health-topics/environment-and-health/urban-health/publications/2016/urban-green-spaces-and-health-a-review-of-evidence-2016>
- WHO. (2017). **Urban green space interventions and health: A review of impacts and effectiveness**. Retrieved from <https://www.euro.who.int/en/health-topics/environment-and-health/urban-health/publications/2017/urban-green-space-interventions-and-health-a-review-of-impacts-and-effectiveness.-full-report-2017>
- Zhang, C., Luo, L., Xu, W. & Ledwith, V. (2008). Use of local moran's I and GIS to identify pollution hotspots of Pb in urban soils of Galway, Ireland. **Science of The Total Environment**, **398**(1-3), 212-221.
- Zhang, Z., Ming, Y. & Song, G. (2019). Identify road clusters with high-frequency crashes using spatial data mining approach. **Applied Sciences**, **9**, 5282.
- Zhang,T. & Lin, G. (2016). On Moran’s I coefficient under heterogeneity. **Computer Statistics & Data Analysis**, **95**, 83-94.