Research Journal of Applied Sciences, Engineering and Technology 13(4): 283-294, 2016
DOI:10.19026/rjaset.13.2944
ISSN: 2040-7459; e-ISSN: 2040-7467

Submitted: July 13, 2015 Accepted: August 30, 2015 Published: August 15, 2016

Research Article

Modeling of Annual Maximum Storm Intensity with Bayesian Markov Chain Monte Carlo (MCMC) and L-moment

[1]Rado Yendra, [1]Ari Pani Desvina, [1]Rahmadeni, [2]Kuseiri, [3]Abdul Aziz Jemain and [4]Ahmad Fudholi
[1]Department of Mathematics, Faculty of Science and Technology, Universitas Islam Sultan Syarif Kasim (UIN Suska), 28293 Pekanbaru, Riau, Indonesia
[2]Department of Mathematics and Science Education, Faculty of Education and Teacher Training, Islamic State University of Sunan Ampel, Surabaya, Indonesia
[3]School of Mathematical Sciences, Faculty of Science and Technology,
[4]Solar Energy Research Institute, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia

Abstract: This study presents the best fitting distribution to describe the siries MSI based on hourly rainfall form 1970 to 2008 for three rain gauge stations in Peninsular Malaysia namely Bertam, Dungun and Pekan. Two three-parameter extreme value distributions which are considered are Generalized Extreme Value (GEV) and Generalized Logistic (GL). The parameters of these distributions are determined using the Bayesian MCMC with non-informative prior distribution and L-moments (LMOM) method. The Goodness-Of-Fit (GOF) between empirical data and theoretical distributions are then evaluated for each stations. The result show that the majority of the stations are found that the L-moment method can give the best modelling for MSI, specified for GEV distribution. Based on the model that has been identified, we can reasonably predict the risks associated the MSI for various return periods.

Keywords: AMSE, Bayesian MCMC, MSA, MSD, MSI

INTRODUCTION

Statistical modeling of extreme event is important in various disciplines including hydrology, engineering and environmental science. For environmental processes extreme value theory can be used to estimate the probabilities of extreme levels of the processes. For some processes, such as sea-level and wind speed, this information can help in the design of structures such as sea walls, bridges and buildings. For other processes, such as rainfall and pollution, the information can be used to assess danger due to extreme levels of the process. Extreme value theory can be used in finance to, for example, assess the risks of large insurance claims or predict the probability of rare events.

Extreme rainfall event is often associated with climate change, which may be followed by siries of natural disasters such as flash floods and landslides. Considering this phenomenon, the analysis of extreme rainfall data can be utilized for decision makers to set-up measures for reducing or preventing the impact of disasters. In Malaysia, extreme analysis on rainfall data has been explored for all sorts of purposes such as tracing patterns and trends of daily rainfall during monsoon seasons (Suhaila *et al*., 2010a, 2010b), detecting recent changes in extreme rainfall events (Zin *et al*., 2010) and fitting probability distributions to annual maximum rainfalls by implementing various methods (Zin *et al*., 2009; Zawiah *et al*., 2009; Zin *et al*., 2010).

Previous literatures provide a few methods of viewing Storm Event Analysis (SEA) in their analysis, among them (Eagleson, 1972; Adams *et al*., 1986; Guo and Adams, 1998a, 1998b; Adams and Papa, 2000; Rivera *et al*., 2005). The statistical characterization of Maximum Storm Intensity (MSI) will be analysis involves fitting two extreme values distributions which are considered are Generalized Extreme Value (GEV) and Generalized Logistic (GL). The estimation parameters of these distributions is determined using Bayesian with non-informative prior and L-moment.

Bayesian inference is having a fundamental impact on virtually every statistical methodology. The Bayesian analysis has enormous potential for the various research fields. Especially, there are important literatures which dealt with the Bayesian approach in

Corresponding Author: Ahmad Fudholi, Solar Energy Research Institute, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia, Tel.: +60-17-3328637; Fax: +60-3-8925-4519

the fields of water resources engineering (Coles and Powell, 1996; Kuczera and Parent, 1998). The most important step in the processes of the Bayesian method is the construction of the prior distribution. The prior distribution represents the information about an uncertain parameter that is combined with the probability distribution of new data to yield the posterior distribution, which in turn is used for the future inferences and decisions involving parameter. The decision to use a particular prior distribution should be based on any available knowledge about the parameters. Estimated optimal values of the parameters from previous studies can be helpful for establishing the prior distribution and constraints on the model parameters. There are two types of prior distributions, 'Data-based prior distribution' and 'Non data-based prior distribution'. When the prior distribution is derived through the objective analysis using data, it is called as the data-based or the informative prior distribution. Also, when it is derived from subjective judgments or theoretical considerations, it is called as the non-data-based prior distribution. Especially, the non-informative prior distribution is a special case of the non-data-based prior distribution. When the non-informative prior distribution is used, the posterior distribution only reflects the information in the sample. One of the arguments to Bayesian statistics is the potential effect by the non-informative prior distribution. Jeffreys (1961) suggested that efforts for the elicitation of the informative prior distribution should be performed. While theoretical background to apply the non-informative prior distribution is plentiful (Bernado and Smith, 1994; Gelman *et al*., 1995; Carlin and Louis, 1996), there appeared to be a little literature in which the analysis is via the informative prior distribution in spite of the importance of the application with the informative prior distribution.

Table 1: Main characteristics of the rain stations, SD standard deviation

No.	Station name	Mean	S.D.	Coef of variation	Kurtosis	Skewness
1	Bertam	24.41	7.5	0.31	1.13	4.27
2	Dungun	20.46	7.87	0.38	1.39	5.39
3	Pekan	23.64	8.28	0.35	1.57	6.82

