International Journal of Engineering, Mathematical and Physical Sciences
ISSN: 2517-9934
Vol:10, No:5, 2016

Daily Probability Model of Storm Events in Peninsular Malaysia

Mohd Aftar Abu Bakar, Noratiqah Mohd Ariff, Abdul Aziz Jemain

***Abstract*—**Storm Event Analysis (SEA) provides a method to define rainfalls events as storms where each storm has its own amount and duration. By modelling daily probability of different types of storms, the onset, offset and cycle of rainfall seasons can be determined and investigated. Furthermore, researchers from the field of meteorology will be able to study the dynamical characteristics of rainfalls and make predictions for future reference. In this study, four categories of storms; short, intermediate, long and very long storms; are introduced based on the length of storm duration. Daily probability models of storms are built for these four categories of storms in Peninsular Malaysia. The models are constructed by using Bernoulli distribution and by applying linear regression on the first Fourier harmonic equation. From the models obtained, it is found that daily probability of storms at the Eastern part of Peninsular Malaysia shows a unimodal pattern with high probability of rain beginning at the end of the year and lasting until early the next year. This is very likely due to the Northeast monsoon season which occurs from November to March every year. Meanwhile, short and intermediate storms at other regions of Peninsular Malaysia experience a bimodal cycle due to the two inter-monsoon seasons. Overall, these models indicate that Peninsular Malaysia can be divided into four distinct regions based on the daily pattern for the probability of various storm events.

***Keywords*—**Daily probability model, monsoon seasons, regions, storm events.

I. Introduction

RAINFALL analysis is important, especially for countries which experiences rain all year round like Malaysia. By analyzing rainfall characteristics, researchers could model rainfall dynamics, build rainfall’s spatial and temporal profile as well as predict future rainfall events. All this will subsequently help in the development and management of various sectors such as agriculture and economics.

There are various methods used in processing and extracting useful information from rainfall data. Two common methods are the moving windows approach and SEA. The moving windows approach uses accumulated rainfalls within predetermined time intervals and hence is usually referred to as a univariate analysis [1]. Meanwhile, the SEA approach uses the definition of storms to analyze rainfalls in terms of storm events. The definition of storms relies on the inter-event storm definition (IETD) which is the minimum time between two consecutive events. The IETD is chosen such that the serial correlation between two events is minimized [2].

Researches regarding the value of IETD are mostly done on rainfall data from Canada and the United States of America. These researches indicate that, for small urban catchments like most rainfall stations in Peninsular Malaysia, the IETD is six hours [3]–[5]. Meanwhile, based on rainfalls data in Korea and Australia, the value of IETD is determined to be between one and twelve hours with the interval of one to six hours is suggested for most urban catchments [6], [7]. In this study, the IETD value chosen is six hours due to the insignificance difference between the mean annual total storms produced between taking IETD values of six and seven hours.

For hourly rainfalls data, if two consecutive wet hours (hours with rainfall values more than 0 mm) are separated by a dry period which is longer than the IETD value, then the two wet hours are considered to be from two different storm event. Meanwhile, if the opposite is true, then both wet hours are part of the same storm event. The accumulated rainfalls within a storm event is called the storm amount while the length of the storm event is known as the storm duration [8].

Spatial and temporal analysis of rainfalls characteristics are frequently done over the years in Peninsular Malaysia. However, most of the analyses are based on the moving windows approach. Among studies that have been done on rainfalls in Peninsular Malaysia are trend analyses of daily rainfalls during monsoon seasons [9]–[11] and pattern analyses on the difference between daily rainfall distribution at different regions [12], [13]. There are also studies that investigate the various rainfalls characteristics in Peninsular Malaysia [14]–[16], and their dependence towards space and time [17].

The study of seasonality is important since climate cycles is a source of variation in many natural world processes [18] and it can be linked to other processes in the ecosystem such as wildfire [19]. Due to the importance of seasonality, it has been incorporated into various rainfall models [20]–[22]. This study aims to model the daily probability of storm events in Peninsular Malaysia. By modelling the probability of daily storm occurrence, researchers as well as experts in various fields could investigate the interannual variability, the onset and end of season, and the factors and the types of storms affecting the regional hydrologic cycle of different areas of a country.

Mohd Aftar Abu Bakar is with the School of Mathematical Sciences, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, Bangi, Selangor 43600 Malaysia (phone: 603-8921-5784; fax: 603-8925-4519; e-mail: aftar@ukm.edu.my).
Noratiqah Mohd Ariff is with the School of Mathematical Sciences, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, Bangi, Selangor 43600 Malaysia (phone: 603-8921-5755; fax: 603-8925-4519; e-mail: tqah@ukm.edu.my).
Abdul Aziz Jemain is with the School of Mathematical Sciences, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, Bangi, Selangor 43600 Malaysia (phone: 603-8921-5724; fax: 603-8925-4519; e-mail: azizj@ukm.edu.my).

