International Journal of GEOMATE, April, 2020, Vol.18, Issue 68, pp. 135 - 141
ISSN: 2186-2982 (P), 2186-2990 (O), Japan, DOI: https://doi.org/10.21660/2020.68.9220
Special Issue on Science, Engineering and Environment

REGIONALIZATION OF RAINFALL IN NORTHEASTERN THAILAND

Pongpinid Pinidluek[1], *Supasit Konyai[2] and Vichai Sriboonlue[3]

[1]Faculty of Engineering, Khon Kaen University, Thailand; [2,3] Khon Kaen University, Thailand

*Corresponding Author, Received: 15 June 2019, Revised: 25 Oct. 2019, Accepted: 09 Feb. 2020

ABSTRACT: Rainfall is one of the most valuable natural resources for northeastern Thailand, a region whose livelihood of more than 80% of the population depend predominantly on rainfed agriculture. Rainfall events can also cause severe natural hazard, therefore proper water resources and flood management is crucial to this region. Regionalization of rainfall can facilitate the management of water resources and floods by delineating a region with rainfall patterns that vary in both time and space into areas with more homogeneous rainfall characteristics. This study applied both hierarchical and K-means clustering methods to separate the entire Northeast into three precipitation zones using monthly rainfall data from 38 stations over 30 years from 1987 to 2016. The temporal variation of zoning was also investigated by dividing the data into three 10-year periods, 1987-1996, 1997-2006, and 2007-2016. The regionalization by both methods gave similar results that are comparable to the mean annual rainfall of the entire region. Mean annual rainfall values were highest in the northeastern part of the region with a gradual decreasing trend to the southwest. The clustering methods, on the other hand resulted in the highest rainfall zone in the northern part, the moderate rainfall on the southeastern, and the lowest on the southwestern part. As it relates to temporal variation, the results from 1987-1996 for both methods are mostly the same but those from 1997-2006 and 2007-2016 are not quite the same especially results from the last decade. Homogeneous rainfall regionalization by clustering methods is preferable to merely geographical and mean annual rainfall values because they take into account several factors such as monthly rainfall distribution.

Keywords: Hierarchical clustering, K-means clustering, Homogeneous rainfall, Temporal variation, Northeastern Thailand.

1. INTRODUCTION

Rainfall is the primary cause of flood, drought, and other hydrological consequences. Many hydrological and water resources analyses, designs, and operations are needed to know the rainfall homogeneous regions [1]. Traditionally, homogeneous rainfall region has been determined from the mean annual precipitation and geography of the region [2]. Since the northeastern Thailand climatic condition is monsoonal the area is wet during the southwest monsoon from mid-May to mid-October and dry from mid-October to mid-May by the northeast monsoon [3], [4]. This is actually the result of the dynamics of Intertropical Convergence Zone (ITCZ) [5], [6] that moves across the northeast region during the wetting period by touching the southern boundary of the region at about mid-May and keep on moving up passing the northern boundary at about early July and moving further north then returning to touch the northern boundary again at about mid-July then slowly moving southward across the Northeast region again until mid-October then passing the southern boundary of the region. Mobility of the ITCZ causes rainfall pattern of the Northeast region to be ever changing [7].

The Northeast topography is a plateau-like with mountain ranges on the west and the south boundaries. The rainy season starts when the ITCZ band first approaching the southern boundary about mid-May and ends when the ITCZ about to leave the same boundary at mid-October. During the ITCZ band moving further north, the rainfall of the Northeast decreases allowing dry spell to interfere with its rainfall pattern. Two types of wind bring moist air from the oceans to the Northeast region namely the southwest monsoon from the Indian Ocean and the tropical cyclone from the Pacific Ocean. These phenomena make the rainfall pattern very complicate. Regionalization of rainfall in this part of Thailand cannot be performed just by subjectively observing the mean annual rainfall and its geography but must be applying cluster analyses. Two well-known clustering algorithms being generally used for this type of work are hierarchical and k-means methods [1], [8].

