IDENTIFICATION OF RAINFALL TEMPORAL PATTERNS

[1] Fadhilah Yusof [2] Nur Syereena Nojumuddin and [3] Zulkifli Yusop

[1,2]Department of Mathematical Sciences, Faculty of Science, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia.
[3]Instititute of Environmental and Water Resources Management, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia.

[1]nsyereena@gmail.com, [2]fadhilahy@utm.my, [3]zulyusop@utm.my

Abstract. Temporal pattern for rainfall events is required in the design and evaluation of hydrologic safety for hydraulic structures. In this paper, a method of statistical cluster analysis is applied to identify event-based representative temporal rainfall pattern in fourteen stations in Johor. The 8 hour separation time of no rain is used to divide the 5 years rainfall data into individual rainfall event according to the monsoons. The analysis is implemented from the dimensionless mass curve ordinates as the attributes in statistical cluster analysis of the cumulative storm depth over the non-dimensionalized time. As a result, three representative rainstorm patterns are identified and classified under the two basic types of advanced-type (A) and central-peak type (C) for northeast and southwest monsoon. Meanwhile, only one basic type is identified to represent rainstorm pattern during the inter-monsoon that is advanced type (A). In addition, the rainfall pattern is dependent on rainfall depth and duration, season and geographical location by contingency table test. The rainfall information presently used for design in Malaysia is very dissimilar to the representative curves derived in this study. The identification of three representative rainfall temporal patterns according to the monsoon seasons in Johor can be used as a basis to stochastically generate the plausible rainfall hyetographs of the specified pattern in Johor.

Keywords Mass curve and cluster analysis

INTRODUCTION

Design of hydraulic structures in river systems are often confronted with a lack of available data at specific point in a catchment owing to the models of hydrology and water quality often require long records of short-time increment precipitation data. These data are generally not available except at research experimental watershed facilities. If rainfall data are readily available, they can be served for the further analysis as the process of designing and analyzing of hydro systems need an appropriate rainfall pattern. For this reason, it is imperative that the designer knows the time distribution of rainfall so that runoff can be managed appropriately. For example, when considering storm-water management for a new site, a detention structure may be installed to detain the new runoff and release it at a controlled rate. If the detention system is improperly designed, it may inflate the cost or increase the risk. Therefore, an appropriate temporal pattern for a design storm is importance in the design and evaluation of hydrologic safety. Thus, a variety point of rainfall data products can be used in such design studies such as the historical time series, a synthetically generated time series, intensity-duration –frequency (IDF) relations and design storms.

Design storms are often used in urban drainage design for drainage and storm water management projects. Information about the storm event under consideration is often required to design a hydraulic structure. When designing a hydraulic structure to manage rainfall and runoff, a time distribution pattern (hyetograph) for the rainfall event is used to create a hydrograph. A complete description of a design storm involves the specification of storm depth, duration and its temporal pattern. A first design of hyetograph which is called as the Chicago method is introduced by Keifer and Chu in the year 1957 for the use in sewer design. It is based on an intensity duration curve for a specific return period. Hence, the method for determining the time-varying rainfall input in the absence of data is through mass curves by Huff has been developed in the year 1967. In the study of storm patterns in IIlinois urban areas, he presented a mass rainfall curve which the time distributions were classified into four groups depending on whether the maximum intensity occurred in the first, second, third, or fourth quarter of the duration. For each quarter, dimensionless time distributions were presented for various probability levels. However, according to Bonta (1987,1997), a problem with developing Huff curves is that they can be affected by factors such as season of year, climatic region; sampling-time interval of raw data, storm size, number of storm, etc.

In 1975, Pilgrim and Cordery demonstrated the method based on the rank of the rainfall depth in particular time periods. The average amount of rainfall in each rank was determined and rainfall pattern were developed. Later, in the

1980s, Yen and Chow applied the method of statistical moments to describing hyetographs. Data from over 9000 rainstorms at four locations were used (Boston, Massachusetts; Elizabeth City, North Carolina; San Luis Obispo, California; and Urbana, Illinois), with the analysis focused primarily on the first two moments with respect to the beginning of the event. They found that a general non-dimensional triangular hyetograph could be established, utilizing only the first moment. To characterize rainfall pattern there are methods which parametrically treat the dimensionless rainfall mass curve as a cumulative distribution function (CDF) of the dimensionless time. Hence, the same method of characterizing temporal rainfall events pattern were conducted by Yen and Chow (1983) and Fang and Tung (1996).

Twenty years later, the principle component analysis and cluster analysis have been used to classify hyetograph and hydrographs by Hannah et al (2000), Lana et al (2001) and Lin et al (2004). In addition, Wu S.J (2006) has done the studied on identification of representative rainfall temporal patterns in Hong Kong territory by using statistical cluster analysis. Last but not least, in 2008, Pelczer, I.J and Cisneros-Iturbe H.L. investigated on the identification of rainfall patterns over the Valley of Mexico of years 1993 till 2005 using clustering algorithm. The basic aim of cluster analysis is to find 'natural grouping' of a set of individuals by a relevant attributes. Cluster analysis allocates a set of individuals to a set of mutually exclusive and exhaustive groups such that individuals within the group are similar to one another whereas individuals between different groups are dissimilar. In general, the divisive (K-means method) and hierarchical (average and Ward's methods) clustering techniques can be applied to perform the cluster analysis and each method defines the similarity and the distance measurement somewhat differently. The result from different applications Fang and Tung (1996) and Ramos (2001) revealed that the effects of using methods of K-means, Ward and others on the final classification result were significant.

Moreover, Ramos (2001) also indicated that K-means method seems to be powerful enough to classify the observation. According to the study on the relative performance of using different attributes to classify rainfall patterns, Fang and Tung (1996) reported that by using statistical moments or fitted distribution parameters as the attributes, it is less desirable when compared with a direct use of dimensionless rainfall mass or hyetograph ordinates. By comparing the results of dimensionless rainfall mass (F-based) and hyetograph (P-based) classifications, the results are similar. However, the rainfall patterns obtained by the dimensionless rainfall mass (F-based) classification are more desirable, Wu S.J (2006). The main concern of this study is to identify the representative temporal rainfall pattern in Johor by event-based. Hence, the Euclidean distance based K-means clustering method is adopted herein to identify representative storm pattern occurred in Johor as the dimensionless rainfall mass curves ordinates are used as the attributes in statistical cluster analysis. Hence, the factor that affecting the rainfall pattern occurrence is investigates by using contingency table test. The representative curves result in the study is compare to the dimensionless hyetograph of the design storm given in the Urban Storm water Management Manual for Malaysia (DID, 2000).