ISSN: 2517-9934
Vol:10, No:5, 2016

II. Storm Events in Peninsular Malaysia

Hourly rainfalls data for rainfall stations which are obtained from the Department of Irrigation and Drainage Malaysia are used in this study. The rainfall stations chosen are scattered across Peninsular Malaysia and are shown in Fig. 1. In the figure, four regions are roughly determined based on the location of states and the geographical characteristics of Peninsular Malaysia given by Tourism Malaysia and previous studies [23], [24].

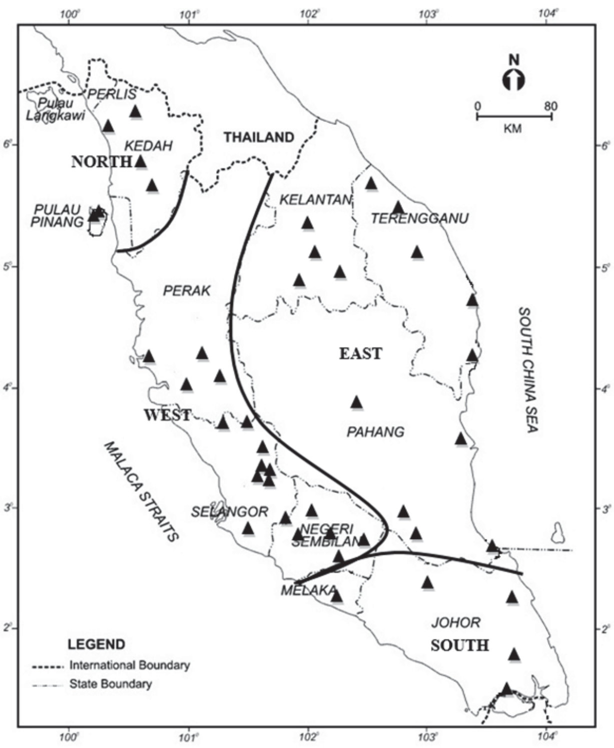

Fig. 1 Rainfall stations in Peninsular Malaysia

On average, Peninsular Malaysia experiences 168 storms annually at each rainfall station. The western region of Peninsular Malaysia received the most storms with an average of 178 storms annually for each station while rainfall stations at the northern and eastern regions have an average of 164 and 169 storms per year respectively. Meanwhile, the rainfall stations at the southern region recorded the least number of storms with an average of 152 storms each year.

Majority of the storms lasted less than 12 hours long. For example, Fig. 2 shows the histograms for storm duration at four rainfall stations; one from each region; for the length of 39 years. The shaded blocks show the frequency of storms with storm duration less than or equal to 12 hours.

The mean storm duration for storms in Peninsular Malaysia is 6.23 hours per storm and 937 hours a year. Regions at the east and south of Peninsular Malaysia experiences longer period of rain compared to other regions since both regions have longer annual total storm durations. East Peninsular Malaysia recorded 1054.92 hours of storm duration a year while west Peninsular Malaysia recorded the shortest time with 845.92 hours of storm duration annually. Individually, storms in the east of Peninsular Malaysia also have longest storm duration with a mean of 6.41 hours per storm event. The western region of Peninsular Malaysia also has the shortest mean of storm duration with 4.99 hours per storm event.

