Proceedings of the Pakistan Academy of Sciences:
A. Physical and Computational Sciences 53 (2): 225–235 (2016)

ISSN: 0377 - 2969 (print), 2306 - 1448 (online)

Pakistan Academy of Sciences

Research Article

Application of Bayesian Monte Carlo Technique to Calculate Extreme Rainfall over Sindh Province in Comparison with Maximum Likelihood Method

Muhammad Ali[1, *], Muhammad Jawed Iqbal[2], and Zohaib Aziz[3]

[1] Department of Computer Science, Sindh Madrasatul Islam University,
Aiwan-e-Tijarat Road, Karachi-74000, Pakistan
[2]Institute of Space and Planetary Astrophysics, University of Karachi, Karachi, Pakistan
[3]Department of Statistics, Federal Urdu University of Arts, Science and Technology,
Karachi, Pakistan

Abstract: The consequences of statistical modeling of extreme rainfall are pivotal for civil engineering and planning division in order to instill the capability of structures of building that can withstand the extreme situation. Yearly maximum rainfall data of Karachi, Badin, Chhor from 1961-2010, and Rohri from 1971-2010 have been used in this study. The method of Maximum Likelihood (ML) and Bayesian techniques have been implemented to estimate the parameters of Generalized Extreme Value (GEV) distribution and also to compute return levels against sundry return periods. Non-informative priors are used to get the posterior densities. To gauge and compare the results of the above mentioned methods, acceptance rates and forecasting errors have been used as Goodness of Fit (GoF) test. Though both the methods are applicable, but the GoF test highlights that ML method is slightly better than Bayesian for observing the annual maximum rainfall in Sindh province of Pakistan.

Keywords: Yearly maximum rainfall, extreme value modeling, return period, goodness of fit test, return levels.

1. INTRODUCTION

The phenomenon of Earth's climate change with connected rainfall are mainly driven by the changes [1], caused by both natural and anthropogenic activities in the twentieth century, but their comparative roles and local influences are still under extreme debate [2]. The weather conditions which lie outside a locality's normal range are termed as extreme weather events. Extreme weather events, such as heat waves, cold waves, fog, thunderstorms, snowstorms, cyclones, hailstorms, floods and heavy rains etc., have a key importance. The economy and people's life generally hinge on these events [3]. Extreme value modeling of climatic processes is a customary practice for large scale construction [4], brought in generalized extreme value distribution. This distribution is repeatedly and extensively applied to model the extreme hydrological events, for example flood flow, extreme rainfall intensity [5], coastal water level [6] and extreme temperature [7]. The most climatic models in recent events suggest that the global warming is likely to augment the frequency of extreme weather events in many places [8]. The hot extremes, heat waves, and heavy precipitation events will persist to become more frequent [9].Maximum Likelihood(ML) method is used to embrace the result of covariates, for example the chance that one or more constraint of the distribution may contain a trend owing to climatic changes [10]. By means of historical data, estimates are calculated for design parameters, which are then used in construction to have a minimal failure probability [5]. These extreme events are frequently allied with climatic changes, which may be succeeded by natural calamities like landslides and flash floods. According to the report of the UNFCCC (United Nations Framework Convention on Climate Change), the Asian climatic change would influence the resources of water, food security and agriculture, biodiversity and ecosystems, public health and coastal zones. For any developing country similar to Pakistan, the natural disasters will

Received, March 2016; Accepted, May 2016
*Corresponding author: Muhammad Ali; Email: m.alishigri@gmail.com