S.D.: Standard deviation

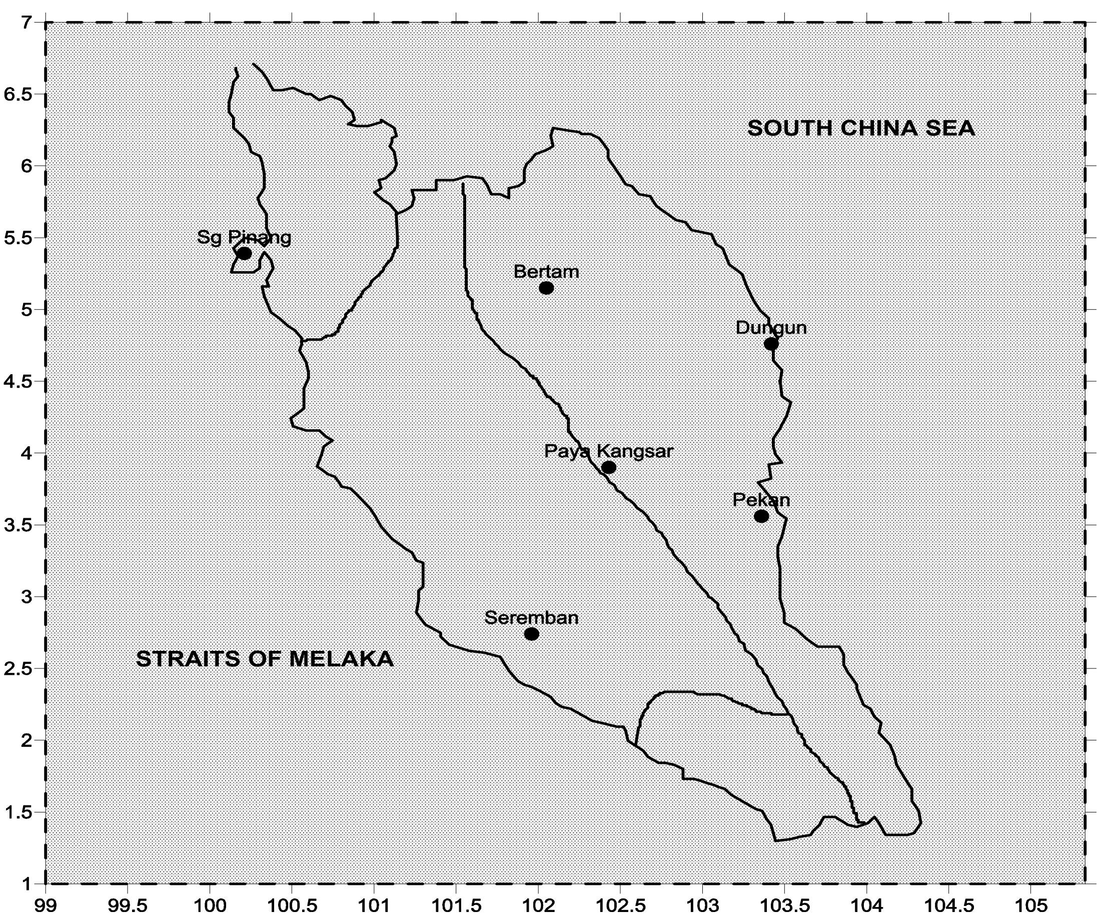


Fig. 1: Location of rain stations used in this study

Few are found information about external measures of Annual Maximum Storm Event (AMSE) in literature to the best of the authors' knowledge. The main objective of this study is fitting distribution to describe the siries MSI based on hourly rainfall form 1970 to 2008 for three rain gauge stations in Peninsular Malaysia namely Bertam, Dungun and Pekan.

DATA AND DEFINITION OF STORM

The data consisting of hourly rainfall data from 3 rain gauge stations in Peninsular Malaysia from 1970 to 2008 have been obtained from the Drainage and Irrigation Department. Data from these stations presented in Table 1. The locations of these six stations are shown in Fig. 1.

The definition of storm-event depends greatly on the interevent time definition. The Inter-Event Time Definition (IETD) is defined as the minimum duration of dry period between two consecutive storm events. Hence, the dry duration between two individual storm events must at least be equal to the IETD value. If not, they would not be considered as two different events but parts of the same storm. The IETD value is chosen such that the serial correlation between the two different storms is minimized (Restrepo-Posada and Eagleson, 1982). For small urban catchments, the IETD is usually taken as 6 h because the time concentration of rainfall which is less than 6 h would make the runoff response of successive storms to appear independent (Palynchuk and Guo, 2008). Storm depth is defined as the accumulated rainfall which begins and ends with at least one wet hour and either contains dry periods with less than 6 h or none at all. Storm duration is defined as the time interval for a storm event and storm intensity is the ratio of storm depth to storm duration. The information extracted from the rainfall data is the annual MSA, MSI and MSD. This information and definition of storm can be explained form Fig. 2. The annual MSA, MSI and MSD and Number of Storm (NS) for two rain gauge stations (Dungun and Bertam) are provided in Table 2.

let x_{ij} is jth rainfall (mm) on ith storm and n_i is duration (hour) on ith storm, for each storm the MSA, MSI and MSD was obtained from the hourly data as follows:

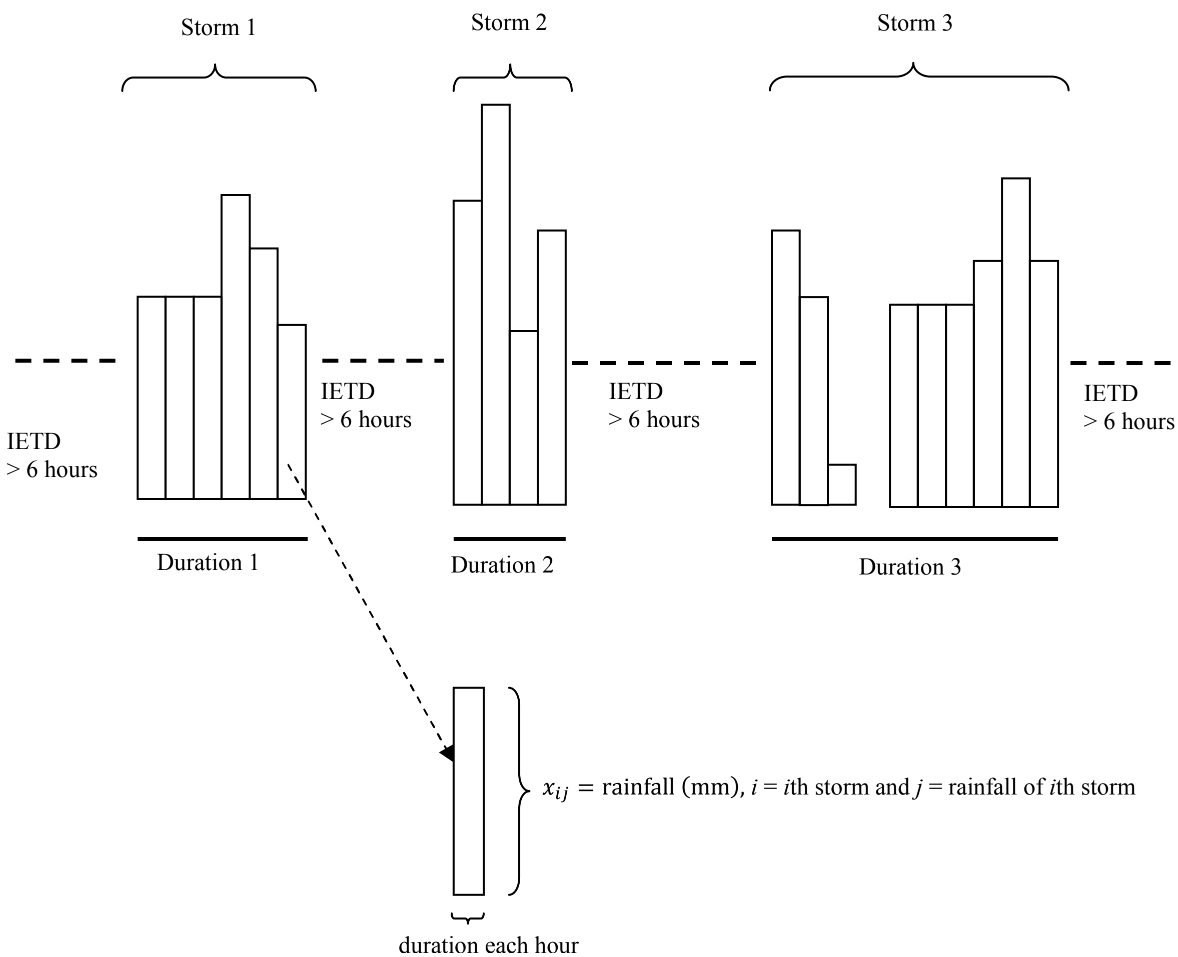


Fig. 2: Definition of storm

Table 2: Annual number of storm (NS) and annual MSA, MSI and MSD for station Dungun and Bertam