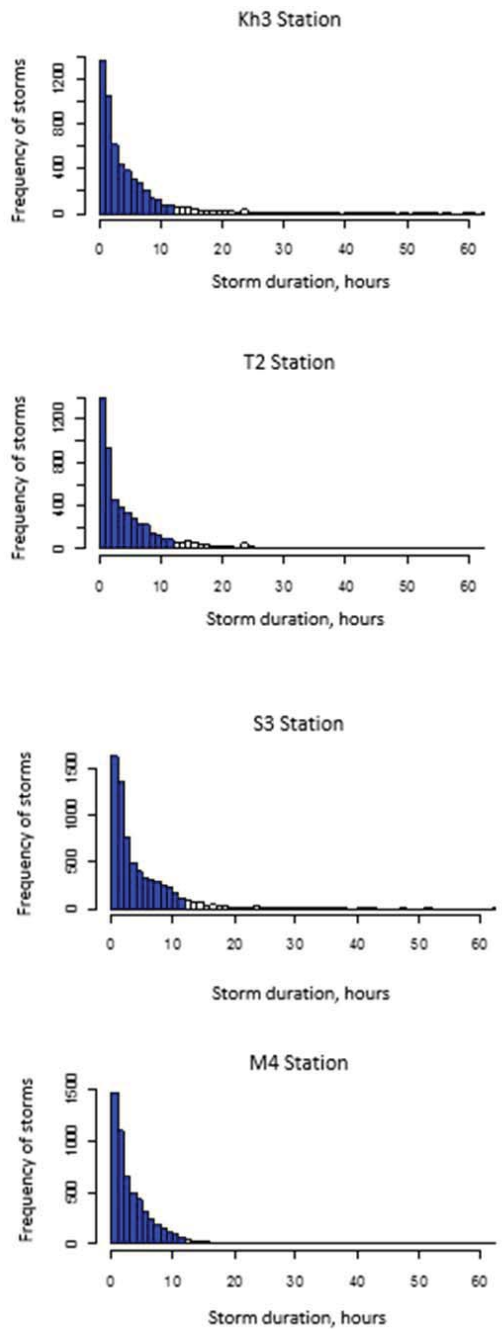

Fig. 2 Histograms of storm durations for four rainfall stations

TABLE I
GENERAL CHARACTERISTICS OF STORM EVENTS FOR 45 RAINFALL STATIONS

Station	Frequency	Annual total storm duration, hour	Mean storm duration, hour	Annual total storm amount, mm	Mean storm amount, mm
Kh1	159.94	866.59	6.26	2184.07	14.34
Kh2	173.10	955.59	5.57	2325.86	13.46
Kh3	150.46	819.97	5.56	2005.88	13.42
Kh4	153.17	919.84	6.38	1805.54	11.90
PP2	158.75	860.06	5.47	2617.33	16.70
PP3	186.56	897.76	4.82	2742.45	14.80
Pk1	192.52	1012.50	5.91	2546.17	13.63
Pk2	165.76	848.82	5.73	2183.26	13.39
Pk3	183.73	991.52	6.04	2643.82	14.78
Pk4	142.29	666.94	4.73	1528.88	10.81
Pk5	205.53	927.11	4.67	3150.04	15.42
S1	154.06	734.27	4.89	1738.97	11.35
S2	156.51	729.85	4.79	1864.69	11.95
S3	178.99	931.83	5.27	2468.10	13.86
S4	207.59	1115.54	5.53	2683.01	13.00
S5	185.36	939.71	5.37	2544.85	13.89
W1	194.72	791.35	4.10	2574.23	13.20
W2	191.80	832.98	4.40	2569.31	13.60
W3	207.79	850.68	4.08	2418.92	11.61
W4	195.00	819.85	4.25	2428.23	12.50
W5	198.98	841.59	4.27	2415.17	12.19
N3	163.46	803.32	5.15	2008.27	12.38
N4	139.17	704.01	5.62	1442.90	10.51
N5	141.60	684.63	5.03	1388.99	9.92
N1	69.04	2251.36	36.74	1681.60	26.31
N2	149.44	833.58	6.51	1685.21	11.61
J1	168.23	1009.87	6.74	2409.75	14.66
J2	169.35	830.81	5.01	2052.04	12.25
J4	187.59	1056.04	5.64	2800.27	14.94
J5	152.47	756.56	4.98	1791.35	11.74
J6	161.98	1041.61	6.52	2838.89	17.70
M4	150.42	678.57	4.58	1691.12	11.36
Pg1	149.82	706.35	4.73	1846.14	12.38
Pg2	160.43	707.74	4.33	1841.54	11.20
Pg3	149.29	817.87	5.57	2348.11	15.92
Pg4	157.57	724.91	4.96	1769.25	11.37
T1	152.60	1095.06	7.23	2712.64	17.88
T2	147.90	1027.76	7.09	2419.91	16.29
T3	180.87	1384.74	7.89	3365.83	18.74
T5	177.87	1412.15	8.10	3483.23	19.57
T6	176.04	1260.97	7.20	3337.81	18.80
Kn1	184.57	939.38	5.16	2381.01	12.91
Kn2	181.93	1030.26	5.93	2585.35	14.51
Kn3	172.34	872.86	5.05	2131.73	12.21
Kn4	179.41	1038.16	6.30	2497.50	14.16

Based on the percentages of storm durations for all storms at the 45 rainfall stations being considered, the storm events in Peninsular Malaysia can be divided into four categories; short, intermediate, long and very long storms. Short storms are storm events with storm duration less than or equal to one hour. Intermediate storms are storms which lasted for two to four hours while long storms are those with storm duration between five and twelve hours. Each category comprises of about 30% of storms while the other 10% is made of very long storms which lasted for more than twelve hours each.

Peninsular Malaysia experiences, on average, about 2309.98 mm of storm amount annually with an average of 13.98 mm for each storm event. The average annual total storm amount for the east of Peninsular Malaysia is larger than other regions with a yearly amount of about 2639.31 mm. Meanwhile, the average annual total storm amount for the north, west and south of Peninsular Malaysia are 2280.19 mm, 2255.43 mm and 2063.79 mm respectively. Similarly, the eastern region recorded the highest mean amount of rain for each storm event with 15.67 mm per storm event. Hence, rainfall stations in the east, not only receive more rain in a year but also experience more rain during each storm event. Although the rainfall stations at the south of Peninsular Malaysia recorded the smallest value for the average annual total storm amount but the mean storm amount is the smallest for storms in the western region with an average of 12.67 mm for each storm. This means that, in general, storms in the west brings less amount of rain per storm event but since the frequency of storm events is higher, hence the value of total annual storm amount is bigger compared to the rainfall stations in the south.