Our study objective is to partition the whole Northeast region into three uniform precipitation sub-regions using the two cluster analyses with monthly rainfall data on scatter rainfall stations all over the region during three 10-year periods.

2. STUDY AREA AND DATA ACQUISITION

Our study area is the whole northeast region of Thailand. It covers the area of 168,854 km^2 of a relatively mild rolling plain with the elevation of about 150 to 500 m above mean sea level. Due to its plateau-like topography it is called the Khorat Plateau [9]. There are mountain ranges on the west and south borders, the Phetchbun and Dong Paya Yen on the west boundary which border with the Central and the Sankambeng and Phanom Dangreg Ranges on the south which border with the Central and Cambodia (Fig. 1). These mountain ranges create orographic effects as rain shadows retarding the flow of southwest monsoon to the region. The Phu Phan Range divides the Khorat Plateau into two basins the upper is smaller called the Sakol Nakhon Basin and the lower part is called the Khorat Basin. Both basins drain to the Mekong River in the north and east boundaries by the Songkhram River for the upper basin and the Chi-Mun System for the Khorat Basin. Geology of the northeast composts of sandstone of the Mesozoic Khorat Group [10]. Following the parent materials of sandstone, typical soils of the northeast are sandy soils which are very infertile and sometime saline [9]. Hydrologically, there are only wet and dry seasons in the northeast, Thailand, each sharing half of the year. Two sources of humidity to produce rainfall in the northeast region namely the southwest monsoon and the tropical cyclone. The topographic and climatic conditions complicate the rainfall pattern of the northeast region, therefore it is essential to partition the whole northeast region into more compact, homogeneous areas. We used monthly rainfall data from 38 rainfall stations for 30 years during 1987 to 2016 from the Thai Meteorology Department for our study. The names, locations, and elevations of the rainfall stations are shown in Table 1 and Fig. 1.

3. METHODOLOGY

Traditional method of rainfall regionalization is by using mean annual depth of precipitation e.g. low, modulate, and high rainfall [11]. Objectively, rainfall regionalization can be performed by mean of a cluster analysis [12], which is a set of algorithms to group the similar objects from a diverse observations. In this study we tried to partition 38 rainfall stations of the whole Northeast region (Fig. 1 and Table 1) into 3 groups of homogenous rainfall characteristics as in [13]. From 30-year data we investigated each of three 10-year periods, i.e. 1987-1996, 1997-2006, and 2007-2016, to evaluate temporal variation.

Two types of clustering algorithms, i.e. hierarchical and k-means, were used with monthly rainfall depths which were arranged in the forms of 3 tables. Each table contains monthly rainfall data with 120 columns of months of the years, say from January 1987 to December 1996 for the first table, and 38 rows of the stations in Table 1. The hierarchical algorithm and the k-means one are totally different [8]. The former produces a dendrogram of the stations with similar rainfall characteristics [14].