	Dungun				Bertam			
Years	NS	MSA	MSI	MSD	NS	MSA	MSI	MSD
1970	127	293.40	12.37	63	142	76.90	19.30	78
1971	158	329.00	17.43	124	196	193.70	17.60	56
1972	127	281.90	14.46	99	206	96.80	22.60	48
1973	164	456.90	17.20	116	321	60.10	17.70	41
1974	166	439.50	22.40	98	187	29.82	12.41	28
1975	178	396.50	18.00	77	185	109.50	29.17	23
1976	147	366.80	34.50	81	189	106.50	18.17	34
1977	150	99.00	22.50	28	212	101.60	23.50	27
1978	154	124.90	18.20	49	242	107.10	20.89	47
1979	152	165.10	46.50	58	193	48.00	38.50	27
1980	165	279.50	18.00	72	197	93.10	36.50	132
1981	135	124.50	24.90	72	173	80.40	21.49	51
1982	100	445.80	10.00	163	165	244.80	48.50	133
1983	146	414.00	26.50	72	187	54.50	23.00	24
1984	165	241.90	19.50	204	208	224.90	21.25	53
1985	168	215.90	14.60	59	183	86.50	18.38	24
1986	140	916.00	17.00	127	168	115.50	18.25	36
1987	175	225.90	41.50	51	167	138.10	30.00	50
1988	187	887.80	29.50	82	197	224.80	19.12	50
1989	127	192.00	29.50	128	210	57.50	21.50	34
1990	148	192.50	21.38	68	110	90.30	15.76	49
1991	150	327.20	20.00	71	210	143.50	26.50	36
1992	158	265.60	19.50	101	167	147.10	15.50	76
1993	166	196.10	13.50	49	194	283.40	38.50	90
1994	142	348.40	19.50	125	203	98.40	16.88	37
1995	127	219.20	12.77	70	192	125.00	23.17	44
1996	148	301.00	11.60	87	199	159.50	34.75	46
1997	142	126.00	20.00	48	158	99.00	24.50	35
1998	180	109.00	22.47	43	177	153.50	24.38	79
1999	245	192.25	11.92	55	223	111.40	25.97	48
2000	198	325.50	20.42	61	161	130.50	28.25	27
2001	156	431.60	15.62	88	125	116.50	29.12	16
2002	99	125.70	14.30	24	133	78.00	20.00	22
2003	160	244.50	22.23	17	166	105.50	28.50	47
2004	145	261.00	22.50	52	161	259.00	19.62	63
2005	156	231.00	17.57	110	200	153.00	28.00	57
2006	148	221.00	25.00	79	161	165.50	32.00	80
2007	143	295.80	24.37	96	209	205.00	20.04	91
2008	90	106.40	8.77	60	122	89.50	22.75	31

MSD = Maximum $(n_i), i = 1,2,3$

MSA = Maximum $\left(\sum_{j=1}^{n_i} x_{ij}\right), i = 1,2,3$

MSI = Maximum $\left(\frac{\sum_{j=1}^{n_i} x_{ij}}{n_i}\right), i = 1,2,3$

METHODOLOGY

Probability distribution: The most common analysis of extreme hydrological events involves the use of annual maximum or annual extreme. When constructing the MSI series for each year in the record is selected; hence, the series obtained would have a length equal to the number of years. Many works that apply the annual maximum series usually involve fitting of a probability model to the rainfall data. Thereafter, several researchers have provided useful applications of annual maximum distributions to rainfall data obtained from different regions of the world.

Two probability distributions associated with modeling extreme events, GEV and GL are considered in this study. The probability density function, probability function and quantile function for each distribution that we consider are as given in Table 3, where x denote the observed values of the random variable representing the event of interest, α is the scale parameter, ε is the location parameter and κ is the shape parameter.

In order to fit a particular theoretical distribution to the observed distribution of AMSE, parameters are estimated using the Bayesian MCMC and LMOM method.

Bayesian MCMC with non-informative prior distribution: This section introduces the idea of Bayesian MCMC using non-informative priors. Suppose that prior beliefs about θ can be formulated and expressed by a probability density function π(θ) with no reference to the data. The likelihood for θ is $L(\theta \mid x)$. The prior information and the likelihood can be combined using Bayes theory to produce a posterior distribution for θ as follows:

Table 3: List of distributions used in this study

Distribution	Probability density function, $f(x)$	Cumulative distribution, $F(x)$	Quantile function, Q (F)
GEV	$f(x)=\alpha^{-1}exp\{-(1-\kappa)y-exp(-y)\}$ $with\ y=\begin{cases}-\kappa^{-1}log\{1-\kappa(x-\xi)/\alpha\}, & \kappa\neq 0\\ (x-\xi)/\alpha, & \kappa=0\end{cases}$ $-\infty<x<\xi+\alpha/\kappa\ for\ \kappa>0$ $-\infty<x<\infty\ for\ \kappa=0$ $\xi+\alpha/\kappa\leq x<\infty\ for\ \kappa<0.$	$F(x)=$ $\begin{cases}exp\left(-\left(1-\frac{\kappa}{\alpha}(x-\varepsilon)\right)^{\frac{1}{\kappa}}\right), if\ \kappa\neq 0\\ exp\left(-exp\left(-\frac{1}{\alpha}(x-\varepsilon)\right)\right), if\ \kappa=0\end{cases}$	$Q(F)=\varepsilon+\frac{1}{\kappa}(1-$ $-lnF\kappa$
GL	$f(x)=\frac{\alpha^{-1}exp\{-(1-\kappa)y\}}{\left(1+exp(-y)\right)^2}$ $with$ $y=\begin{cases}-\kappa^{-1}log\{1-\kappa(x-\xi)/\alpha\}, & \kappa\neq 0\\ (x-\xi)/\alpha, & \kappa=0\end{cases}$ $-\infty<x<\xi+\alpha/\kappa\ \ for\ \kappa>0$ $-\infty<x<\infty\ for\ \kappa=0$ $\xi+\alpha/\kappa\leq x<\infty\ for\ \kappa<0$	$F(x)=\left(1+\left(1-\kappa\left(\frac{x-\varepsilon}{\alpha}\right)\right)^{\frac{1}{\kappa}}\right)^{-1}$	$Q(F)=\varepsilon+\frac{\alpha}{\kappa}\left(1-\left(\frac{1-F}{F}\right)^{\kappa}\right)$

$$\pi(\theta\mid x)=\pi(\theta)L(\theta\mid x)\Big/\int\pi(\theta)L(\theta\mid x)d\theta$$

Often, the function of an extreme value analysis is to describe the extremal behavior of an observed process in order to find the probability of extreme events occurring in the future. Within the Bayesian framework, prediction is possible through the predictive distribution. Let y denote a future observation with probability density function $f(y|\theta)$, then:

$$f(y\mid x)=\int f(y\mid\theta)\pi(\theta\mid x)d\theta$$

Is the predictive distribution of y given x. So, if suitable prior distribution can be specified, there are good reasons to choose Bayesian procedures. The difficulty in computing the integral in predictive distribution makes the simulation techniques such as MCMC can be overcome to simulate realizations of the posterior distribution. The main issue in Bayesian MCMC with non-informative prior distribution is the priors are constructed by assuming there is no information available about the process apart from data. In this study the prior density was chosen to be:

$$\pi(\phi,\xi,\kappa)=\pi_{\alpha}(\phi)\pi_{\xi}(\xi)\pi_{\kappa}(\kappa)$$
$$\text{with }\ \phi=\log\alpha,\ \phi\sim N(0,100),\ \xi\sim N(0,1000),\ and\ \ \kappa\sim N(0,10),$$

The variances are chosen large enough to make the distribution almost flat, corresponding to prior ignorance and the posterior density is:

$$\pi(\phi,\xi,\kappa|y)\propto\pi(\phi,\xi,\kappa)L(\phi,\xi,\kappa|y)$$
$$\text{where } L(\phi,\xi,\kappa|y)\ is\ likelihood\ ,$$
$$y\sim GEV(\phi,\xi,\kappa)\ ,\ and\ \ y\sim GL(\phi,\xi,\kappa)$$