The frequency of storms, the mean and annual total of storm duration as well as the mean and annual total of storm amount are given in Table I. Based on Table I, there exist an outlier station which is station N1 where it is located near the reserved forest of Johol. This outlier is not detected through analysis using moving windows approach. The contour plots for storm frequency, total and mean storm duration as well as total and mean storm amount are shown in Fig. 3.

III. DAILY PROBABILITY MODEL OF STORM EVENTS

In this study, a model is built as a basis to investigate the temporal distribution and to represent the probability of a certain type of storm to happen each day in a year. The model is build based on the model used by [21]. The probability of a storm event occurring for each single day in a year is a binary data with a seasonal cycle. If we denote the occurrence of a certain type of storm under consideration on the i^{th} day of the t^{th} year as $z_t(i)$, then $z_t(i)$ follows a Bernoulli distribution with the occurrence of such storm is equal to 1 ($z_t(i) = 1$ if such storm occurs) and the non-occurrence of that particular storm is equal to 0 ($z_t(i) = 0$ if no such storm occurs).

If $z_t(i)$ is assumed to be a random variable that is independent of time and space, then $z_t(i)$ can be modelled by applying logistic regression, from the generalized linear model, on the first Fourier harmonic series. The Fourier series is used to represent the seasonality and annual cycle of storm events being considered. Specifically, the first Fourier harmonic is chosen to reduce the number of parameters that needed to be estimated in the model. Hence this model can be written as:

$$z_t(i) \sim \text{Bernoulli}(p_t(i)), \tag{1}$$

$$p_t(i) = \text{logit}^{-1}(\beta_{1t} + \beta_{2t}\sin(\omega i) + \beta_{3t}\cos(\omega i)) \tag{2}$$

with $\text{logit}^{-1}(x) = \frac{\exp(x)}{1+\exp(x)}$, $\omega = \frac{2\pi}{\gamma}$ rad/day and the angle of $1 \text{ rad} = \frac{180^{\circ}}{\pi}$. The value of γ depends on the annual cycle of the type of storm being considered which is usually a unimodal or a bimodal cycle [21]. For this study, a unimodal cycle uses γ = 366 while a bimodal cycle uses γ = 183. In this model, i = 1 is the first day for year t which is the 1st of January for that year. Meanwhile, by taking account 29th of February during leap year, i = 366 is the last day representing the 31st of December for year t.