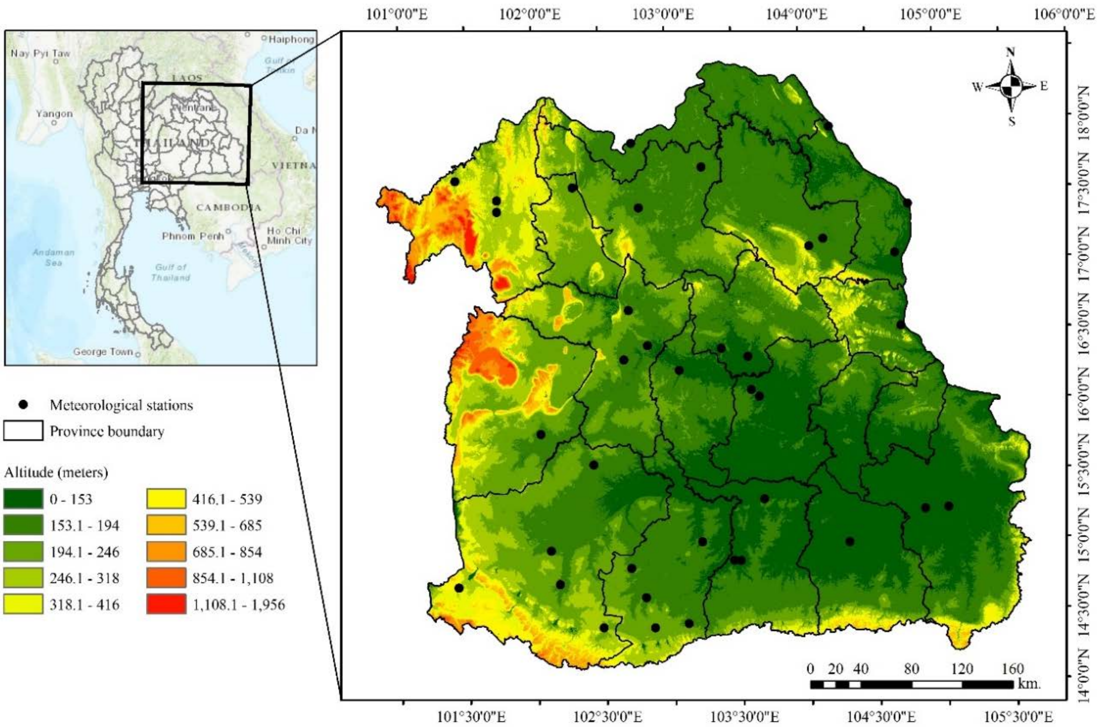


Fig.1 Topography of the northeast and locations of rain gauges.

Table 1 Details of rainfall stations

Code	Name	Elevation (m)	Latitude	Longitude
352201	Nong Khai	168	17°52'44.63"	102°44'14.11"
353004	Tha Li	266	17°38'25.64"	101°25'50.36"
353201	Loei	244	17°28'54.86"	101°42'55.42"
353301	Loei Agrometeorological Station	286	17°23'44.08"	101°44'16.82"
354005	Ban Dung	176	17°41'1.34"	103°16'0.62"
354201	Udon Thani	178	102°47'22.12"	102°47'22.12"
356201	Sakon Nakhon	164	17°10'53.67"	104°9'58.38"
356301	Sakon Nakhon Agrometeorological Station	193	17°6'58.07"	104°2'22.05"
357003	Renu Nakhon	156	17°2'31.12"	104°40'46.36"
357005	Ban Phaeng	153	17°58'52.35"	104°12'46.61"
357301	Nakhon Phanom Agrometeorological Station	150	17°24'13.61"	104°47'12.42"
360004	Suwan Khuha	206	17°34'34.27"	102°18'9.30"
381017	Ubonrat Dam Self-Supporting Settlement	187	16°41'54.34"	102°41'7.94"
381201	Khon Kaen	164	16°25'56.53"	102°49'22.69"
381301	Tha Phra Agrometeorological Station	179	16°19'57.09"	102°38'13.41"
383201	Mukdahan	152	16°31'56.92"	104°42'21.08"
387401	Maha Sarakham	160	16°15'21.23"	103°4'7.05"
388008	Kalasin Seed-Multiplication Station	152	16°24'16.31"	103°22'6.17"
388401	Kalasin	146	16°19'53.19"	103°34'27.44"
403201	Chaiyaphum	185	15°48'31.42"	102°2'28.95"
405201	Roi Et	151	16°2'57.83"	103°38'41.08"
405301	Roi Et Agrometeorological Station	163	16°5'45.84"	103°36'15.77"
407301	Ubon Ratchathani Agrometeorogical Station	132	15°13'55.05"	105°2'3.04"
407501	Ubon Ratchathani	123	15°13'8.59"	104°52'29.69"
409301	Si Sa Ket	132	15°0'36.30"	104°18'37.16"
431002	Bua Yai	173	15°34'25.59"	102°25'20.42"
431013	Soeng Sang	228	14°24'33.08"	102°29'10.29"
431201	Nakhon Ratchasima	191	14°58'56.35"	102°5'39.37"
431301	Pak Chong Agrometeorological Station	313	14°42'13.17"	101°25'6.81"
431401	Chok Chai	196	14°43'24.48"	102°9'19.40"
432201	Surin	149	14°52'23.58"	103°29'20.55"
432301	Surin Agrometeorological Station	142	14°52'26.48"	103°26'52.24"
432401	Tha Tum	134	15°18'36.57"	103°40'51.98"
436002	Lahan Sai	215	14°25'14.20"	102°51'24.6"
436009	Ban Kruat	190	14°26'42.64"	103°5'53.88"
436012	Nong Hong	185	14°50'10.10"	102°40'32.14"
436201	Burirum	157	15°1'17.07"	103°12'52.11"
436401	Nang Rong	190	14°37'40.43"	102°47'44.51"