Study region and data

Johor are situated in the southern part of Peninsular Malaysia and are located between the 1°20"N and 2°35"N latitudes. Johor has total land area of 19,210 km^2 and a population of about 3.2 million as of 2010. Johor has a humid tropical climate with monsoon rain from November to March blowing from the South China Sea. The average annual rainfall is 2355 mm with average temperatures ranging between 25.5 °C and 27.8 °C.

Hourly rainfall data in this study are obtained from Department of Irrigation and Drainage (DID) of the Government of Malaysia. Rainfall measured by 14 rain gauges in Johor was used in the study. The period of data available ranged from 2007 to 2011. The details of rain gauges are shown in Table 1 and the locations of the rain gauges are shown in Figure 1.

Table 1: Rain gauge stations used in Johor.

Station	Station name	Latitude	Longitude
J1	Bt 42 Jalan Kluang	1.85109	103.04854
J2	Ibu Bekalan Kahang	2.30866	103.66719
J3	Kg peta ulu sg Endau	2.54142	103.56136
J4	Kompleks Penghulu Sg Chaah	2.20860	103.03866
J5	Kompleks Perumahan Pontian	1.48682	103.39884
J6	Ladang Chan Wing	2.51052	102.83986
J7	Pusat Pertanian Endau	2.61329	103.63026
J8	Ladang Sg Gemas	2.22141	103.54748
J9	Ladang Kekayaan	1.61048	103.57577
J10	Sekolah Menengah Munshi Sulaiman	1.86838	102.9786
J11	Pintu Kawalan Parit Bintang	1.9344	103.3587
J12	Stor JPS Endau	2.64142	103.6614
J13	Ladang Getah Kukup	1.33194	103.4606
J14	Ladang Gunung Pulai	2.03437	103.3304

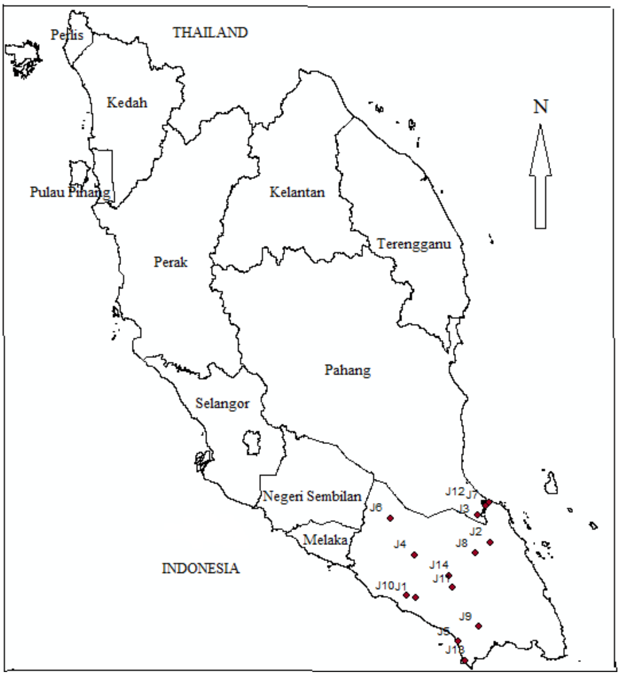

Figure 1: Location of the rain gauge stations in Johor used in the study

IDENTIFICATION OF RAINSTORM PATTERN

Determination of rainfall event

To divide the long continuously rainfall data into rainfall event, the method adapted from Adams (2000) has been applied to all the fourteen stations to define a suitable minimum inter-event time value. The results obtained, suggested that the rainfall event is identified by using 8 hour of separation time of no rainfall period. Therefore, 8 hour separation time of no rain is used to divide the rainfall data into individual rainfall event. The rainfall data is divided into northeast, southwest and inter-monsoons. According to the study of Wu, S.J (2006), the storm duration of at least 3 h might be adequate to reflect the overall temporal variation of hyetograph as the rainfall duration equal to the multiples of 12 are generally not directly available in the computation of the cumulative rainfall fraction. Hence, rainfall events with duration less than or equal to 2h are excluded from the analysis in the study.

Characterization of Rainfall Patterns

During a rainfall event, the precipitation amounts differ with respect to time at a given location. To identify the different storm patterns for the observed storm events, temporal patterns of storm events should be characterized. As rainfall duration and total depth vary from one event to another, characterization of similarity or dissimilarity of different rainfall patterns can best be made through the use of a dimensionless scale. By using dimensionless rainfall mass curves or hyetograph, the effect of rainfall duration and depth are removed, leaving the temporal variation of rainfall event as the only factor for differentiating events of different patterns. After non-dimensionalization, rainfall event with different depths and durations can be combined, examined and categorized for identifying representative rainfall patterns.

To characterize the storm temporal pattern using the rainfall mass curve, the selection of the number of time points over the storm duration is subjective. Too few points may not accurately describe the underlying variation of storm pattern. On the other hand, too many points would capture unwanted sample noises masking the essential feature of storm pattern. According to a study on Wyoming by Fang and Tung, Y.K, the entire storm duration was divided into six equal intervals and Wu, S.J (2006) divide the entire duration into 12 equal intervals. Therefore, in this study, a non-dimensionalized rainfall mass curve is divided into 10 intervals and the corresponding dimensionless rainfall mass ordinates $F\tau$ at t= $j/10$ with j=1, 2… 10 are used to represent a rainstorm pattern as its suitable with the rainfall characteristic used in the study. Non-dimensionalized rainstorm pattern can be obtained by adjusting the scale of the duration and depth of a rainfall mass curve as

$$\tau = \frac{t}{d}\ ;\ \ F_\tau = \frac{D_{\tau\times d}}{D_d} \qquad (3)$$

in which τ is the dimensionless time, $\tau \epsilon (0,1]$; d is the storm duration; F_τ is the dimensionless cumulative rainfall representing the cumulative fraction of rainfall depth, $F_\tau \epsilon (0,1]$; D_τ is the cumulative rainfall depth at time $t(t = \tau \times d)$; and D_d is the is the total rainfall depth. The dimensionless mass curve shows the cumulative fractions of storm depth, $0 \leq F_\tau \leq 1$ over the non-dimensionalized time, $0 \leq \tau \leq 1$.