assuredly affect the country's production. There is a conclusion from the analysis of climate change that it also influences negatively on paddy cultivation of rice [11]. It has been noted that such extreme climatic events are predictable at all. From statistical analysis of climate change, it is found that, the devastating effect of extreme rainfall may also be decreased by precautionary measures. In case of an extreme value modeling, one must consider other sources of knowledge like the known physical constraints, the maximum possible value or may be derived from an understanding of associated processes and possibly the same variable at a different location. Sometimes we may observe data, not to be completely representative for the whole period and there may be historical evidences, though not in the form of data, but in the shape of behavior , which is significantly more extreme than that which has been considered. As a result, there are a number of reasons that why it is to be expected that an expert with knowledge of the physical processes may have information that is pertinent to extreme behavior , which is also independent of the existing data. This leads obviously to the Bayesian inferential framework as a beginning for undertaking an extreme value analysis. Several attempts have been made to use Bayesian methodology in the extreme value analysis. For example, the most comprehensive analysis by Smith and Naylor [12] shows the effect of different prior assumptions on the posterior distributions of parameters of the Weibull distribution. The theoretical background of Bayesian estimation of extreme quantities, with partial regard for the significance of prior structure, has been explained by Pickands [13]. The advent of Markov chain Monte Carlo methodology estimates the marginal densities using Gibbs MCMC sampler [14]. In this way the current work is an attempt to calculate different return levels of extreme rainfall in Pakistan having best fitted model.

2. DATA DESCRIPTION AND METHODOLOGY

Peak Over Threshold (POT) and Annual Maximum Series (AMS) are commonly used in extreme rainfall analysis [8]. The AMS contains the maximum rainfall value occurring in a particular day of a year while POT involves all maximum rainfall values greater than the defined threshold value. POT is usually used to resolve the wastage of the data in AMS method. Due to the complex situation in choosing a suitable threshold in POT, the AMS is preferable method in extreme value analysis [15]. Therefore we use annual maximum series of rainfall of four meteorological stations of Sindh, from Jan, 1961 to Dec, 2010 for Karachi, Badin, Chhor and Jan, 1971 to Dec. 2010 for Rohri station. It is clear from Fig. 1, that there is no any particular trend in the data series. 40 years normal values of extreme rainfall of Karachi, Badin, and Chhor are 62.8 mm, 69.1 mm, and 69.8 mm respectively, while 30 years normal value of rainfall at Rohri is 41.2 mm. The maximum amount of extreme rainfall from 1961 to 2010 are 207.0mm that occurred in 1977, 241.0mm that occurred in 1979, 251.2mm that occurred in 1998 and 184.5mm in 1978 at Karachi, Badin, Chhor and Rohri, respectively.

Suitable probability distributions must be used to model extreme behavior of rainfall in order to get the best inference. Numerous literatures for example [16, 17] available to get the best-fitted distribution for annual maximum rainfall data; recommend GEV distribution as the best fitted one. It is found that for sample size $n \geq 30$, maximum likelihood method generally performs well over all values of shape parameter (k) in the range $-0.5 < k < 0.5$ [1]. Here we use probabilistic approach in order to get the return levels of extreme rainfall against different return periods. In this study we use two approaches (Bayesian approach and Maximum likelihood approach) and also compare their results to get a specific forecasting model for future extreme rainfall scenario of Sindh.

2.1. Bayesian Approach

Bayesian statistics investigate the uncertainty about the unknown parameters by using probability statements so that the unknown parameters are here regarded as random variables. These probability statements are conditional on observed values of rainfall. To compute the posterior distribution Markov Chain Monte Carlo (MCMC) simulation technique is repeatedly used for estimating parameters. To make use of MCMC, a parent distribution for producing simulated value of parameter is necessarily to be introduced. To select the anticipated distribution, it is very essential to check the appropriateness of that distribution, as the bad choice may significantly delay the convergence for equilibrium point. The appropriate acceptance rates in getting suitable proficiency of Morkov Chain Monte Carlo simulation are around 10 to 40% [6, 15]. So the likelihood function for $Z_1, Z_2, \dots, Zn$ is given by:

$$L(\mu,\varphi,\xi : Z_1, \dots, Z_n) = \prod_{i=1}^{n} f(Zi|\mu, \varphi, \xi), \quad (1)$$

The density of posterior distribution is directly proportional to the product of Prior and Likelihood distribution as;

$$f(\mu,\varphi,\xi|Z_1, \dots, Z_n) \propto L(\mu,\varphi,\xi : Z_1,\dots,Z_n) \times g(\mu,\varphi,\xi) \quad (2)$$

Where g, L and f are prior, likelihood and posterior distribution, respectively. The prior distribution shows the set of confidence about the required parameters. In this analysis 'g' is the non-informative prior distribution which shows that the significantly prior information about extreme rainfall over Sindh is not available at the moment. The location (μ) shape (ξ) and scale (φ) parameters are assumed to be normally distributed.