The full details of the algorithm are as follows:

1. Initialize the chain a $\theta^0=(\phi^0,\xi^0,\kappa^0)$ and the couter at $j = 1$
2. Put $\phi^*=\phi^{(j-1)}+\omega_{\phi},\ \omega_{\phi}\sim N(0,4)$
3. Accept $\phi^{(j)}=\phi^*$ with probability $a(\phi^{(j-1)},\phi^*)=\min\{1,A\}$ where $A=\dfrac{\pi(\phi^*|\xi^{(j-1)},\kappa^{(j-1)})}{\pi(\phi^{(j-1)}|\xi^{(j-1)},\kappa^{(j-1)})}$ and $\phi^{(j)}=\phi^{(j-1)}$ otherwise
4. Put $\xi^*=\xi^{(j-1)}+\omega_{\xi},\ \omega_{\xi}\sim N(0,0.3)$
5. Accept $\xi^{(j)}=\xi^*$ with probability $a(\xi^{(j-1)},\xi^*)=\min\{1,A\}$ where $A=\dfrac{\pi(\xi^*|\phi^{(j)},\kappa^{(j-1)})}{\pi(\xi^{(j-1)}|\phi^{(j)},\kappa^{(j-1)})}$ and $\xi^{(j)}=\xi^{(j-1)}$ otherwise
6. Put $\kappa^*=\kappa^{(j-1)}+\omega_{\kappa},\ \omega_{\kappa}\sim N(0,0.1)$
7. Accept $\kappa^{(j)}=\kappa^*$ with probability $a(\kappa^{(j-1)},\kappa^*)=\min\{1,A\}$ where $A=\dfrac{\pi(\kappa^*|\phi^{(j)},\xi^{(j)})}{\pi(\kappa^{(j-1)}|\phi^{(j)},\xi^{(j)})}$ and $\kappa^{(j)}=\kappa^{(j-1)}$ otherwise
8. Increase counter from j to $j+1$ and return to step 2

L-moment (LMOM) and Goodness-Of-Fit (GOF): The LMOM method, populared by Hosking and Wallis (1997), is widely applied in the field of applied research such as hydrology, meteorology and civil engineering for estimating parameters of a distribution. It is based on a linear combination of order statistics where the first- until the fourth-order statistics correspond to measures of location, scale, skewness and kurtosis, respectively. When compared to maximum likelihood methods and method of moments, estimators found based on LMOM are more robust, proven to have smaller mean square error and easier to compute. As described by Vogel and Fennessey (1993), LMOM should be preferred for small sample sizes due to its robust property. The *r*th LMOM, denoted as λ_r is defined as:

$$\lambda_r = \frac{1}{r}\sum_{k=0}^{r-1}(-1)^k \binom{r-1}{k} E(X_{r-k:r});\ r = 1,2$$

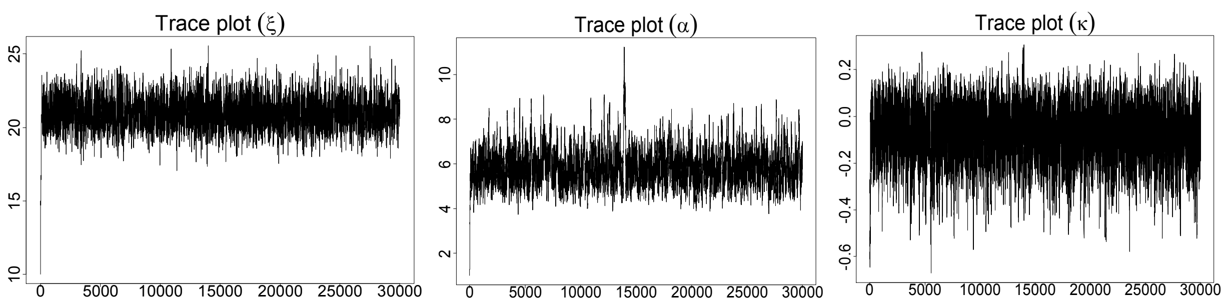
where, $X_{r-k:r}$ is the random variable variable for $(r-k)$ th order statistics.

Once the distribution of the observed values is determined for MSI series, the expected frequencies under the assumed distribution are computed for each station. The most appropriate distribution for each station is identified using results found based on several goodness-of-fit tests.

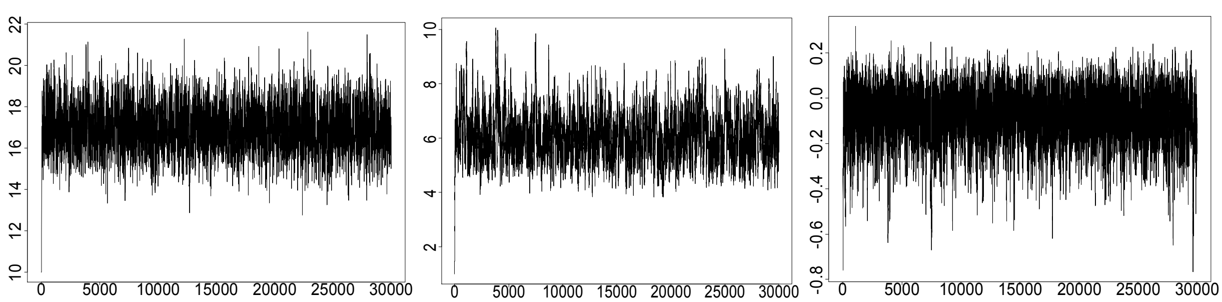
Three GOF tests considered are Relative Root Mean Square Error (RRMSE), Relative Absolute Square Error (RASE) and Probability Plot Correlation Coefficient (PPCC). The first two methods involve the assessment on the difference between the observed values and the expected values under the assumed distribution while the third method involves measuring the correlation between the ordered values and the associated expected values. The formulas for the tests are:

$$RRMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_{i:n}-\hat{Q}(F_i)}{x_{i:n}}\right)^2}\ RASE = \sum_{i=1}^{n}\left|\frac{x_{i:n}-\hat{Q}(F_i)}{x_{i:n}}\right|\ PPCC = \frac{\sum_{i=1}^{n}(x_{i:n}-\bar{x})\left(\hat{Q}(F_i)-\bar{Q}(F_i)\right)}{\sqrt{\sum_{i=1}^{n}(x_{i:n}-\bar{x})^2}\sqrt{\sum_{i=1}^{n}\left(\hat{Q}(F_i)-\bar{Q}(F_i)\right)^2}}$$

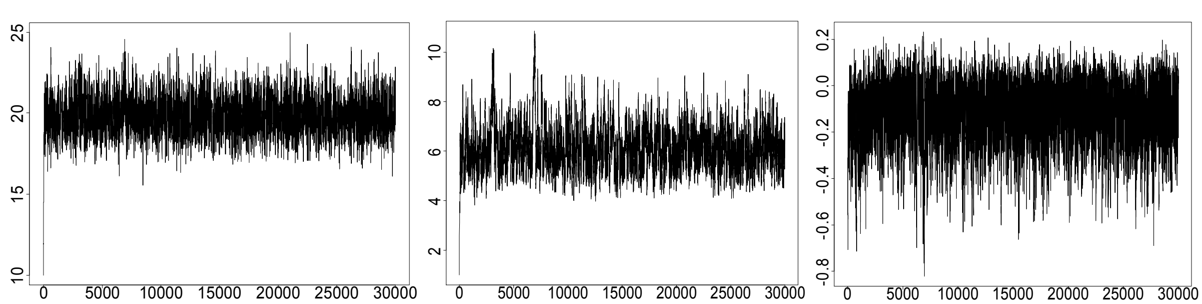
where, $x_{i:n}$ is observed values for *i*th order statistics of random sample of size n, $\bar{Q}(F_i) = \frac{1}{n}\sum_{i=1}^{n}\hat{Q}\,(F_i)$ is the estimated quantile values associated with Gringorton plotting position F_i. If RRMSE and RASE are used to compare the models, the smallest value of RRMSE and RASE will indicate best fitting distribution. However, when PPCC test is used,