Cluster Analysis for Identification Representative Rainfall Patterns

As rainstorm pattern varies among events, it is useful to categorize them into several representative types so that individual rainfall patterns within each type are similar to one another, but not necessarily identical, whereas individual rainfall patterns between different types are dissimilar. Euclidean distance based on K-means clustering method by MacQueen (1967) is adopted herein to classify pattern typically occurred in Johor. Hence, the study used dimensionless rainfall mass curves ordinates denoted as F-based ordinates as the attributes in statistical cluster analysis to identify representative rainfall patterns in Johor. The representative of rainfall pattern can be determined after the rainfall event data is classified into 3,4,5 and 6 group. To remove the scale effects of the attributes used in the cluster analysis, attributes are standardized so that they have a zero mean and a unit standard deviation. The proper number of rainfall patterns can be determined by identifying the appropriate number of group that results from the k-means clustering analysis. Since the appropriate number of group to represent rainfall pattern is not known in advanced, it is commonly determined by a trial-and-error process of visually examining the averaged dimensionless mass curve for each group resulting from the cluster analysis. The better clustering is when the similarity within a group is greater and the difference between groups is greater. The K-means clustering algorithm is performed by using the three steps until its convergence. The three steps are (1) determine the centroid coordinate and by using k=3, 4, 5 and 6; (2) determine the distance of each object to the centroids and (3) group the object based on minimum distance. Repeat the three steps above until convergence or no object move group.

The centroid coordinates of the points $(x_1\ ,\ y_1),\ (x_2\ ,\ y_2),\ \ldots,\ (x_n, y_n\ )$is given by equation (1);

$$\left(\frac{x_1+x_2+x_3+\ldots+x_n}{n}\ ,\ \frac{y_1+y_2+y_3+\ldots+y_n}{n}\right) \qquad (1)$$

The Euclidean distance formula is given in equation (2);

$$D = \sum_{j=1}^{k}\sum_{i=1}^{n}\left\|x_i{}^j - c_j\right\|^2 \qquad (2)$$

Where $\left\|x_i{}^j - c_j\right\|^2$ is a chosen distance measure between a data point, $x_i{}^j$ and the cluster center, c_j, is an indicator of the distance of the *n* data points from their respective cluster centers.

Factors Affecting the Rainfall Pattern

As the identification of rainfall pattern based on the statistical cluster analysis, the rainfall pattern is derived from the dimensionless rainfall mass curve of many rainfall events with different durations and depths that have occurred at different seasons at various locations in Johor. The investigation of whether the occurrence frequency of one particular rainfall pattern might be affected by the season, geographical location or rainfall depth and duration can be done by statistical contingency table introduced by (Conover, 1980). In the study, the chi-square test is conducted to compute the p-value. The degree of freedom associated with the chi-square test, which equal to $(n_c-1)(n_r-1)$ with n_c and n_r being the number of column and row, respectively. In practice, the significance level of 5% is used for decision making. If the p-value is lower than 5%, the factor under consideration is considered to have significant effect on the occurrence frequency of various rainfall patterns.

Comparison of Representative Rainstorm Pattern with the Malaysian Design Storm

The dimensionless hyetograph of the design storm given in the Urban Storm water Management Manual for Malaysia (DID, 2000) is also compared with the best representative rainfall pattern identify in the study. The development of temporal rainfall patterns in the manual is based on an approach discussed in ‘‘Australian rainfall and runoff’ (DID, 2000, chapter 13, pp. 5–16). The dimensionless hyetographs are presented for durations of 10, 15, 30, 60, 120, 180, and 360 min. The comparison is between the best representative rainfall pattern result in the study and 2-h West Coast design storm given in the manual (DID, 2000).

RESULT AND DISCUSSION

Determination of rainfall event

The total number of storm events occurred in Johor by monsoon season during the five years period is shown in Table 2. It is noted that northeast (November-Mac) has the highest rainstorm occurrence with 3404 rainstorm events followed by southwest (May-September) with 3245 rainstorm events. Meanwhile, the two inter-monsoons reported 1324 number of storm event occurrences.

Table 2: Number of rainstorm event occurrences by monsoon season during 2007-2011.

Station	Number of storm event during southwest (May-September)	Number of storm event during northeast (November-Mac)	Two inter-monsoon (April-October)
J1	352	295	121
J2	212	172	86
J3	187	197	80
J4	188	189	61
J5	233	213	102
J6	191	205	74
J7	245	257	95
J8	123	171	64
J9	195	195	83
J10	247	264	100
J11	260	253	114
J12	218	275	82
J13	296	352	132
J14	298	366	130
Total	3245	3404	1324

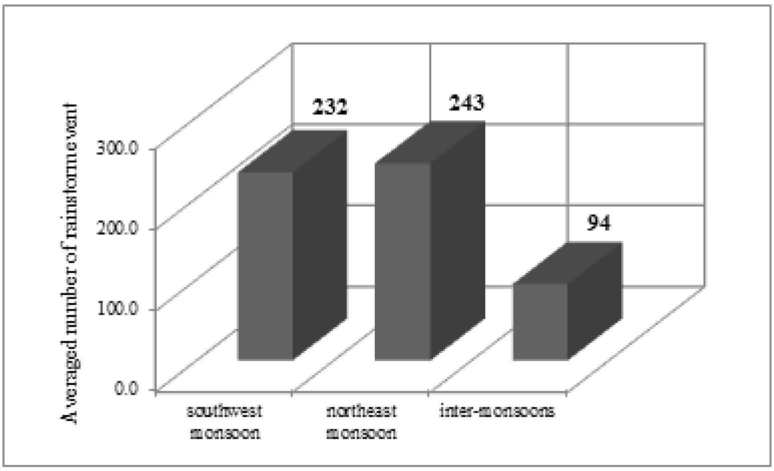

Figure 2: Average number of rainstorm event by monsoon

Figure 2 shows the average number of rainstorm events for all the fourteen stations by monsoon. Northeast monsoon received the highest average number of rainstorm event with 243, followed by southwest monsoon with 232 and inter-monsoon with 94.