Generally the purpose of extreme rainfall analysis is to calculate the expected values of rainfall (return levels). If 'y' denotes the future rainfall values with pdf:

$$h(y|Z_1, \dots, Z_n) = \iiint f(y|\mu, \varphi, \xi) f(\mu, \varphi, \xi|Z1, \dots, Zn) d\mu d\varphi d\xi \quad (3)$$

Then the estimates of probability of 'n' year return levels for sample $\theta_1, \theta_2, \dots, \theta_R$ will be:

$$\Pr[y \le q_n|z_1, \dots, z_n] \approx \frac{1}{R}\sum_{i=1}^{R} \Pr(y \le qn|\theta i) \quad (4)$$

Since Eq. (3) is very complicated to solve directly, so we use MCMC simulation technique to find the posterior distribution. For MCMC simulation we use R-statistical tool with loading two appropriate extreme value modeling (EVM) packages named as texmex and mvtnorm.

2.2. Probability Distribution and Maximum Likelihood Method

Let $Z_1, Z_2 \dots, Z_n$ be the daily rainfall series where $M_n = \max\{Z_1, Z_2 \dots, Z_n\}$ with n=365, is the annual maximum rainfall data. Asymptotic considerations recommend that the distribution of M_n should be something like that of a member of the generalized extreme value (GEV) family having probability distribution function (PDF):

$$f(x) = \frac{1}{\varphi}\left[1 + \xi\left(\frac{Zi-\mu}{\varphi}\right)\right]^{-1-\frac{1}{\xi}} \times e^{-\left[1+\xi\left(\frac{Zi-\mu}{\varphi}\right)\right]^{-\frac{1}{\xi}}} \quad (5)$$

Where μ,φ and ξ are location, scale and shape parameter with parameter space $-\infty < \mu < \infty$, $\varphi > 0$ and $-\infty < \xi < \infty$. Eq. (5) can be integrated analytically to obtain cumulative density function (CDF) as;

$$F(x) = e^{-\left[1+\xi\left(\frac{Zi-\mu}{\varphi}\right)\right]^{-\frac{1}{\xi}}} \quad (6)$$

This equation can be reversed to find an appropriate formula for return levels (q_r) of rainfall corresponding to 1/r years return period [3].

$$q_r = \mu + \frac{\varphi}{\xi}\{1-[-\log(1-p)]^{-\xi}\} \quad (7)$$

The above mentioned parameters are then calculated by both ML method and Bayesian method.

3. GOODNESS OF FIT TEST

Now we will discuss about the best fitted model between Bayesian and maximum likelihood method for the estimation of GEV parameters and future predicted extreme rainfall of above said stations. In this work we use three Goodness of Fit (GoF) tests, named as Relative Absolute Squared Error (RASE), Relative Root Mean Squared Error (RRMSE) and Probability Plot Correlation Coefficient (PPCC). RRMSE and RASE include the measurement of inconsistency among observed and predicted values through the parent distribution (GEV), while PPCC measures the correlation between observed and forecasted values. Now we use annual maximum rainfall data from 1971 to 2000, and forecast the return levels from 2002 to 2010 through both Bayesian and Maximum likelihood method. A method is said to be better in forecasting, if it has less values of RRMSE and RASE and closest value of PPCC to 1 and vice versa. Now as suggested by Zin [17],we will also observe the above said three GOF tests by using the following formulas as;

$$\text{RRMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{X_{i,o}-Q(x_i)}{X_{i,o}}\right)^2} \tag{8}$$