At the beginning procedure the two most similar rainfall characteristics being merged together and then the third station (or another cluster of stations) to be merged with the Ward's criteria that is the sum of euclidean distances among the stations in the new cluster must be minimum [8]. Each dendrogram (not shown herein) illustrates 38 clusters at its base and 1 cluster at its top.

The k-means algorithm needs to be provided with k seeds randomly [12] among each of the three data sets. In our case, we set up 3 seeds since we need to group into three clusters as probably of low, medium, and high rainfall. From the 3 predetermined seeds, the rainfall data that are more similar to a seed are grouped together with the seed, therefore we have three groups for the whole data. Consequently, new seed is to be determined for each group e.g. at the centroid of the group, then the process to be performed all over again until the stations in each group do not change [8].

The similarity are always measured by euclidean distance which is not really a distance but it is a square root of sum square of the coordinates difference between two points. The points are the rainfall stations and their coordinates are monthly rainfall data from each ten year period. Since rainfall depths in the wet seasons are much larger than those in the dry seasons, therefore the data in each station must be normalized so that the wet season data do not dominate the dry season ones.

Tables of monthly rainfall data from 38 stations were prepared using Excel software then transformed into text files to be used with the R programming. The text file tables of rainfall data were input as data frames. The data frames were normalized using *scale* function. The normalized table can be clustered by hierarchical method using *hclust* function together with the Ward's criteria, e.g.

```
ne.rainh1  <-  hclust(table1, method = 'ward')
```

This is an R code to cluster by hierarchical algorithm the data fame of 10-year monthly rainfall data of 38 rainfall stations called table1 with Ward's method for merging criteria. The result of hierarchical clustering on the data frame is kept in *ne.rainh1*.

For k-means algorithm we use *kmeans* function to group the similar rainfall stations from the same data frame *table1* into 3 clusters in R code as:

```
ne.raink1  <-  kmeans(table1, 3)
```

4. RESULTS AND DISCUSSIONS

The three 10-year rainfall data (1987-1996, 1997-2006, and 2007-2016) were grouped into three clusters using the hierarchical and K-means algorithms. The results are shown in Fig. 2 as triangles, squares, and hexagons for each group, the results from hierarchical on the left and those from K-means on the right, the three pairs represent each period from the first to last. The isohyets are also shown on Fig. 2 illustrating that the highest rainfall is in the northeastern part of the region and the lowest is in the southwestern part for all three periods. The annual means rainfall of the region were 1306, 1391, and 1438 mm, for the first to the last periods, respectively. The increases of annual means from the first to the second period is 6.6 % and from the second to the third is 3.4 %. These are likely to be the effect of climate change [15].