Table 3: Statistics of storm duration

Station	Duration (h)					Total
	3-5	6-11	12-17	18-23	≥ 24	
J1	113	396	73	20	50	652
J2	63	146	25	16	146	396
J3	52	133	30	20	171	406
J4	73	141	26	20	94	354
J5	113	134	27	17	110	401
J6	67	133	20	23	150	393
J7	71	254	50	23	55	453
J8	54	97	28	12	65	256
J9	61	128	27	13	152	381
J10	120	237	44	15	19	435
J11	127	250	46	15	15	453
J12	78	217	50	24	51	420
J13	142	358	60	18	18	596
J14	144	374	61	13	36	628
Total	1278	2998	567	249	1132	6224

Table 3 shows the statistics of storm duration during the five years rainfall data from 2007-2011 for each station. The storm event duration which is less than 3h was removed. As a result, 6224 storm event is extracted for further analysis. From the table, the rainstorm event was grouped into five durations. It can be observed that the rainstorm event with duration of 6-11 hour has the highest number of occurrence. The second highest of rainstorm event occurrences was during the duration of 3-5 hour. Figure 3 shows the average number of rainstorm events for all the fourteen stations by different storm duration. The highest average number of rainstorm event for each station is 214 during duration of 6-11h, followed by 91 during duration of 3-5h.

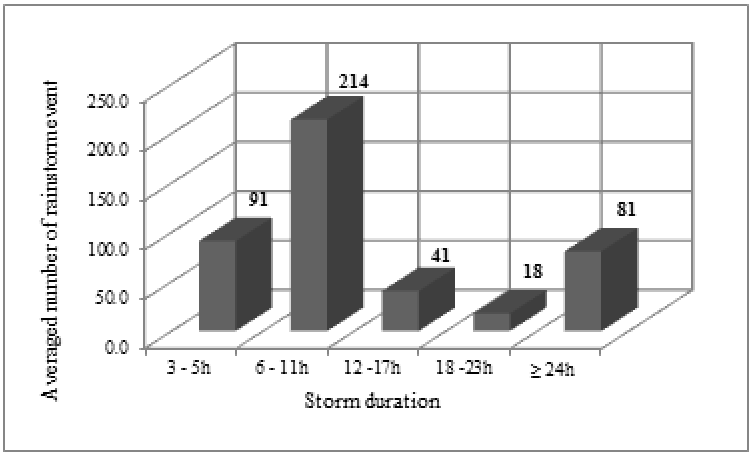

Figure 3: Average number of rainstorm events by different storm duration.

Characterization of Rainfall Patterns

As a non-dimensionalized rainfall mass curve is divided into 10 intervals to represent rainfall mass curve ordinates, therefore, Figure 4 shows the mean of event depth according to the duration interval according to the monsoon season. Each interval in the figure represents the two intervals of the rainfall mass curve ordinates used in the study. It can be observed that during the northeast and inter-monsoons, the largest depth occurs in the beginning part of interval of 3-12h and the smallest depth occurs when the duration is greater than 43h which is at the ending part of duration. However, during the inter-monsoon period, the smallest depth occurs in all intervals (second, third, fourth and fifth) except for the first interval. Meanwhile, the southwest monsoon has the largest depth during the center part of duration in interval of 13-22h. The ratio of the largest depth (first interval) to the smallest depth (last interval) during northeast, southwest and inter-monsoon is 1:7, 1:6, and 1.9, respectively.

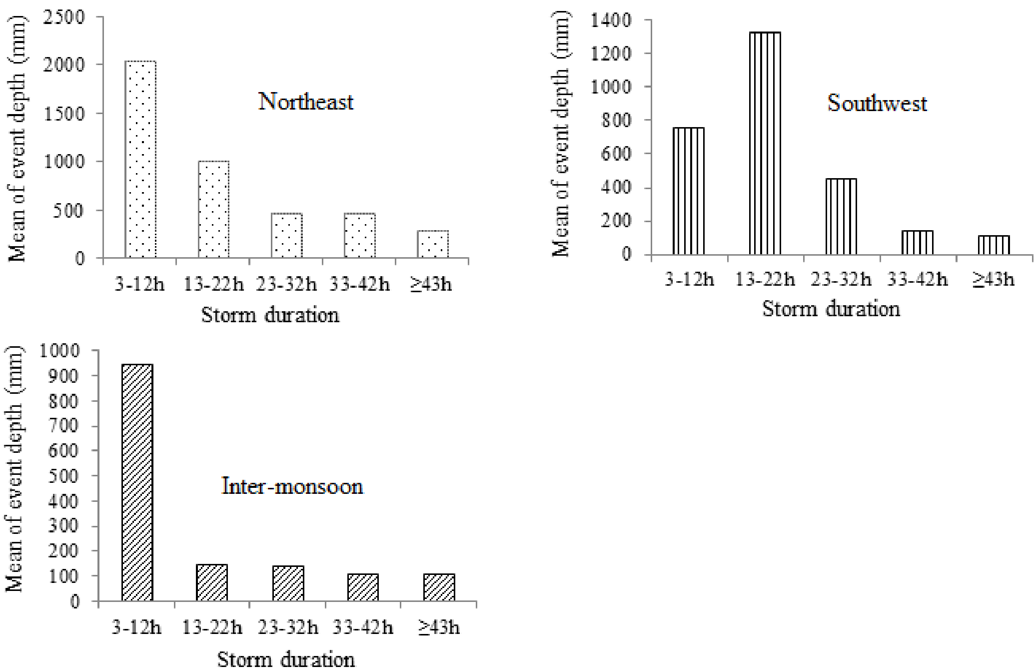

Figure 4: Mean of event depth by monsoon seasons

Cluster Analysis for Identification Representative Rainfall Patterns

Figures 5-7 show the average rainfall mass curve under 3, 4, 5 and 6 groups of classification based on the monsoon seasons. In these figures the graphs of the cumulative fraction of rainfall depth over dimensionless time are shown, hence the gradient of the curve represents the intensity of the rainfall pattern. Figure 5, 6 and 7 represent the rainfall temporal patterns during the northeast, southwest and inter-monsoon, respectively. It can be observed that most of the curves recorded the highest intensity during the early part for northeast and inter-monsoon whereas during southwest the curves recorded the highest intensity during the center part.