$$\text{RASE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{X_{i,o}-Q(X_i)}{X_{i,o}}\right| \tag{9}$$

$$\text{PPCC} = \frac{\sum_{i=1}^{n}\left(X_{i,o}-\bar{X}\right)\left(\hat{Q}_{(X_i)}-\bar{Q}_{(X_i)}\right)}{\sqrt{\sum_{i=1}^{n}\left(X_{i,o}-\bar{X}\right)^2\sum_{i=1}^{n}\left(\hat{Q}_{(X_i)}-\bar{Q}_{(X_i)}\right)^2}} \tag{10}$$

Where $X_{i,o}$ is observed values of i^{th} order statistics of yearly extreme rainfall from 2002 to 2010, and $Q(x_i)$ is the estimated values for this period. Now the summary of GoF tests is shown in table (2). The statistical analysis of table (2) indicates that, although there is no any big difference between ML and Bayesian method, showing that both the methods are suitable to forecast the extreme rainfall values, but a little bit difference in forecasting errors (i.e. less values of RASE and RRMSE), indicates that ML is more appropriate method for the observed data of Sindh. On the basis of AIC and MSE we can also compare the accuracy of our model for extreme yearly rainfall of observed meteorological stations of Sindh as shown in table (3).

Table 1. Comparison of Estimates of GEV parameters for different cities of Sindh with Standard Deviation (SD) in parenthesis.

GEV Parameters	Rohri		Chhor		Badin		Karachi	
	Bayesian	**ML**	**Bayesian**	**ML**	**Bayesian**	**ML**	**Bayesian**	**ML**
μ (location)	21.65 (3.56)	21.59 (3.4)	40.69 (5.28)	40.72 (5.15)	44.43 (6.03)	44.59 (5.9)	39.21 (5.6)	39.11 (5.42)
Φ (Scale)	2.96 (0.16)	2.9 (0.16)	3.48 (0.13)	3.45 (0.13)	3.6 (0.13)	3.57 (0.13)	3.51 (0.14)	3.46 (0.14)
ξ (Shape)	0.39 (0.166)	0.37 (0.164)	0.19 (0.13)	0.17 (0.13)	0.11 (0.13)	0.08 (0.14)	0.195 (0.15)	0.18 (0.15)

Table 2. Comparative study of forecasting errors between Bayesian and maximum likelihood estimates of Sindh.

Station Name	Method of Estimation	Goodness of Fit Test		
		RASE	RRMSE	PPCC
Rohri	Bayesian	2.36	7.08	0.945
	ML	2.13	6.39	0.956
Chhor	Bayesian	1.56	4.68	0.249
	ML	1.44	4.31	0.276
Badin	Bayesian	1.24	3.72	0.217
	ML	1.13	3.38	0.241
Karachi	Bayesian	0.73	2.19	0.266
	ML	0.70	1.09	0.273

Table 3. Comparative study of AIC, MSE and acceptance rate for observed meteorological stations of Sindh.

S.No	Sation's Name	WMO Number	Latitude	Longitude	AIC	MSE	Acceptance Rate
1	Rohri	41725	27^o 40'	68^o 54'	382	0.110	0.311
2	Chhor	41685	29^o 53'	69^o 43'	518	0.013	0.315
3	Badin	41785	24^o 38'	68^o 54'	526	0.082	0.309
4	Karachi	41780	24^o 54'	66^o 56'	521	0.327	0.312

Table 4. Comparison of return levels against different return period of Bayesian and ML methods for Sindh.

Station Name	Method of estimation	Return Level (mm)				
		10 year	25 year	50 year	75 year	100 year
Rohri (1971-2010)	ML	86	134	122	216	243
Chhor	ML	126	173	213	239	258
Badin	ML	133	175	209	230	245
Karachi	ML	128	178	222	250	271

4. RESULTS AND DISCUSSION

The amount of rainfall expected to occur at a station in future for a defined period of time is said to be return level of that station. The likelihood function can be assembled for complex situation of modeling for example non-stationary, effects of covariates and regression model etc. [18, 19].