The top pair of Fig. 2, the first period of 1987-1996, demonstrate hexagons to occupy the northern part of the region, the squares the southeastern, and the triangles the southwestern part. The positions of the three symbols from both algorithms, left and right in Fig. 2, are exactly the same. We presumed that the hexagons represent the high rainfall region, the squares the moderate, and the triangles are the low rainfall.

The positions of the three symbols from both methods for the second period are not exactly the same as in the first period. For this 2nd period, the hexagons still occupy the northern region, but the squares cover most of the middle zone including the southeast portion. The K-means method give preferable pattern for this period.

For the last period, the hierarchical algorithm gives preferable pattern of the 3 symbols to the K-means one. The hexagons cover the north and northwest zones, the squares on the eastern, while the triangles occupy the southwestern region. We cannot conclude that the hexagons are indicated the high rainfall zone and the triangle the low rainfall one. For the K-means method of the 3rd period showing 4 hexagons in the zone of low rainfall (Fig. 2).

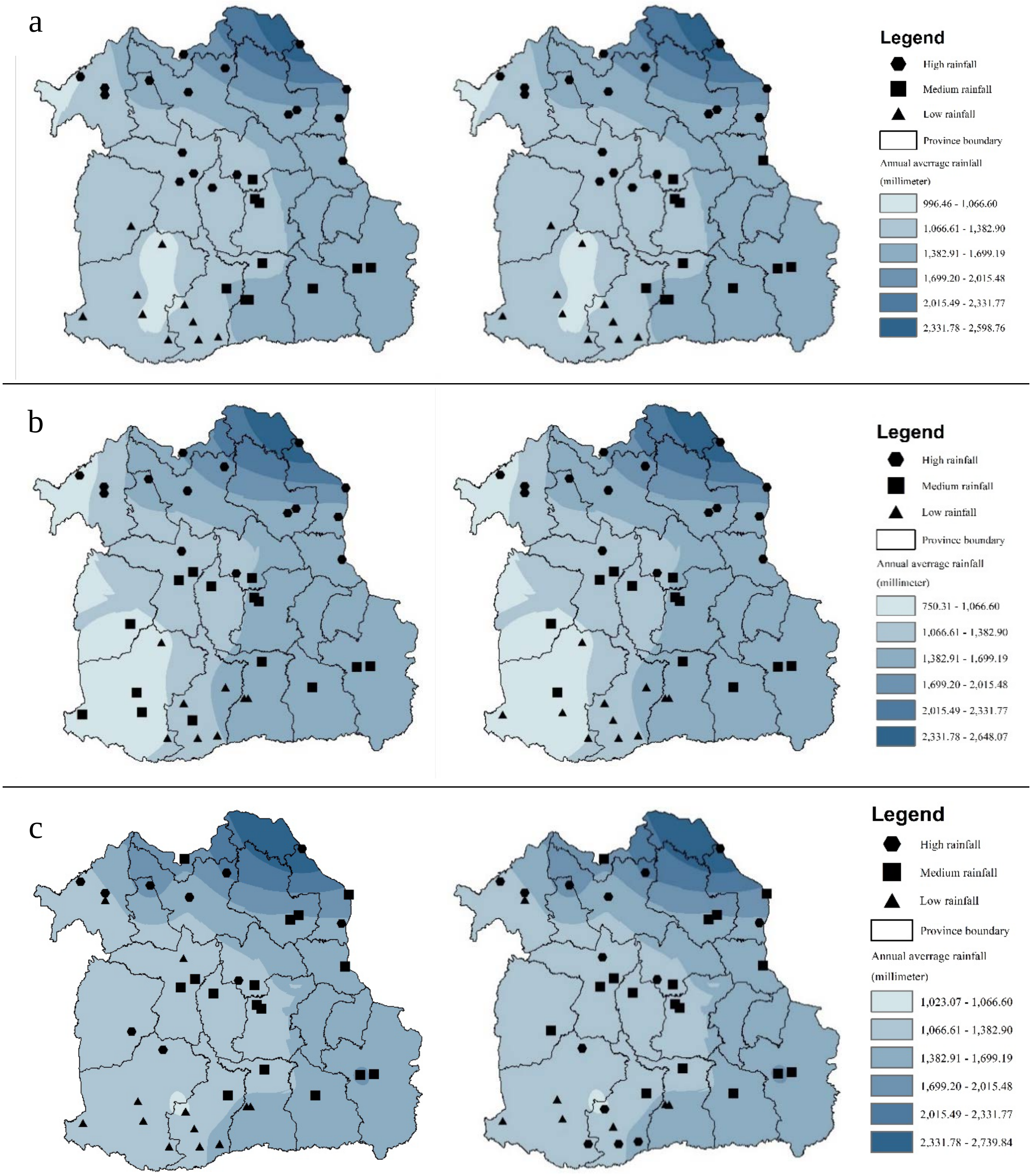
Fig. 3 illustrates rainfall distributions throughout the year. The triangles are of the top pair, the squares of the middle, and the hexagons of the bottom one. The hierarchical clustering is on the left and the K-means on the right. Fig. 3 and Table 2 show that peaks of hexagon group of rainfall always being in the month of August but for those of square and triangle groups being in both August and September. The dry spell during rainy season were in the month of July for all groups of the 1st period. For the 2nd and 3rd periods, there was no dry spell during the rainy season except for the triangles were in June. These phenomena were the consequence of the movement of ITCZ rain band. The ITCZ band moved much farther north creating dry spell in July for all groups (triangle, square, and hexagon). Only the triangle group did show the dry spell in June, there was no dry spell for other groups.