(a)



(b)



(c)

Fig. 3: Trace plots of the GEV parameters using MCMC for station rainfall; (a): Bertam; (b): Dungun; (c): Pekan

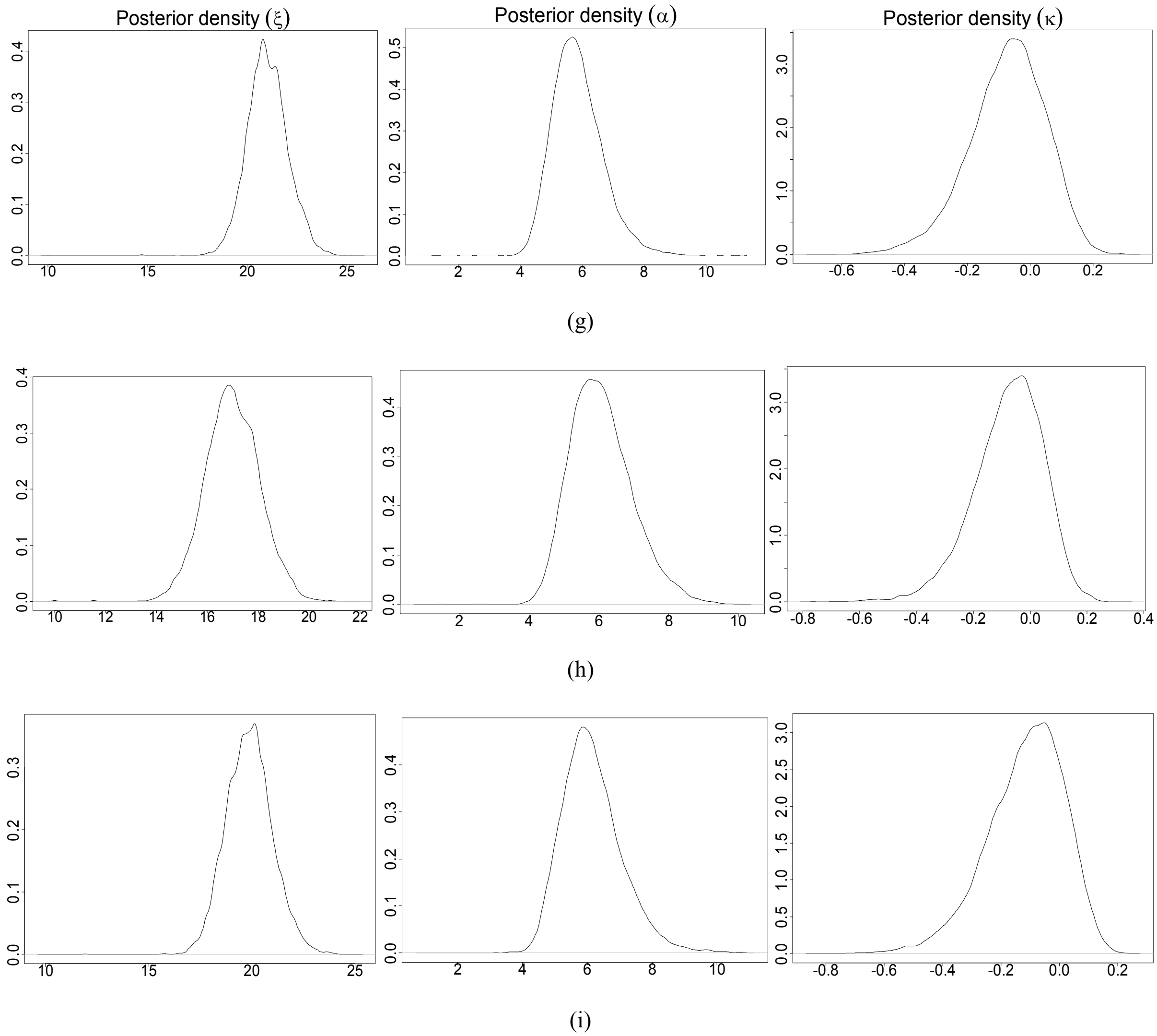


Fig. 4: Posterior densities of the GEV parameters using MCMC for station rainfall; (g): Bertam; (h): Dungun; (i): Pekan

the model with the computed PPCC value closest to 1 is the best.

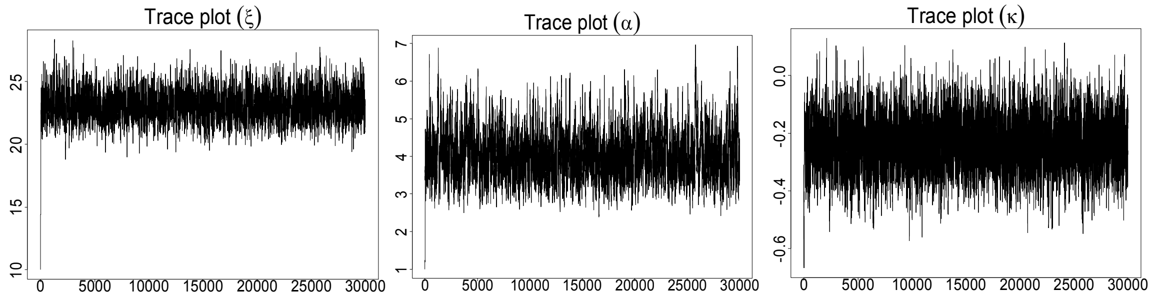
RESULTS

The parameters in the two probability distributions, the GEV distribution and the GL distribution, are estimated using L Moment and the Bayesian MCMC with the random walk chain algorithm.

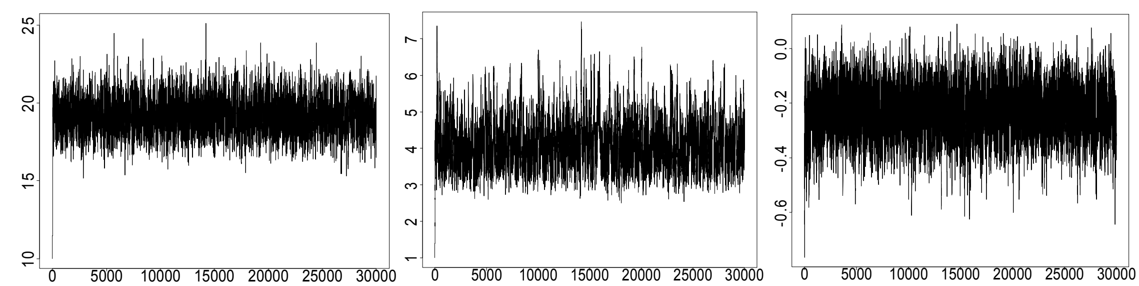
In the case of LMOM, The *r*th LMOM, denoted as λ_r can be solved using the simple mathematic to obtain estimates of parameters for GEV and GL distribution and in Bayesian framework the algorithm in 3.2 can be used to obtain parameters.