MCMC technique is a way of simulation, for a complex distribution. The simulated values for all parameters of GEV distribution have been found together in a similar zone as revealed in Fig (2). The trace plots for scale and location parameters of GEV distribution for 40000 iterations are also shown in Fig (2). For non-informative priors, the variances have to be chosen large enough to get flat priors.

Fig (3) shows the posterior density plots for location, scale and shape parameters of parent distribution (i.e. GEV distribution), from which we can observe that all these figures are symmetrical in shape. The wide spread distributions in these figures indicate a big variance in non-informative prior distribution. In this study we have used Gibbs sampling in combination of Metropolis-Hasting scheme to get the desired posterior distribution. Our calculation also shows that the acceptance rate of Karachi, Badin, Chhor and Rohri are 31.2%, 30.9%, 31.5% and 31.1% respectively.

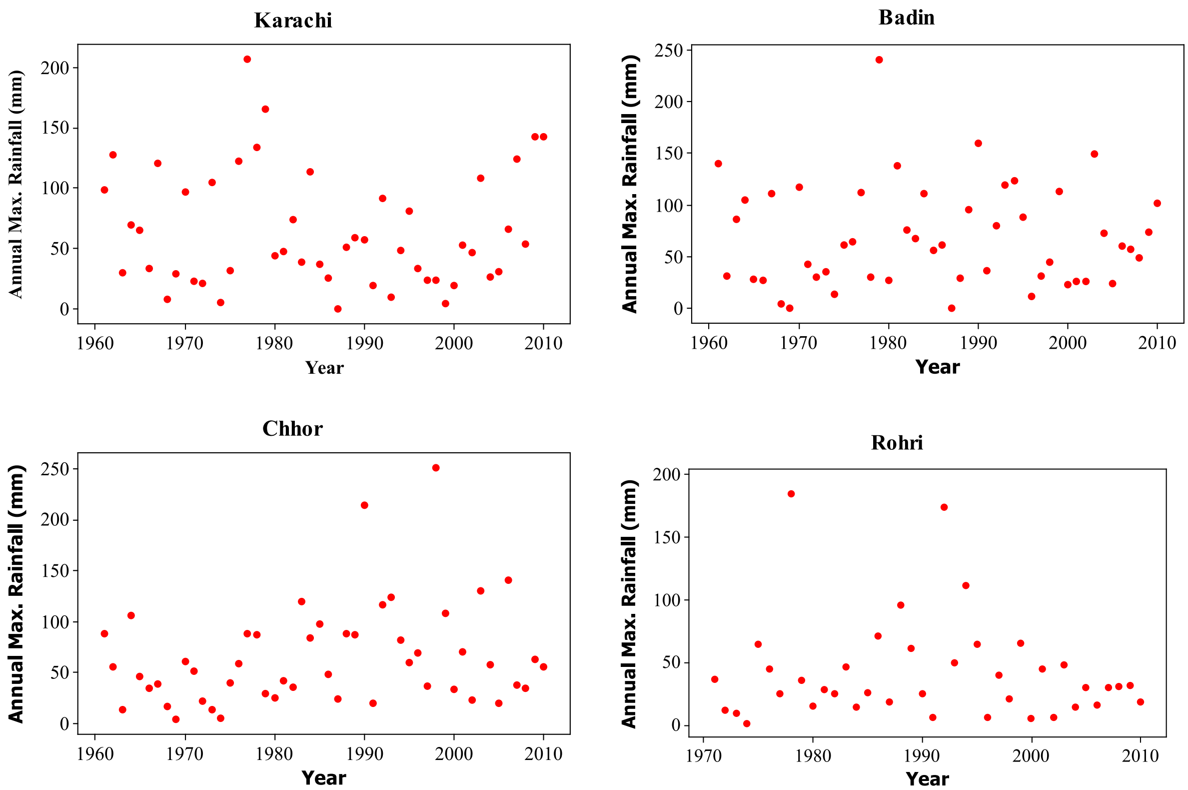
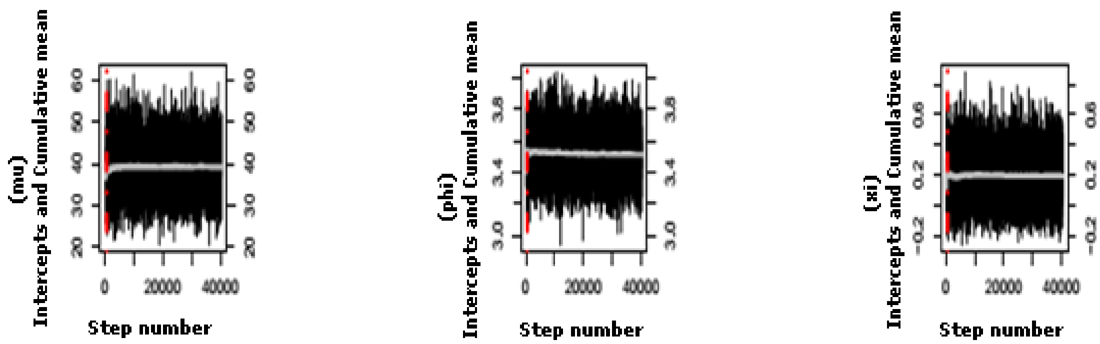