The above phenomena demonstrate that the hexagons always peak before the others since the ITCZ band, on the way back, reaching the hexagons before the others. The absent of dry spell with a few in June, for the 2nd and 3rd periods, show the ITCZ band did not move too far off from the north border of the Northeast.

5. CONCLUSIONS

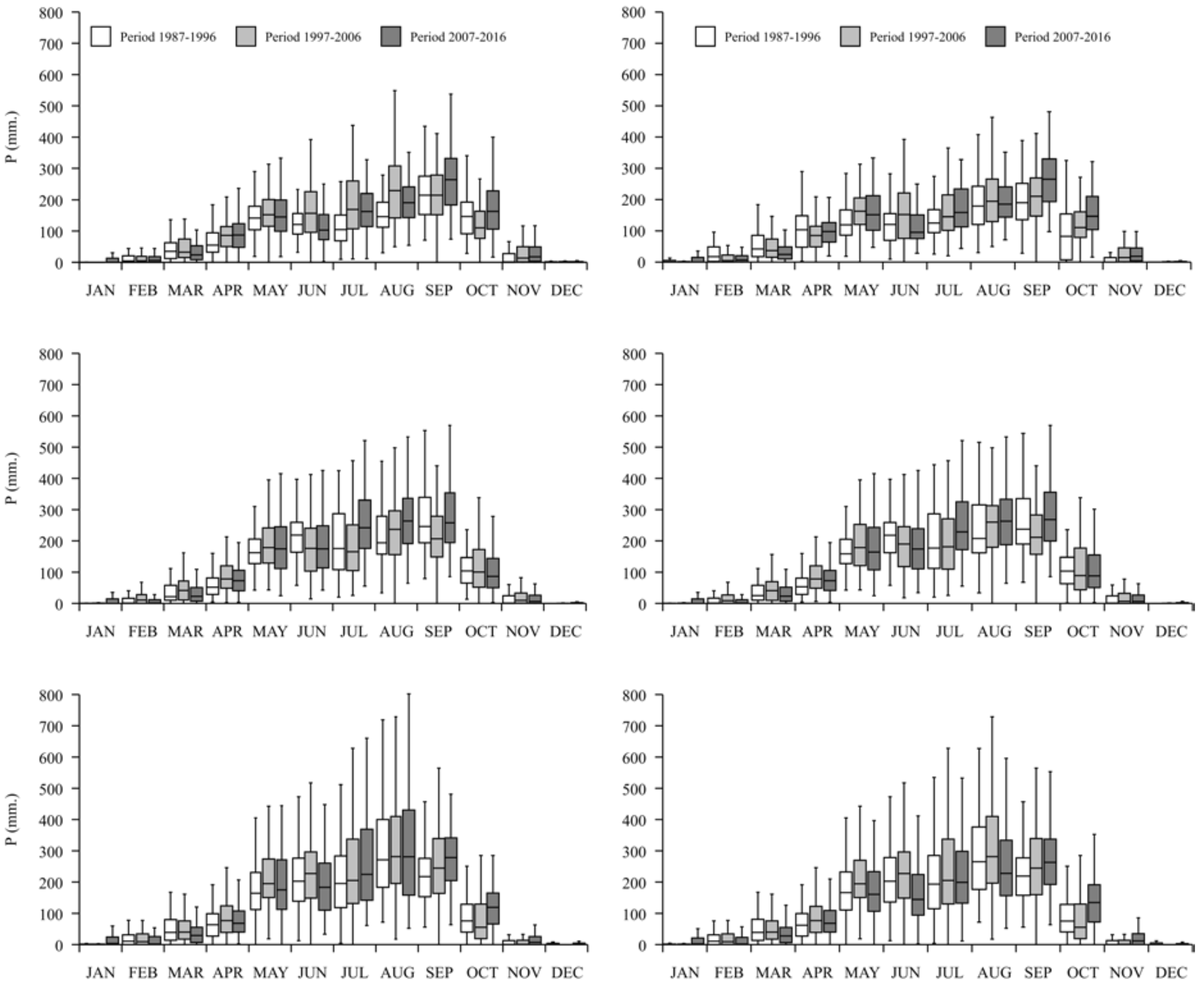
The regionalization of homogeneous rainfall pattern of the Northeast, Thailand, was done by hierarchical and K-means methods with data from