Hourly rainfall data for six rain gauge stations will be analyzed by algorithm in 3.2. In each case 30000 iterations of the algoritm were carried out. The MCMC trace plots and estimated posterior densities for GEV and GL parameters for six rain gauge are given in Fig. 3 to 6, respectively. To check that chains had converged to the correct place, the same algorithm was carried out using the various starting points. The chains for the six sites all converged very well within the first 10000 iterations. Therefore, it can be suggested that the developed proposal distribution works well.

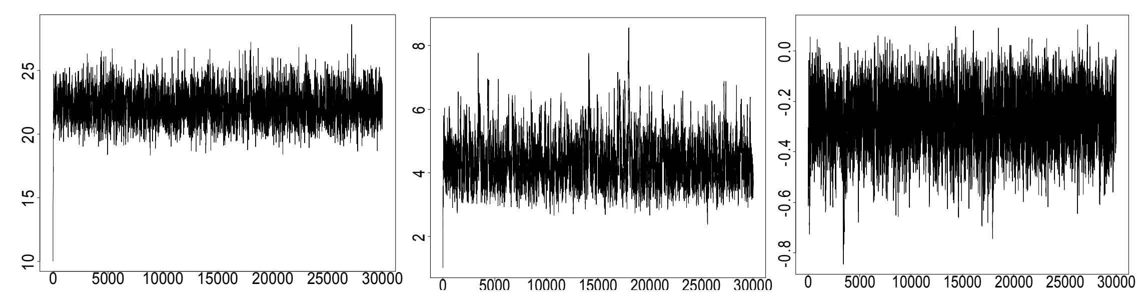
The result of the parameter estimation at six rain gauge station using GEV and GL distributions are summarize in Table 4. In the Table 4, the posterior means of Bayesian MCMC and the value of LMOM are very similar to each other regardless of the rain gauge stations and probability distributions. Therefore, it is suggested that the Bayesian MCMC has no advantage over the LMOM. However, from the point of the uncertainty, the Bayesian MCMC is more meaningful than the LMOM. Therefore, it is suggested that the Bayesian MCMC exhibits an advantage over the LMOM when the quantification of the uncertainty in the parameters is required.



(a1)

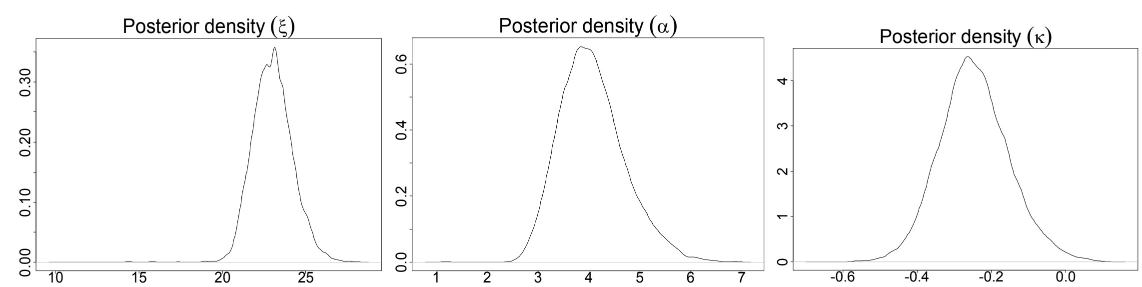


(b2)

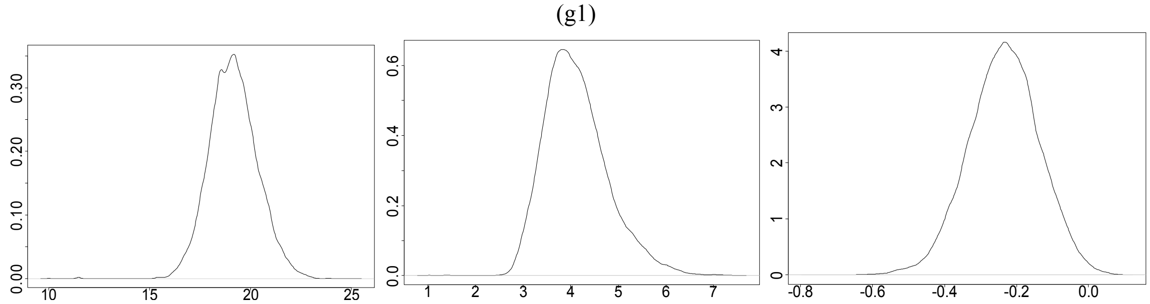


(c1)

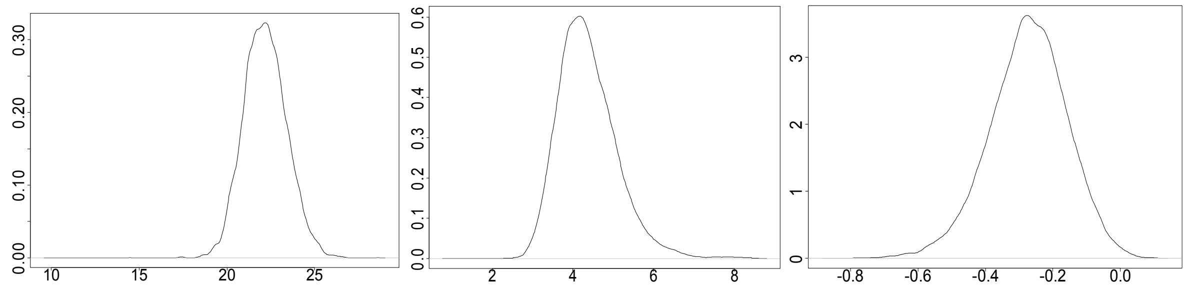
Fig. 5: Trace plots of the GLO parameters using MCMC for station rainfall; (a1): Bertam; (b1): Dungun; (c1:) Pekan



(g1)



(h1)



(i1)

Fig. 6: Posterior densities of the GLO parameters using MCMC for station rainfall; (g1): Bertam; (h1): Dungun; (i1): Pekan

Table 4: Posterior means by MCMC and L-moment for the GEV and GLO parameters for each station rainfall

	Bayes						L-mom					
	GEV			GLO			GEV			GLO		
Station	ξ	α	κ	ξ	α	κ	Ξ	α	κ	ξ	α	κ
Bertam	20.97	5.78	-0.06	22.99	4.01	-0.25	20.78	5.37	-0.09	22.91	3.73	-0.23
Dungun	16.97	5.99	-0.07	19.12	4.04	-0.23	16.81	5.57	-0.07	19.01	3.84	-0.22
Pekan	19.86	6.05	-0.10	22.15	4.28	-0.27	19.83	5.85	-0.07	22.13	4.02	-0.22

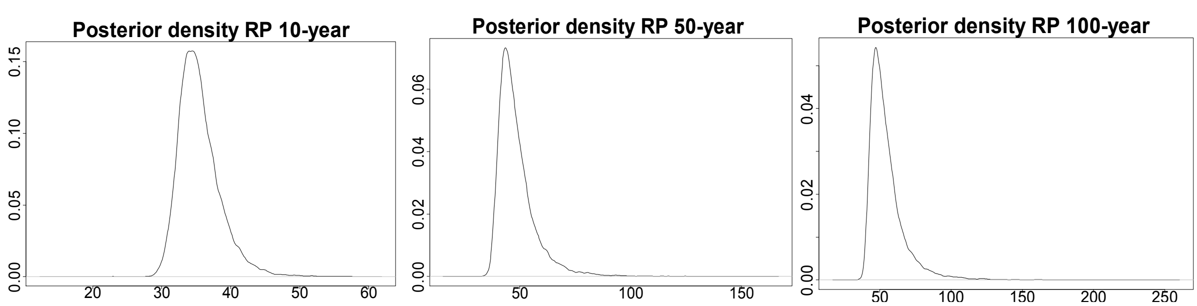
Table 5: Comparison of performance of MCMC versus L-moment under different GOF for GEV and GLO distributions