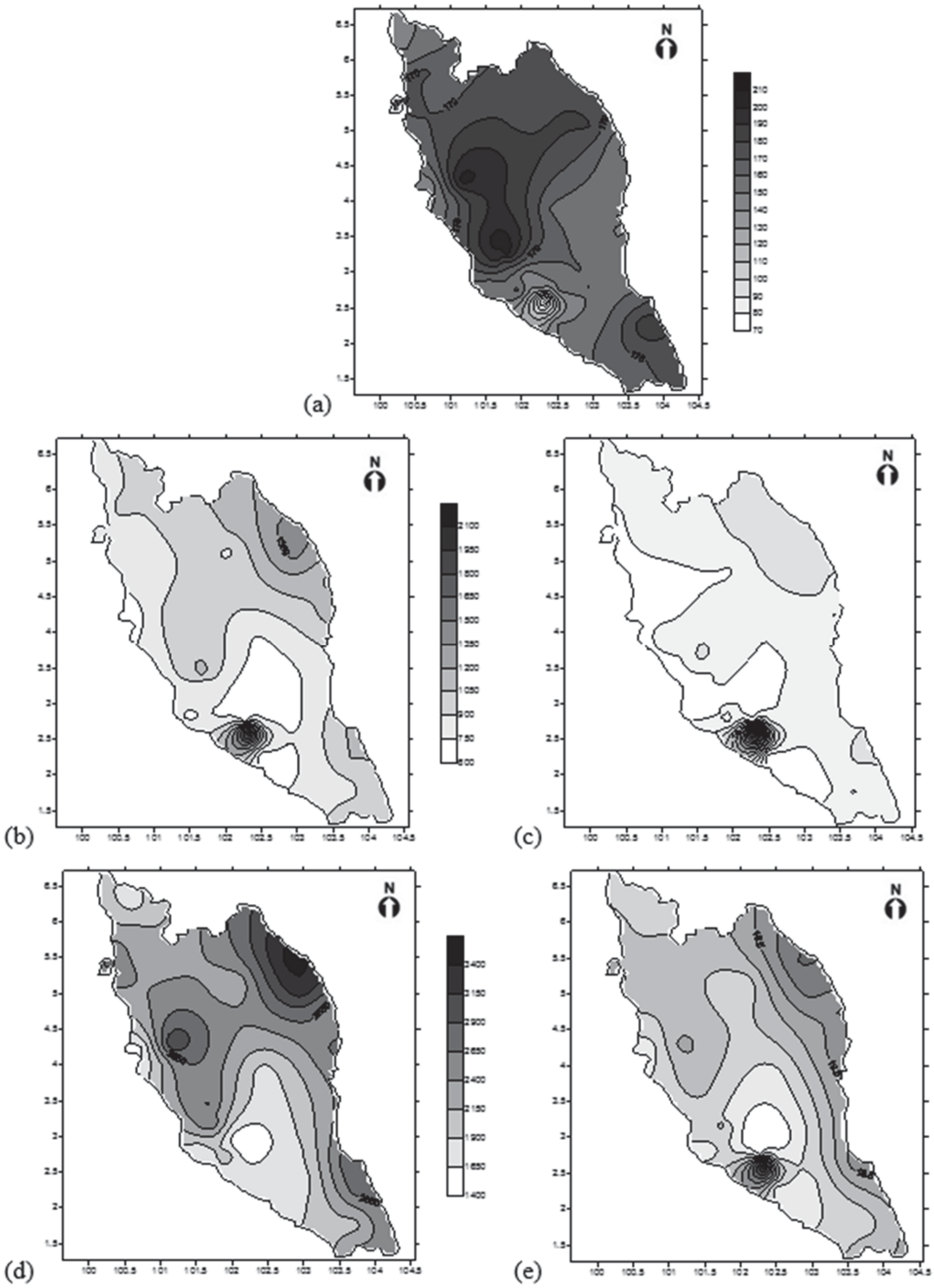

Fig. 3 Contour plots for (a) frequency of storm events, (b) annual total storm duration (hours), (c) mean storm duration (hours), (d) annual total storm amount (mm), and (e) mean storm amount (mm)

For validation purposes, the coefficient of variation of root mean squared error (CV_{RMSE}) is used. The value of CV_{RMSE} is calculated to measure the difference between observed daily probabilities of storm event being studied which are obtained from the data and the predicted probabilities of that particular type of storm event found from the model. The observed daily probabilities are the ratios of the number of times such storm event exist for the i^{th} day against the number of years in the data set. Thus, the value of CV_{RMSE} is calculated as:

$$CV_{RMSE} = \frac{\sqrt{\frac{1}{366}\sum_{i=1}^{366}\left(\hat{p}(i)\text{-}p(i)\right)^2}}{\bar{p}} \tag{3}$$

with $\hat{p}(i)$ is the probability that the particular type of storm event occuring on the i^{th} day of a year based on the model while $p(i)$ is the probability found from the observed storm event in the data set. The average of $p(i)$ which is denoted as $\bar{p}$ is computed as

$$\bar{p}= \frac{1}{366}\sum_{i=1}^{366} p(i).$$

The value of CV_{RMSE} is also used to determine whether the annual cycle for the storm event being studied is a unimodal or a bimodal cycle. The annual cycle which gives the smaller value of CV_{RMSE} is deemed to be able to represent the seasonal cycle of the storm event better.

IV. Application of Model on Storm Events in Peninsular Malaysia

The daily probability for four categories of storms; short, intermediate, long and very long storms; are modelled and the annual cycle are determined for all 45 rainfall stations in Peninsular Malaysia. The average values of CV_{RMSE} for the four categories of storms are 0.5621, 0.4704, 0.6061 and 1.7692 respectively. These show that the root mean squared errors (RMSE) found from the model are small with the errors for short, intermediate and long storms are smaller than the mean observed probabilities found from the data. The values of CV_{RMSE} are larger for very long storms since the data for this type of storm are small; comprises of about 10% of the whole data set. Hence, longer data set are needed to model it more accurately. However, based on the plots of the observed daily probability from the data and the predicted daily probability found from the model, it can be seen that the model is able to capture the cyclic pattern of all four categories of storms in Peninsular Malaysia. Examples of the plots for selected rainfall stations are shown in Fig. 4.