38 rain gauges for three periods, 1987-1996, 1997-2006, and 2007-2016. The rainfall characteristics were partitioned into three groups illustrating by triangle, square, and hexagon. These groups did not exactly match with the amounts of the mean annual rainfall. However, the triangle group got along with the low rainfall zone, the square with the moderate, while the hexagon agreed with the high rainfall region. The hexagon group occupied the northern part of the Northeast, the square the southeastern, whereas the triangle covered the southwestern part. The movement of the ITCZ band went farther from the northern boundary of the region creating dry spell during rainy season in July for the 1st period (1987 to 1996). For the 2nd and 3rd periods, rainy season came early and the ITCZ band did not move further away from the northern border, by majority no dry spell existed. The mean annual rainfall of the region increased from the 1st period to the 2nd then to the 3rd at 1306, 1391, and 1438 mm, respectively. This is a sign of climate change**.**



Hierarchical clustering method K-mean clustering method

Fig.2 Identification of clusters for hierarchical and K-mean clustering method: (a) period 1987-1996; (b) period 1997-2006 and (c) period 2007-2016.



Hierarchical clustering method K-mean clustering method

Fig.3 Monthly rainfall distributions in three 10-year periods.

Table 2 Monthly rainfall cluster by using Hierarchical and K-means methods. (H = Hierarchical, K = K-means, N = None)

Periods	Zones	Method	Peak	Dry spell	Mean annual rainfall
1987-1996	low	H	Sep.	Jul.	1,078
		K	Sep.	Jul.	1,078
	medium	H	Sep.	Jul.	1,331
		K	Sep.	Jul.	1,346
	high	H	Aug.	Jul.	1,417
		K	Aug.	Jul.	1,413
1997-2006	low	H	Aug.	N	1,294
		K	Aug.	Jun.	1,245
	medium	H	Aug.	N	1,309
		K	Aug.	N	1,357
	high	H	Aug.	N	1,555
		K	Aug.	N	1,555
2007-2016	low	H	Sep.	Jun.	1,258
		K	Sep.	Jun.	1,274
	medium	H	Aug.	N	1,475
		K	Sep.	N	1,463
	high	H	Aug.	N	1,594
		K	Aug.	N	1,488

6. ACKNOWLEDGMENTS

The authors would like to thank the Graduate School and Applied Engineering for Important Crops of North East Research Group, Khon Kaen University and the Water Engineering Group, Faculty of Engineering, Khon Kaen University for funding and data supporting.