Fig. 1. Annual maximum rainfall over Sindh, 1960-2010.

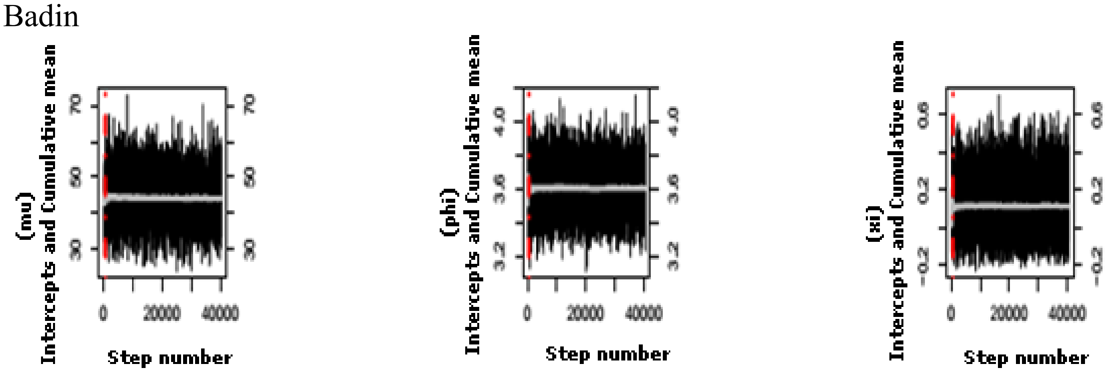
The ML method selects those parameters which maximize the likelihood of the given data. In ML method, parameters are supposed to be unidentified but fixed, and are calculated approximately. The suitability of this method for observed data points of given meteorological stations, can also be seen from Figs. (4).The model values and empirical values of Karachi, Badin, Chhor and Rohri approximately overlap the actual line in ML method. Therefore we can conclude that ML can explain more accurately the yearly maximum rainfall behavior of these stations. The values of parameters and their standard deviations shown in Table (1), also indicate that the ML method gives slightly better results than Bayesian.