	GOF Bayes						GOF L-mom					
	GEV			GLO			GEV			GLO		
Station	RRMSE	RASE	PPCC	RRMSE	RASE	PPCC	RRMSE	RASE	PPCC	RRMSE	RASE	PPCC
Bertam	0.06	0.04	0.99	0.08	0.04	0.96	0.05	0.03	0.98	0.06	0.03	0.97
Dungun	0.07	0.05	0.99	0.08	0.04	0.97	0.06	0.04	0.99	0.06	0.04	0.98
Pekan	0.07	0.05	0.99	0.09	0.05	0.98	0.04	0.03	1.00	0.05	0.04	0.99

Comparing with the results of the GEV distribution and GL distributions using numerical GOF from Table 5, it can be concluded that for all GOF tests, the MSI of the six rain gauge stations follow a Generalized Extreme Value (GEV) distribution. When the performance of the methods of estimation are compared under a particular GOF test based on the proportion of time where one is better than the other, it is found that the LMOM is more superior than the Bayesian MCMC for all the three type of GOF tests considered. These results are summarized in Table 5.

Return period: Based on the best-fitted models, we can calculate the return values of the periods 10, 50 and 100 years for six rain gauge stations, by substituting the vectors of observations from the marginal posterior distributions of α, ε and κ into quantile function in Table 3, for $0 < F < 1$, samples from the posterior distribution of return levels can be obtained. This procedure was carried out for $p = 0.1, 0.5, 0.01$, to obtain the posterior distributions of the 10, 50 and 100-year return levels. Plot the posterior densities of the 10, 100 and 1000-year return levels for six rain gauge station using Bayesian MCMC for GEV and GL distribution are given in Fig. 7 and 8 respectively.

Due to the positive skew of the posterior distributions, seen in Fig. 7 and 8, the posterior medians are considered to be more suitable measures of location than the posterior means. Posterior medians and value of return period using LMOM for the three return levels for each rain gauge are given in Table 6. In Table 6, with the exception of rain gauge stations Seremban and



(m)

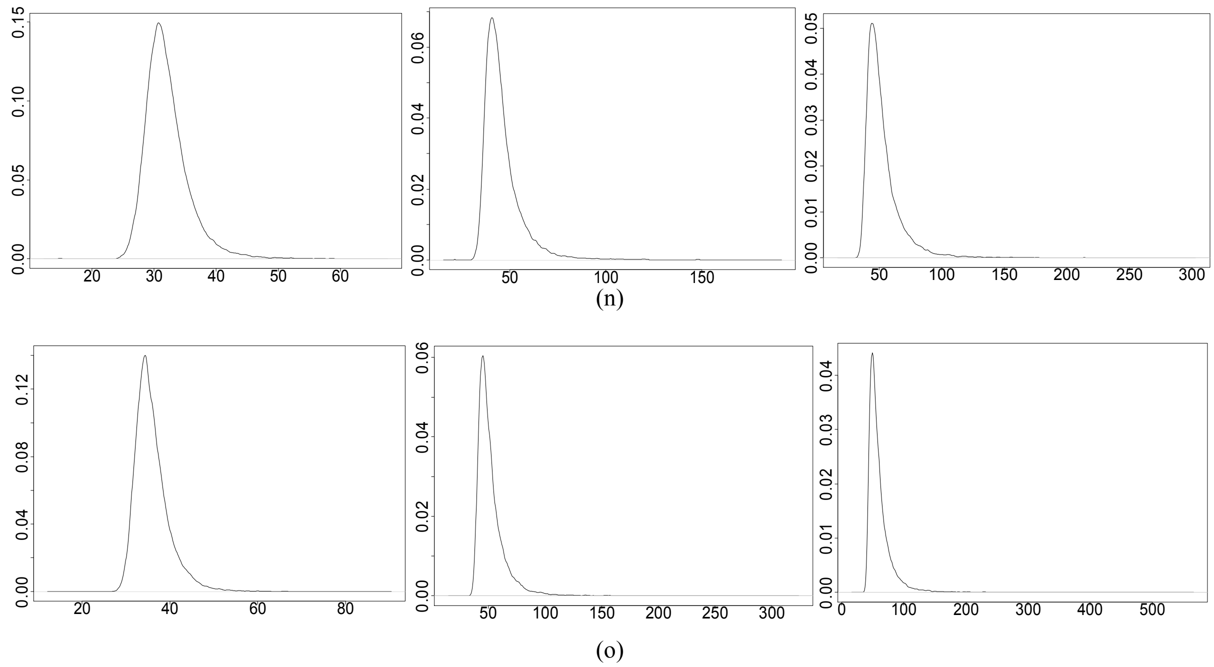


Fig. 7: 10, 50 and 100-year return period for station rainfall; (m): Bertam; (n): Dungun; (o): Pekan

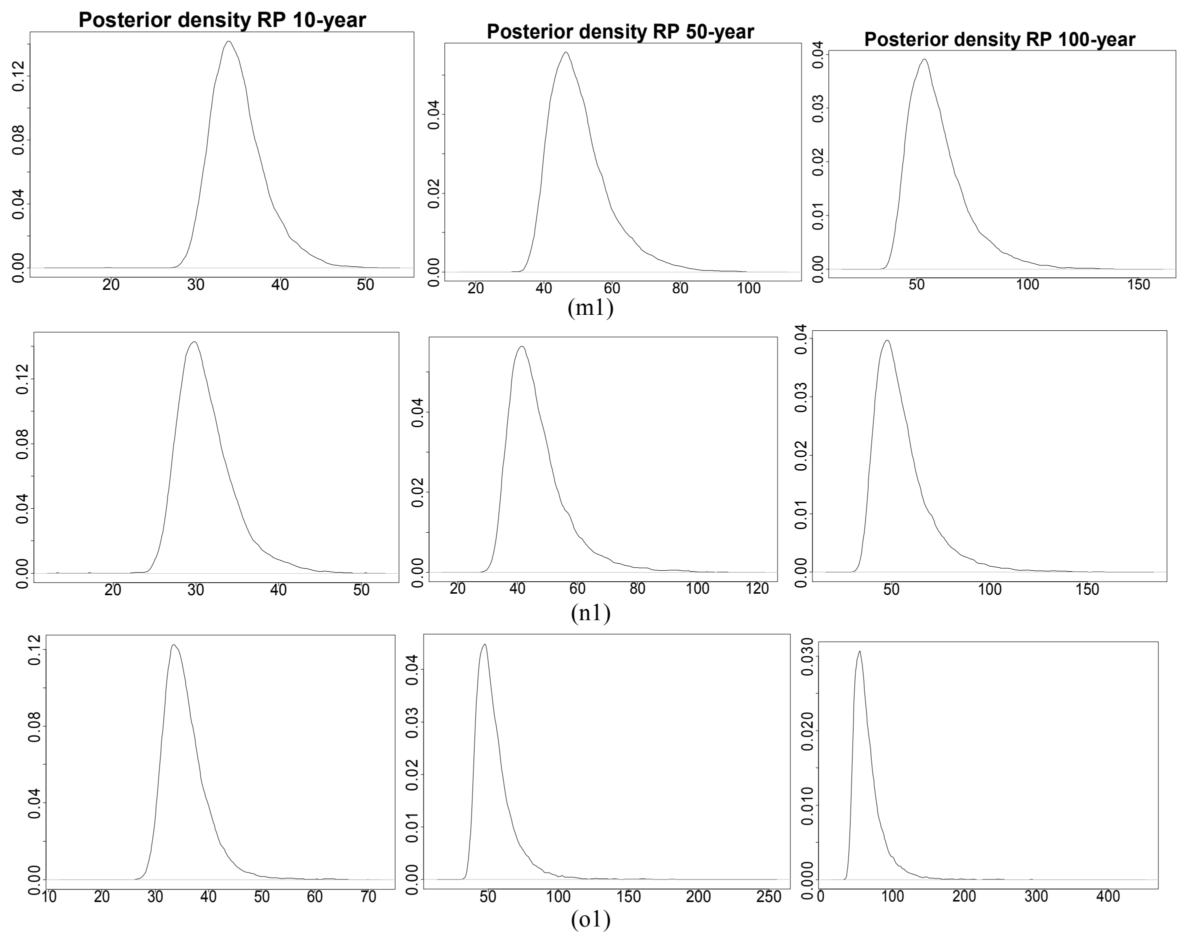


Fig. 8: 10, 50 and 100-year return period for station rainfall; (m1): Bertam; (n1): Dungun; (o1): Pekan using MCMC (GLO distribution