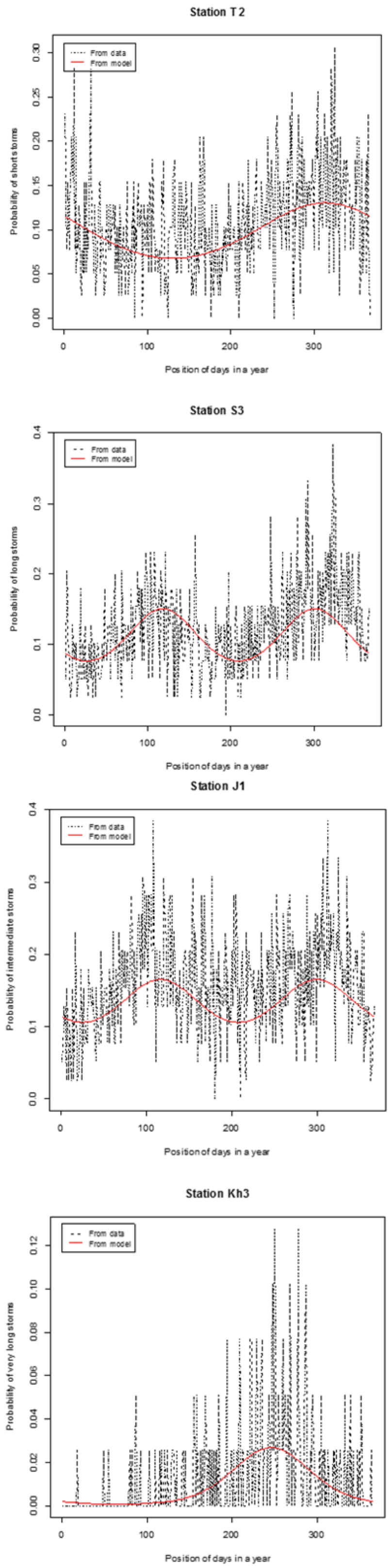

Fig. 4 Plots of daily probability of storm events observed from the data and predicted from the model

From the comparison of the unimodal and bimodal cycle at each rainfall stations, it can be observed that there is a difference between the annual cycles of storm events at different regions in Peninsular Malaysia. These cycles also differ between types of storms at each rainfall station. The annual cycle for each category of storms at all 45 rainfall stations are shown in Fig. 5. Fig. 5 shows that short and intermediate storms at most rainfall stations located at the north, west and south of Peninsular Malaysia have a bimodal cycle. This may be due to the two inter-monsoon seasons which occur around April and October each year. During these inter-monsoon seasons, Peninsular Malaysia received a lot of convective storms which are storms with small values of storm duration like those of short and intermediate storms.

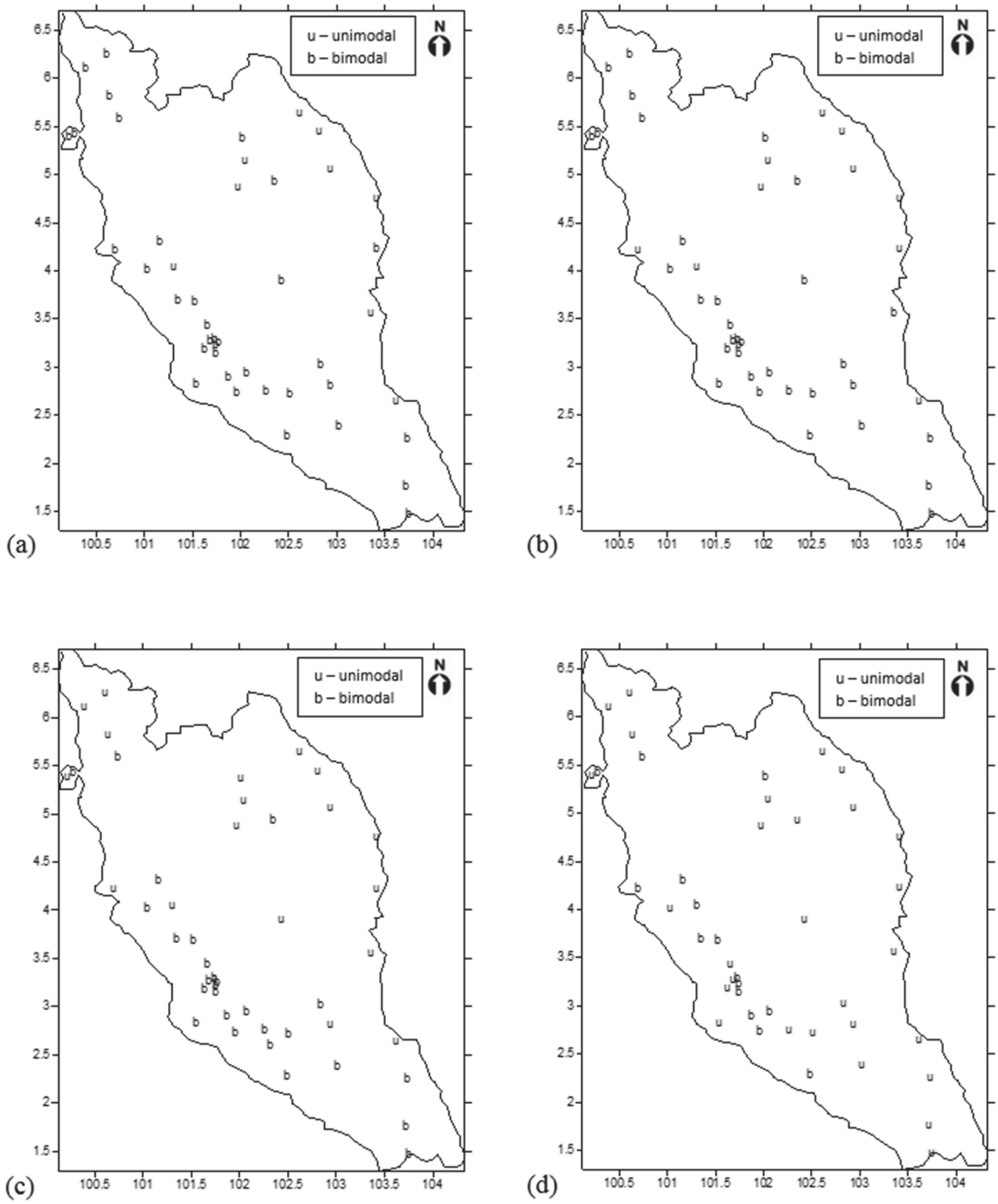

Fig. 5 Annual cycles of (a) short, (b) intermediate, (c) long, and (d) very long storms