Now with the help of R-statistical tool for extreme value modeling (EVM), we have first predicted the return levels of the above mentioned stations of Sindh from 2002 to 2010 designed for both methods Bayesian and ML. Comparing these forecasted values with observed values of yearly maximum rainfall from 2002 to 2010, we have calculated forecasting errors using eq (8), eq (9) and eq (10). Forecasting errors conclude that maximum likelihood is slightly better than Bayesian for yearly maximum rainfall of above mentioned stations of Sindh. Hence we have forecasted different return levels for 10, 25, 50, 75 and 100 years return periods as shown in Table 4. Table 4 depicts that the 50 years return levels estimated by ML method for Karachi, Badin, Chhor and Rohri are 222 mm, 209 mm, 213 mm and 122 mm respectively. While that of 100 years return levels are 271 mm, 245 mm, 258 mm and 243 mm yearly maximum rainfall in 24 hours, respectively. Comparing these predicted return levels of the above stations except Rohri, the amount of yearly maximum rainfall values of Karachi is the greatest while that of Badin is the least. Our analysis also suggests that, for short term forecast, the results obtained by Bayesian and ML are comparable to each other but for long range forecast they differ from one another. When these extreme rainfalls remain continuous for few days, it may lead floods, which will be very destructive and devastating for that region. Therefore, the purpose of calculating these return levels is to inform engineers and higher authorities, about the incoming situation of extreme rainfall over Sindh, in order to take precautionary steps to save our country from different kinds of losses.