Table 6: Comparison of performance of MCMC versus L-moment under different return period for GEV and GLO distributions

	Return period bayes						Return Period L-mom					
	GEV			GLO			GEV			GLO		
Station	10-year	50-year	100-year	10-year	50-year	100-year	10-year	50-year	100-year	10-year	50-year	100-year
Bertam	35.00	46.42	51.59	34.58	48.83	56.66	34.20	45.95	51.47	33.58	46.40	53.38
Dungun	31.51	43.41	48.82	30.61	44.38	51.89	30.45	41.98	47.30	29.82	42.51	49.32
Pekan	35.17	48.61	55.00	34.90	51.22	60.50	34.10	46.13	51.65	33.44	46.68	53.78

Sungai Pinang, the posterior medians of the return levels are all higher than the LMOM estimates, particularly for the 50 and 100-year return levels, it seem to be more sensible estimates for the return levels than LMOM estimates. This is due to the large difference between the posterior mean of κ and the maximum likelihood estimates of κ.

CONCLUSION

In this study, the occurrence probability of the annual Maximum Storm Intensity (MSI) events was analyzed at the six rain gauge stations, in Peninsular Malaysia. The two probability distributions, the GEV distribution and the GL distribution, were selected to fit of the data. The two types of data in this study were analyzed by LMOM and the Bayesian MCMC, specially for estimate the parameters of the two probability distributions. In this study could be showed that the Bayesian MCMC worked well and efficient with the non informative prior distribution in this study by checking of the acceptance rate. From the results of the parameter estimation and quantile estimation, it was seen that the Bayesian MCMC had no advantage over the LMOM when the median or mean value was required. However, in the aspect of the uncertainty analysis, the Bayesian MCMC could remarkably reduce the range of the uncertainty. The reduction of the uncertainty in the results of the frequency analysis may not always give a good description for the all the cases. Also, Bayesian analysis cannot always provide the reduction of the uncertainty. Especially, if we have much information such as large sample size for the defining the unknown parameters, the influence of the uncertainty is relatively weak to determine a specific decision. However, if we have a little information, the analysis of the uncertainty has a strong influence on the final selection of the parameters. Therefore, the reduction of the uncertainty in the frequency analysis with the extreme event such as the rare rainfall event in this study can provide the meaningful description.

ACKNOWLEDGMENT

The authors are indebted to the staff of the Department of Irrigation and Drainage for providing hourly rainfall data to make this study possible. This research would not have been possible without the sponsorship from Universiti Kebangsaan Malaysia. This research was funded by UKM- DIP-2012-15.

REFERENCES

Adams, B.J., H.G. Fraser, C.D.D. Howard and M.S. Hanafy, 1986. Meteorological data analysis for drainage system design. J. Environ. Eng., 112(5): 827-848.

Adams, B.J. and F. Papa, 2000. Urban Stormwater Management Planning with Analytical Probabilistic Models. Wiley, New York.

Bernado, J.M. and A.F.M. Smith, 1994. Bayesian Theory. Wiley, New York, USA.

Carlin, B.P. and T.A. Louis, 1996. Bayes and Empirical Bayes Method for Data Analysis. Chapman & Hall, London.

Coles, S.G. and E.A. Powell, 1996. Bayesian methods in extreme value modelling: A review and new developments. Int. Stat. Rev., 64(1): 119-136.

Eagleson, P.S., 1972. Dynamics of flood frequency. Water Resour. Res., 8(4): 878-898.

Gelman, A., J.B. Carlin, H.S. Stern and D.B. Rubin, 1995. Bayesian Data Analysis. Chapman & Hall, London, UK.

Guo, Y. and B.J. Adams, 1998a. Hydrologic analysis of urban catchments with event-based probabilistic models: 2. Peak discharge rate. Water Resour. Res., 34(12): 3433-3443.

Guo, Y. and B.J. Adams, 1998b. Hydrologic analysis of urban catchments with event-based probabilistic models: 1. Runoff volume. Water Resour. Res., 34(12): 3421-3431.

Hosking, J.R.M. and J.R. Wallis, 1997. Regional Frequency Analysis: An Approach based on L-moments. Cambridge University Press, London, pp: 224.

Jeffreys, H., 1961. Theory of Probability. 3rd Edn., Clarendon Press, Oxford, UK.

Kuczera, G. and E. Parent, 1998. Monte Carlo assessment of parameter uncertainty in conceptual catchment models: The metropolis algorithm. J. Hydrol., 211(1-4): 69-85.

Palynchuk, B. and Y. Guo, 2008. Threshold analysis of rainstorm depth and duration statistics at Toronto, Canada. J. Hydrol., 348(3-4): 535-545.

Restrepo-Posada, P.J. and P.S. Eagleson, 1982. Identification of independent rainstorms. J. Hydrol., 55(1-4): 303-319.

Rivera, P., J. Gironás, J.P. Montt and B. Fernández, 2005. An analytical model for hydrologic analysis in urban watersheds. Proceeding of the 10th International Conference on Urban Drainage. Copenhagen, Denmark.

Suhaila, J., S.M. Deni, W.Z.W. Zin and A.A. Jemain, 2010a. Spatial patterns and trends of daily rainfall regime in Peninsular Malaysia during the southwest and northeast monsoons: 1975-2004. Meteorol. Atmos. Phys., 110(1): 1-18.

Suhaila, J., S.M. Deni, W.Z.W. Zin and A.A. Jemain, 2010b. Trends in Peninsular Malaysia rainfall data during the southwest monsoon and northeast monsoons seasons: 1975-2004. Sains Malays., 39(4): 533-542.

Vogel, R.M. and N.M. Fennessey, 1993. L moment diagrams should replace product moment diagrams. Water Resour. Res., 29(6): 1745-1752.

Zawiah, W.Z.W., A.A. Jemain, K. Ibrahim, J. Suhaila and M.D. Sayang, 2009. A comparative study of extreme rainfall in Peninsular Malaysia: With reference to partial duration and annual extreme series. Sains Malays., 38(5): 751-760.

Zin, W.Z.W., A.A. Jemain and K. Ibrahim, 2009. The best fitting distribution of annual maximum rainfall in Peninsular Malaysia based on methods of L-moment and LQ-moment. Theor. Appl. Climatol., 96(3): 337-344.

Zin, W.Z.W., S. Jamaludin, S.M. Deni and A.A. Jemain, 2010. Recent changes in extreme rainfall events in Peninsular Malaysia: 1971-2005. Theor. Appl. Climatol., 99(3): 303-314.