It can be observed that each monsoon has its own rainfall pattern. For northeast monsoon, by referring to Figure 5, most of the curves have the highest gradient at the beginning part of dimensionless time except for one curve where the highest gradient is at the center of the dimensionless time. Moreover, for southwest monsoon, by referring to Figure 6, the most curves have highest gradient at the center part of dimensionless time except for two curves where the highest gradient is at the beginning part of the dimensionless time for each group. Meanwhile, for inter-monsoon, by referring to Figure 7, all curves have the highest gradient at the beginning part of dimensionless time for all the four groups.

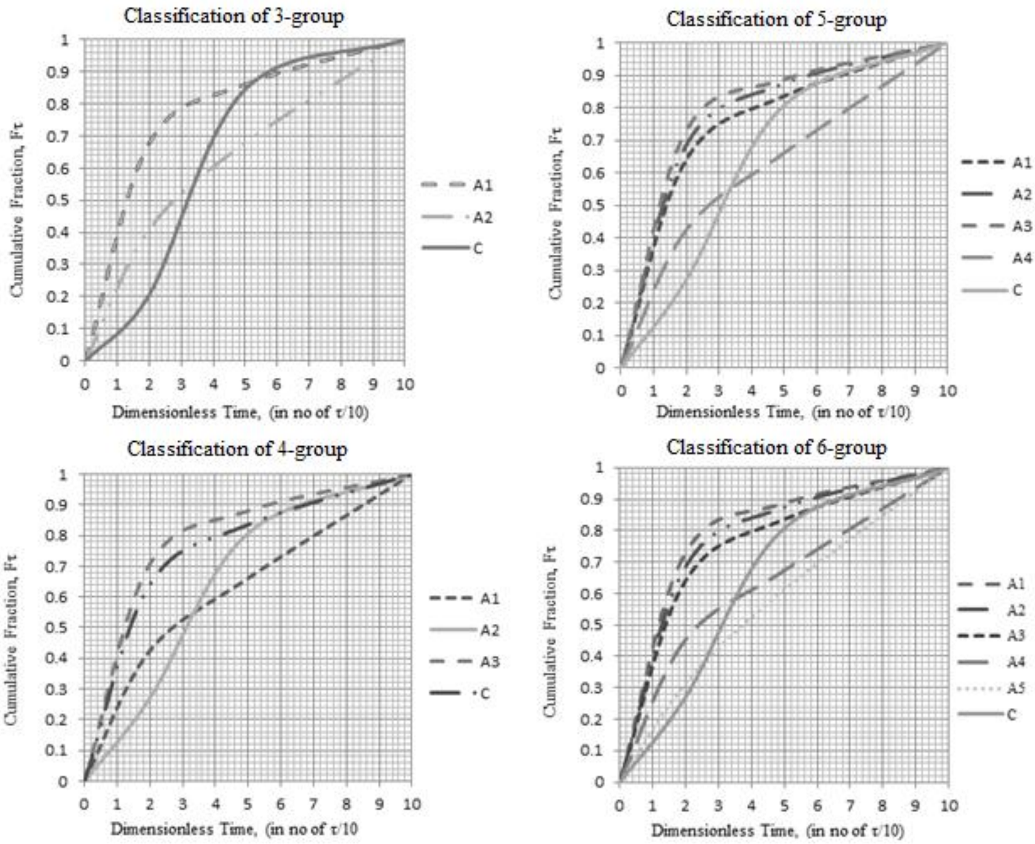

Figure 5: Rainstorm pattern for northeast

The shape of the rainfall patterns can be categorized according to the Wu, S.J (2006) into three basic types of rainfall patterns namely; advanced type, central-peaked type and delayed type which are denoted as A, C and D respectively. The advanced patterns have relatively high rainfall intensity during early part of the rainfall event. The central-peaked pattern has relatively high intensity in the center part of the rainfall event and the intensity tapers off towards the beginning and ending of the rainfall event. The delayed type D on the other hand has relatively high rainfall intensity during later part of the rainfall event. However in this study, the delayed type, D is not found.