Without the influence of convective storms during the inter-monsoon seasons, storms at the northern part of Peninsular Malaysia experiences a unimodal cycle such as those seen for long and very long storms. Very long storms for the south region also show a unimodal cycle. Storm events at most rainfall stations in the western region have bimodal cycles for all types of storm categories. This is because the west of Peninsular Malaysia is shielded from any influence of the Northwest and Southeast monsoon season by the Titiwangsa Range and Sumatera Island respectively.

Mostly storms in the east of Peninsular Malaysia have unimodal cycles. This is because the eastern region is very much influenced by the Northwest monsoon season which occurs from November to March each year. Hence, the probability of storm occurring during this monsoon season is higher compared to other times of the year. Consequently, this results in the unimodal cycle of daily probability of storm events with the cycle starting near the end of each year and lasts until the beginning of the following year. This can be clearly seen from the annual cycles of long and very long storms. The influence of the two inter-monsoon seasons which affect other regions could also be seen evident for short and intermediate storms at the east of Peninsular Malaysia. For short and intermediate storms, the unimodal cycle for daily probability of such storm events starts earlier and ends later compared to long and very long storms.

From the results obtained through applying this model on all rainfall stations, it can be seen that there exist four distinct regions of storms in Peninsular Malaysia due to the unique geographical landscape and monsoon seasons. By performing cluster analysis based on the annual cycles of all four categories of storms, the four regions found is shown in Fig. 6.

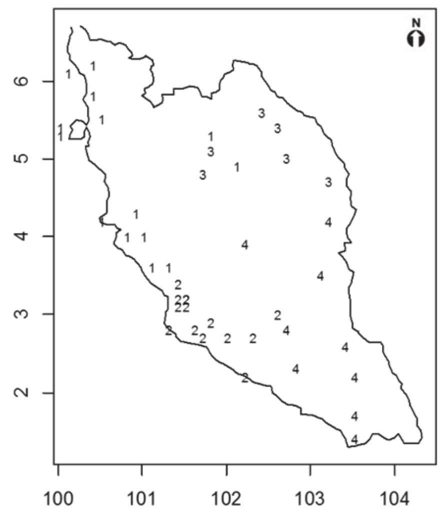

Fig. 6 Regions of storms based on the annual cycles of all the four categories of storm events in Peninsular Malaysia

V. Conclusion

Rainfall analysis using SEA provides a method to analyze rainfall characteristics as storms where storm duration and storm amount can be taken as random variables. Specifically modelling the seasonality or daily occurrence of storms helps in understanding the interannual variability of rainfalls which affects various other fields such as agriculture and hydrology. This study models the daily probability of storms for four different categories of storm events using the generalized linear model on the first Fourier harmonic series.

The results obtained show that the annual cycle of storm events in Peninsular Malaysia is very much dependent on the location of rainfall stations where these locations are affected by the unique geographical landscape of Peninsular Malaysia and the monsoon seasons as well as the two inter-monsoon seasons. Four distinct regions are seen based on the different types of storms and their annual cycles.

Acknowledgment

The authors would like to express utmost appreciations to the Department of Irrigation and Drainage Malaysia for the use of hourly rainfall data in this study. The authors are also very grateful towards the Ministry of Education (MOE) for the allocation of the research grant (FRGS/1/2014/SG04/UKM/03/2) and to the Universiti Kebangsaan Malaysia (UKM) for the equipment and facilities utilized during the research process.

References

[1] S. Vanderberghe, N. E. C. Verhoest, C. Onof, and B. De Baets, "A comparative copula-based bivariate frequency analysis of observed and simulated storm events: A case study on Bartlett-Lewis modeled rainfall," *Water Resources Research*, vol. 47(7), W07529, 2011.

[2] P. J. Restrepo-Pasada, and P. S. Eagleson, "Identification of independent rainstorms," *Journal of Hydrology*, vol. 55(1-4), pp. 303-319, 1982.

[3] S. C. Kao, and R. S. Govindaraju, "Probabilistic structure of storm surface runoff considering the dependence between average intensity and storm duration of rainfall events," *Water Resources Research*, vol. 43(6), pp. W06410, 2007.

[4] S. C. Kao, and R. S. Govindaraju, "Trivariate statistical analysis of extreme rainfall events via the Plackett family of copulas," *Water Resources Research*, vol. 44(2), pp. W02415, 2008.