7. REFERENCES

[1] Zhang Y., Moges S., and Block P., Optimal cluster analysis for objective regionalization of seasonal precipitation in regions of high spatial-temporal variability: Application to western Ethiopia. Journal of Climate, Vol. 29, 2016, pp. 3697-3717.

[2] Ahuja S., and Dhanya C.T., Regionalization of rainfall using RCDA cluster ensemble algorithm in India. Journal of Software Engineering and Applications, Vol. 5, 2012, pp. 568-573.

[3] Sangwaldach P., Wongwises P., Exell R.H.B., and Sukawat D., The surface features of southwest monsoon onset over Thailand. The 2nd Joint International Conference on "Sustainable Energy and Environment (SEE 2006)" D-015 (O).

[4] Naeimi V., Leinenkugel P., Sabel D., Wagner,W., Apel H., and Kuenzer C., Evaluation of soil moisture retrieval from the ERS and Metop scatterometers in the Lower Mekong Basin. Remote Sensing, Vol. 5, 2013, pp. 1603-1623.

[5] Byrne M.P., Pendergrass A.G., Rapp A.D., and Wodzicky K.R., Response of the intertropical convergence zone to climate change: Location, width, and strength. Current Climate Change Reports, Vol. 4, 2018, pp. 355-370.

[6] Chokngamwong R. and Chiu L.S., Thailand daily rainfall and comparison with TRMM products. Journal of Hydrometeorology, Vol. 9, 2008, pp. 256-266.

[7] Weesakul U., and Lowanichchai S., Rainfall forecast for agricultural water allocation planning in Thailand. Thammasart International Journal of Science and Technology, Vol. 10, Issue 3, 2005, pp. 18-27.

[8] Rogerson P.A. (2010) Statistical Methods for Geography: A Student's Guide. Third edition, Sage.

[9] Loffler E., Thompson W.P., and Liengsakul M., Geomorphological development of the Thung Kula Ronghai, Proceedings of the 1st Symposium on Geomorphology and Quaternary Geology of Thailand, 1983, pp.123-130.

[10] Wongsomsak S., Salinization in Northeast Thailand, Southeast Asian Studies, Vol. 24, Issue 2, 1986, pp. 133-153.

[11] Limsakul A., and Singhruck P., Long-term trends and variability of total and extreme precipitation in Thailand. Atmospheric Research, Vol. 169, 2016, pp. 301-317.

[12] Carvalho M.J., Melo-Goncalves P., Teixeira J.C., and Rocha A., Regionalization of Europe based on a k-means cluster analysis of the climate change of temperatures and precipitation. Physics and Chemistry of the Earth, Vol. 94, 2016, pp. 22-28.

[13] Venkatesh B., and Jose M.K., Identification of homogeneous rainfall regimes in parts of Western Ghats region of Karnakata. Journal of Earth System Science, Vol. 16, Issue 4, 2007, pp. 321-329.

[14] de Bodas Terassi P.M., and Galvani E., Identification of homogeneous rainfall regions in the eastern watersheds of the State of Parana, Brazil. Climate, Vol. 5, Issue 3, 2017, pp. 53.

[15] Babel M.S., Agawal A., Swale D.K., and Herath S., Evaluation of climate change impacts and adaptation measures for rice cultivation in Northeast Thailand. Climate Research, Vol. 46, Issue 2, 2011, pp. 137-146.