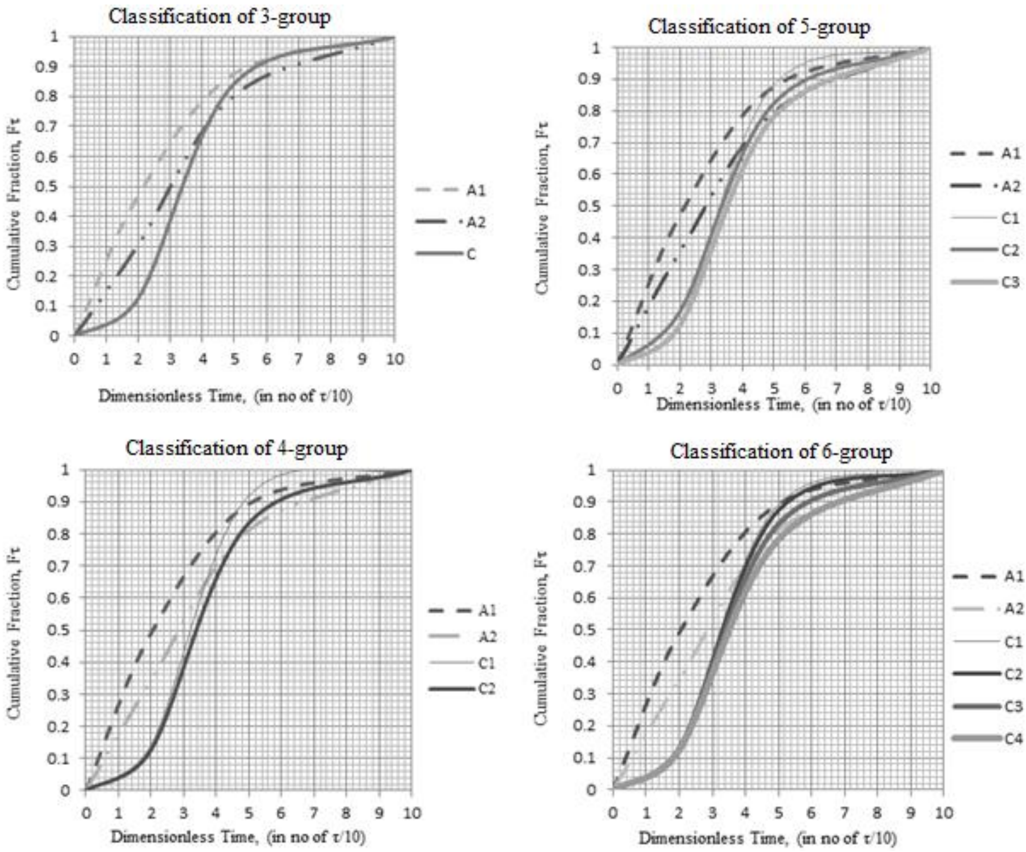

Figure 6: Rainstorm pattern for southwest

According to the above guidelines, Figure 5 demonstrates that in each group for the northeast monsoon, the rainfall patterns consist of central-peaked type (C) and advanced type (A1 and A2); central-peaked type (C) and advanced type (A1, A2 and A3); central-peaked type (C) and advanced type (A1, A2, A3 and A4); and central-peaked type (C) and advanced type (A1, A2, A3, A4 and A5) respectively. Figure 6 indicates that in the southwest monsoon, the rainfall pattern consist of central-peaked type (C) and advanced type (A1 and A2); central-peaked type (C1 and C2) and advanced type (A1 and A2); central-peaked type (C1, C2 and C3) and advanced type (A1 and A2) and central-peaked type (C1, C2, C3 and C4) and advanced type (A1 and A2) for each group. Figure 7 illustrates that all group consists of advanced type (A1, A2 and A3); (A1, A2, A3 and A4); (A1, A2, A3, A4 and A5) and (A1, A2, A3, A4, A5 and A6) during the inter-monsoon period.

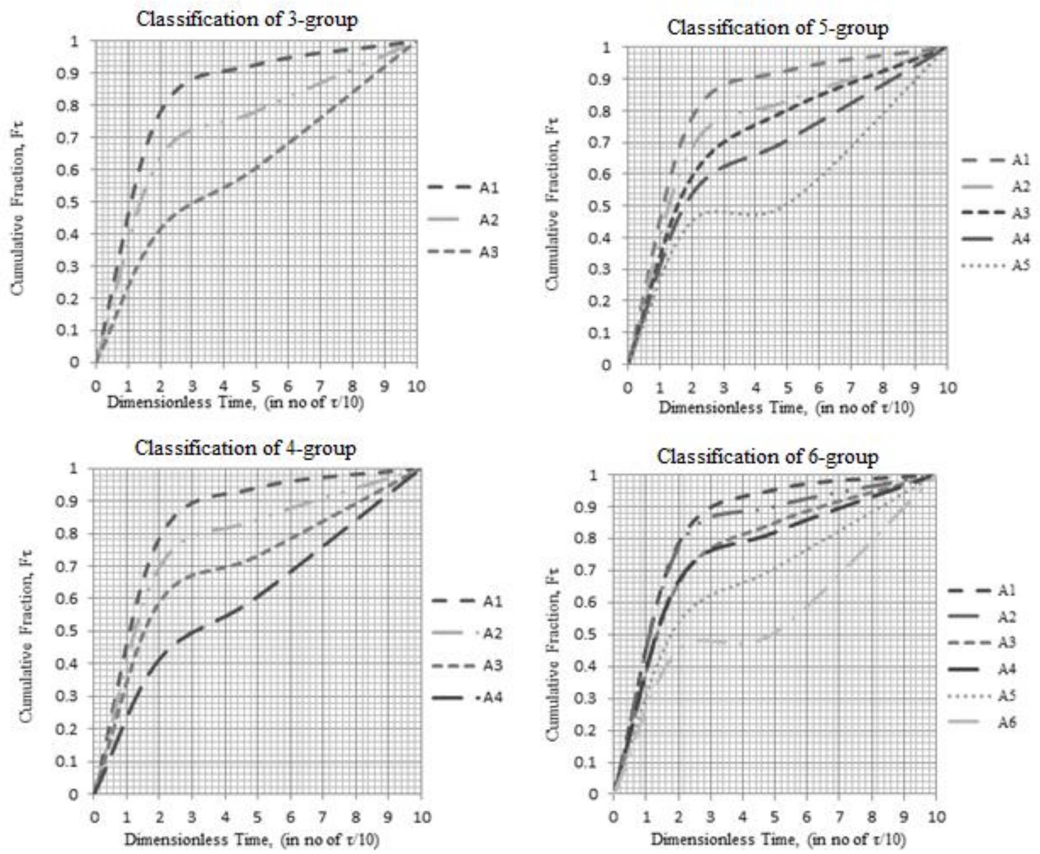

Figure 7: Rainstorm pattern for inter-monsoon

The k-means clustering analysis used in the study is to divide the rainfall event data into 3, 4, 5 and 6 group to represent a rainfall pattern. But, the fixed number of group can make it difficult to predict what the best representative rainfall group should be. However, the best judgment can be made is that the object within a group must be similar to one another and different from the objects in other group. Therefore, by comparing the result of the representative of all groups in Figure 5 for northeast, it can be observed that the entire group has a central peak type (C) and advanced type (A1 and A2) but (A) increases as the number of group increases. Hence, the figure demonstrates that the curve shape of advanced type in group 4-6 (A3, A4 and A5) is similar to one and another and not different with (A1 and A2). . Therefore, the advanced type in group 4-6 is characterized as the same group. By comparing all groups in Figure 6 for southwest, it can be observed that the entire group has the advanced peak type (A1 and A2) but the central peak type (C) increases as the number of group increases. In other words, the central peak type in group 4-6 is characterized as the same group as the shape of the curve (C2, C3 and C4) is similar to one and another and not different from the (C1). In addition, the (A1 and A2) in the 3 group of classification in Figure 5 and 6 are characterized as the different group as the shape of these two curves is not same although they are the same type. As a result, the 3 group is the best representative rainfall pattern present during northeast and southwest monsoon in Johor. The three patterns that are under the two basic types of rainfall patterns are advanced-type (A1 and A2) and central-peak type (C). Meanwhile, by comparing the result of all groups in Figure 7, the three group of classification is the representative of rainfall temporal distribution profiles during inter-monsoon as the three advanced type (A1, A2 and A3) are not similar in curve shape and difference between each group although it represent the same basic type of rainfall pattern. Therefore, basic type for inter-monsoon is advanced pattern (A1, A2 and A3).