[5] B. Palynchuk, and Y. Guo, "Threshold analysis of rainstorm depth and duration statistics at Toronto, Canada," *Journal of Hydrology*, vol. 348(3-4), pp. 535-545, 2008.

[6] T. M. Heneker, M. F. Lambert, and G. Kuczera, "A point rainfall model for risk-based design," *Journal of Hydrology*, vol. 247(1), pp. 54-71, 2001.

[7] J. Joo, J. Lee, J. H. Kim, H. Jun, and D. Jo, "Inter-event time definition setting procedure for urban drainage systems," *Water*, vol. 6(1), pp. 45-58, 2013.

[8] N. M. Ariff, A. A. Jemain, K. Ibrahim, and W. Z. Wan Zin, "IDF Relationships using Bivariate Copula for Storm Events in Peninsular Malaysia," *Journal of Hydrology*, vol. 470-471, pp. 158-171, 2012.

[9] S. M. Deni, J. Suhaila, W. Z. Wan Zin, and A. A. Jemain, "Spatial trends of dry spells over Peninsular Malaysia during monsoon seasons," *Theoretical and Applied Climatology*, vol. 99(3), pp. 357-371, 2010.

[10] S. M. Deni, J. Suhaila, W. Z. W. zin, and A. A. Jemain, "Trends of wet spells over Peninsular Malaysia during monsoon seasons," *Sains Malaysiana*, vol.38(2), pp.133-142, 2009.

[11] J. Suhaila, S. M. Deni, W. Z. Wan Zin, and A. A. Jemain, "Spatial patterns and trends of daily rainfall regime in Peninsular Malaysia during the southwest and northeast monsoons: 1975-2004," *Meteorology and Atmospheric Physics*, pp. 1-18, 2010.

[12] J. Suhaila, A. A. Jemain, M. F. Hamdan, and W. Z. W. Zin, "Comparing rainfall patterns between regions in Peninsular Malaysia via a functional data analysis technique," *Journal of Hydrology*, vol. 411, pp. 197-206, 2011.

[13] J. Suhaila, and A. A. Jemain, "A comparison of the rainfall patterns between stations on the East and the West coasts of Peninsular Malaysia using the smoothing model of rainfall amounts," *Meteorological Applications*, vol. 16(3), pp.391-401, 2009.
[14] M. N. Desa, "Characterisation of urban rainfall in Kuala Lumpur, Malaysia (Ph.D. Thesis)," Lund University, Sweden, 1997.
[15] M. M. N. Desa, and J. Niemczynowicz, "Spatial variability of rainfall in Kuala Lumpur, Malaysia: long and short term characteristics," *Hydrological Sciences Journal*, vol. 41(3), pp. 345-362, 1996.
[16] M. N. Desa, and J. Niemczynowicz, "Temporal and spatial characteristics of rainfall in Kuala Lumpur, Malaysia," *Atmospheric Research*, vol. 42(1), pp. 263-277, 1996.
[17] F. Yusof, and I. L. Kane, "Volatility modeling of rainfall time series," *Theoretical and Applied Climatology*, vol. 113(1-2), pp. 247-258, 2012.
[18] J. Samson, P. Thomson, and T. Carey-Smith, "Stochastic seasonality of rainfall in New Zealand," *Journal of Geophysical Research: Atmospheres*, vol. 118(10), pp.3944-3955, 2013.
[19] M. G. Slocum, W. J. Platt, B. Beckage, S. L. Orzell, and W. Taylor, "Accurate quantification of seasonal rainfall and associated climate-wildfire relationships," *Journal of Applied Meteorology and Climatology*, vol. 49(12), pp. 2559-2573, 2010.
[20] T. Carey-Smith, J. Samson, and P. Thomson, "A hidden seasonal switching model for multisite daily rainfall," *Water Resources Research*, vol. 50(1), pp. 257-272, 2014.
[21] C. H. R. Lima, and U. Lall, "Hierarchical Bayesian modeling of multisite daily rainfall occurrence: Rainy season onset, peak, and end," *Water Resources Research*, vol.45, pp. W07422, 2009.
[22] J. Samson, and P. Thomson, "A hidden seasonal switching model for high-resolution breakpoint rainfall data," *Water Resources Research*, vol. 46(8), pp. W08510, 2010.
[23] Z. M. Daud, A. H. M. Kassim, M. N. M. Des, and V. T. V. Nguyen, "Statistical analysis of at-site extreme rainfall processes in Peninsular Malaysia," *IAHS-AISH Publication*, vol. 274, pp. 61-68, 2002.
[24] C. L. Wong, R. Venneker, S. Uhlenbrook, A. B. M. Jamil, and Y. Zhou, "Variability of rainfall in Peninsular Malaysia," *Hydrol. Earth Syst. Sci. Discuss*, vol. 6, pp. 5471-5503, 2009.