The result is satisfactory with the result shows in Figure 4 which shows that the largest depth occurs in the beginning part of interval during the northeast and inter-monsoons. Meanwhile, the southwest monsoon has the largest depth during the center part of duration. In addition, the smallest depth occurs at the ending part of duration for all monsoons. As the intensity is proportionally to the depth, therefore the largest depth give the meaning of the highest intensity and smallest depth give the meaning of the lowest intensity during that interval. Therefore, only two basic type occurred in the study in Johor which are advanced-type and central-peaked type with the shape of actual rainfall event are relatively high rainfall intensity during early part of the rainfall event and high intensity in the center part of the rainfall event. The identification of three representative rainfall temporal patterns in Johor can be used as the basis to stochastically generate the plausible rainfall hyetographs of the specified pattern for hydro system engineering applications by realizing the existence of inherent variability in rainfall pattern of individual event within each type. In addition, the development of the event-based storm patterns is useful for synthesizing and simulating storm events.

Factors Affecting the Rainfall Pattern

Table 4: Contingency table of four typical rainfall in Johor.

	Rainfall patterns				Column Total
	A1	A2	A3	C	
Rainfall duration (h)					
3-12	2426	1766	96	1614	5902
13-22	261	256	12	184	713
23-32	63	122	13	47	245
33-42	35	65	7	25	132
≥43	34	99	10	21	164
Row Total	2819	2308	1891	138	7156
χ^2 = 187		DF = 12		*p*-value = 0.00	
Rainfall depth (mm)					
0-70	1302	1022	62	929	3315
71-140	505	420	20	305	1250
141-210	314	224	16	194	748
211-280	190	157	9	132	488
281-350	155	96	7	100	558
351-490	160	133	10	112	415
≥490	89	129	7	62	287
Row Total	2715	2181	131	1834	6861
χ^2 = 39.3		DF = 18		*p*-value = 0.003	
Season					
Northeast	1657	1333	0	217	3207
Southwest	797	328	0	1674	2799
Inter-monsoon	365	647	138	0	1150
Row Total	2819	2308	138	1891	7156
χ^2 = 0.0004		DF = 6		*p*-value = 0.00	
Region					
Western	1353	599	0	297	2249
Northern	1596	586	0	114	2296
Southern	101	238	0	218	557
Eastern	281	508	138	297	1224
Row Total	3331	1931	138	926	6326
χ^2 = 0.0002		DF = 9		*p*-value = 0.00	

By referring to the Table 4, the contingency table shows the p-values of 0.00 (< 0.05) for rainfall duration, depth, season and region. As the significance level of 5% is used in the study, therefore it can be concluded that the occurrence frequency of various rainfall patterns in Johor are affected by the rainfall depth, duration, season and the geographical location. The result from the contingency tables analysis provides the basis for choosing rainfall types that are most likely to occur in Johor. The selection of the appropriate rainfall pattern cannot be randomly done as the frequency of occurrence of different rainfall patterns is dependent on rainfall depth and duration, seasonal and regional

Comparison of Identification Rainstorm Pattern with the Malaysian Design Storm

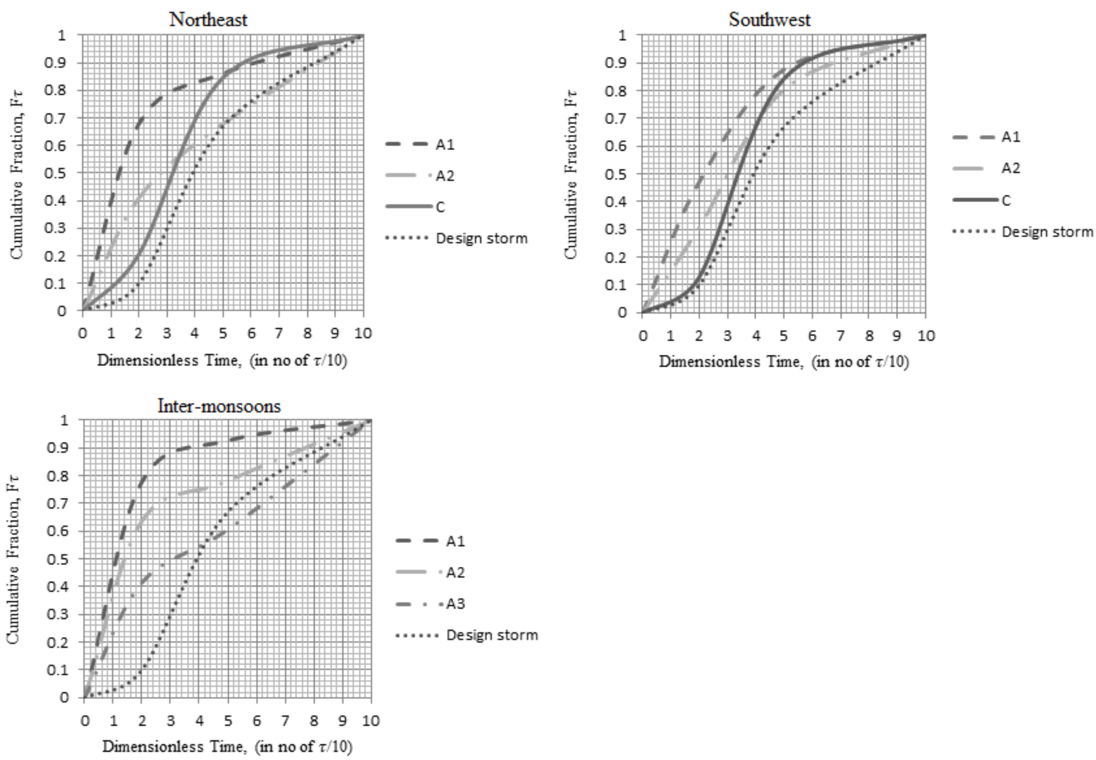

Figure 8: Comparison between the best representative rainfall pattern in Johor and 2-h West Coast design storm given in the manual (DID, 2000)

As expected, the design storm given in the design manual (DID, 2000) cuts across the representative rainfall curve. As storms identified in west coast Peninsular Malaysia are influenced by southwest monsoon, the design storm curve is closest to the southwest monsoon in Figure 8. The situation appears to get worse in the inter-monsoon representative rainfall curve presented. However, the design storm represents the central-peaked type which has a high intensity at the central part of the rainfall event. Therefore, it may be concluded that the design storm in DID (2000) does not represent the ‘‘typical” temporal pattern but it represent the basic type of rainfall in Malaysia.

SUMMARY AND CONCLUSION

The temporal rainfall pattern is required in rainfall-runoff modeling to produce flow hydrographs in design and analysis of hydro systems. The event-based representative temporal rainfall pattern in fourteen stations in Johor has been identified by using statistical cluster analysis. The 8 hour separation time of no rain is used to divide the rainfall data into individual rainfall event and the rainfall event is divided into northeast, southwest and inter-monsoons. The study used dimensionless rainfall mass curves ordinates as the attributes in statistical cluster analysis. A non-dimensionalized rainfall mass curve is divided into 10 intervals and the corresponding dimensionless rainfall mass curve ordinates are used to represent a rainstorm pattern. The Euclidean distance based K-means clustering method is adopted herein to classify pattern typically occurred. Since the appropriate number of representative rainfall patterns is not known in advanced, it is commonly determined by a trial-and-error process. The appropriate number of rainfall patterns is determined by examining the averaged dimensionless rainfall mass curves for each group resulting from the cluster analysis. As a result, three representative rainstorm patterns are identified. The three patterns classified under the two basic patterns are advanced-type (A) and central-peak type (C) during northeast and southwest monsoon. Meanwhile, only one basic type is identified to represent rainstorm pattern during the inter-monsoon that is advanced type (A). The statistical contingency table is used to investigate whether the occurrence frequency of one particular rainfall pattern might be affected by the season, geographical location or rainfall depth and duration. As a result, the occurrence frequency of rainfall patterns in Johor is affected by the rainfall depth, duration, season and the geographical location. The comparison between the best representative rainfall pattern and 2-h West Coast design storm given in the manual (DID, 2000) reveal that the design storm given in (DID, 2000) does not correspond to the representative curves derived from Malaysian data. This identification of three representative rainstorm patterns can be used as a basis in generating the rainfall hyetographs for Johor.

ACKNOWLEDGMENTS

The authors are grateful to the Malaysian Drainage and Irrigation Department for providing the hourly precipitation data. The work was financed by the Research University Grant (RUG) Q.J130000.2526.08H21 by University Technology Malaysia.

REFERENCES

[1] Adams, B.J and Papa, F. 2000. *Urban stormwater management planning with analytical probabilistic models*, John Wiley & Sons, Inc, Canada.
[2] Bonta, J.V. 1997. Proposed use of Huff Curves for hyetograph characterization. *In: Proceedings of the Workshop on Climate and Weather Research. Denver, Colorado.*223: 111-124.
[3] Bonta, J.V., Rao, A.R. 1987. Factors affecting development of Huff curves. *Transactions of the ASAE 30(6)*: 1689-1693.
[4] Fang, T.Q., Tung, Y.K. 1996. Analysis of Wyoming extreme precipitation patterns and their uncertainty for safety evaluation of hydraulic structure. Technical Report, WWRC-96.5, Wyoming Water Resource Centre, University of Wyoming, Laramie, Wyoming.
[5] Hannah, D.M., Smith B.P.G., Gurnel, A.M., McGregor RM (2000). An approach to hydrograph classification. *Hydrol Process* 14: 317-338.
[6] Huff, F.A. 1967. Time distribution of rainfall in heavy storms, *Water Resources Research* 3(4): 1007-1019.
[7] Keifer, C.J and Chu, H.H. 1957. Synthetic rainfall pattern for drainage design, *ASCE J Hydraulic Div* 83 (HY4) :1-25.
[8] Lana, X; Serra, C; Burgueno, A. 2001. Patterns of monthly rainfall shortage and excess in terms of the standardized precipitation index for Catalonia (Ne Sapin). *Int J Climatol* 21: 1669-1691.
[9] Lin, G.F., Chen, L.H., Kao, S.C. (2004). Development of regional design hyetograph. *Hydrol Process (in press).*
[10] MacQueen, J. 1967. Some methods for classification and analysis of multivariate observations. *In: Proceedings of the 5th Berkeley symposium* 1: 281-297.
[11] Pilgrim, D.H.and Cordery, I. 1975. Rainfall temporal patterns for design floods, *J Hydraulic Div ASCE* 101 (HY1): 81-95.
[12] Pelczer, I.J and Cisneros-Iturbe H.L. (2008): Identification of rainfall patterns over the Valley of Mexico. *Proceedings of 11th International Conference on Urban Drainage, Edinburgh,* Scotland, UK.

[13] Ramos, M.C. 2001. Divisive and hierarchical clustering techniques to analyse variability of rainfall distribution patterns in a Mediterranean region. *Atmos Res* 57: 123-138.
[14] Yen, B.C & Chow, V.T. 1980. Design hyetographs for small drainage structures. *J. Hydraul. Div. ASCE,* 106(HY6), 1055-1076.
[15] Yen, B.C., Chow, V.T. 1983. Local design rainfall, vol II-methodology and analysis. Report, FHWA/RD-82/064, Federal Highway Administration, US Department and of Transportation, Washington, DC,USA.
[16] Wu, S.J., Tung, Y.K., Yang, J.C. 2006. Identification and stochastic generation of representative rainfall temporal patterns in Hong Kong territory. *Stoch Environ Res Risk Asse*ss 20(3): 171-183.