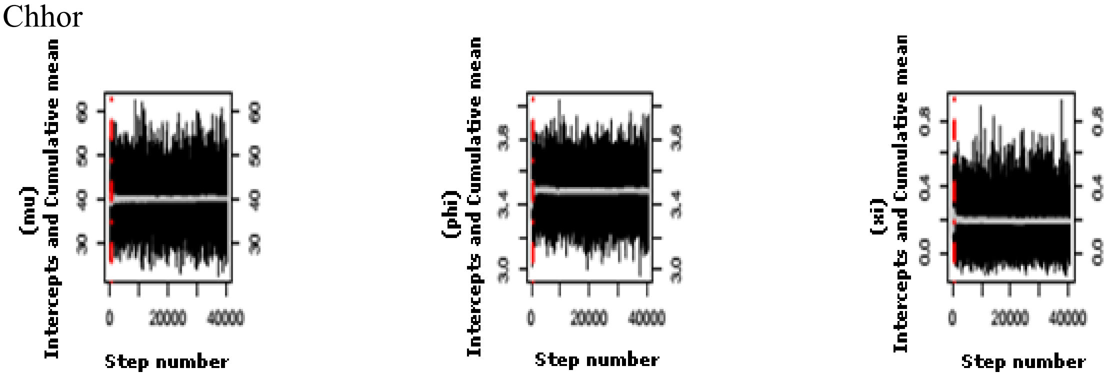
Karachi



Badin



Chhor



Rohri

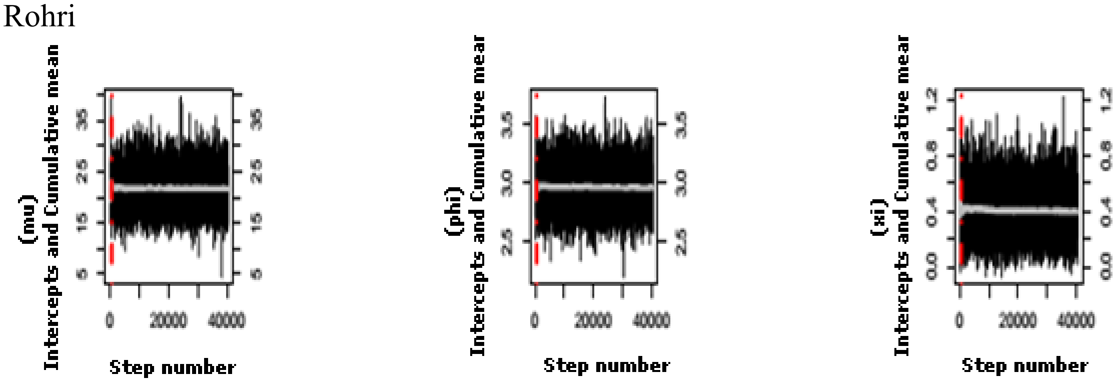
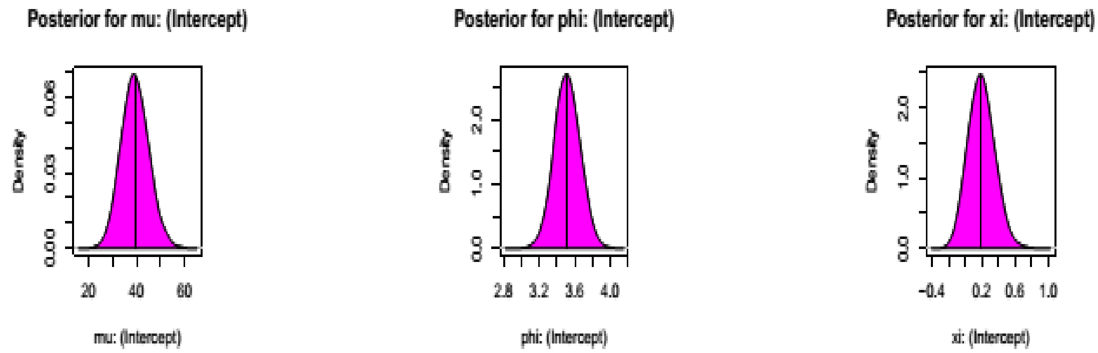
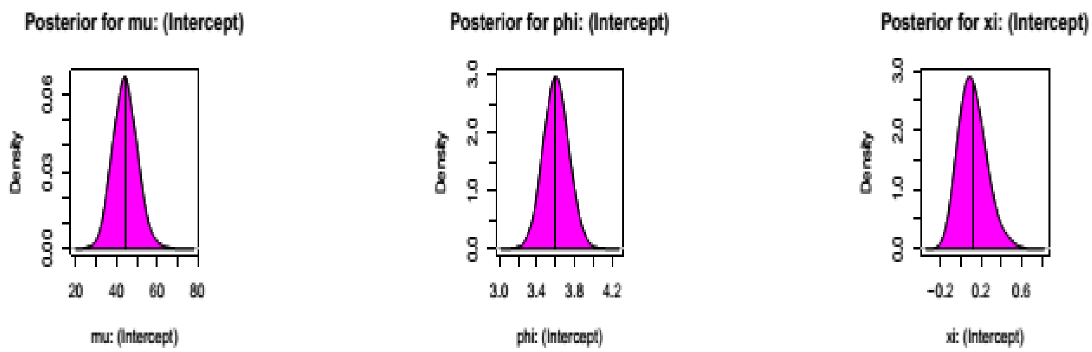


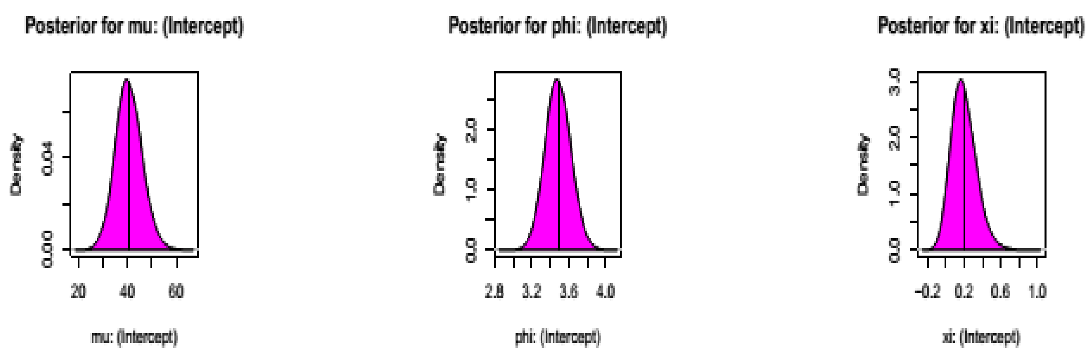
Fig. 2. Trace plots for location, scale and shape parameters of GEV distribution for 40000 iterations for Sindh.



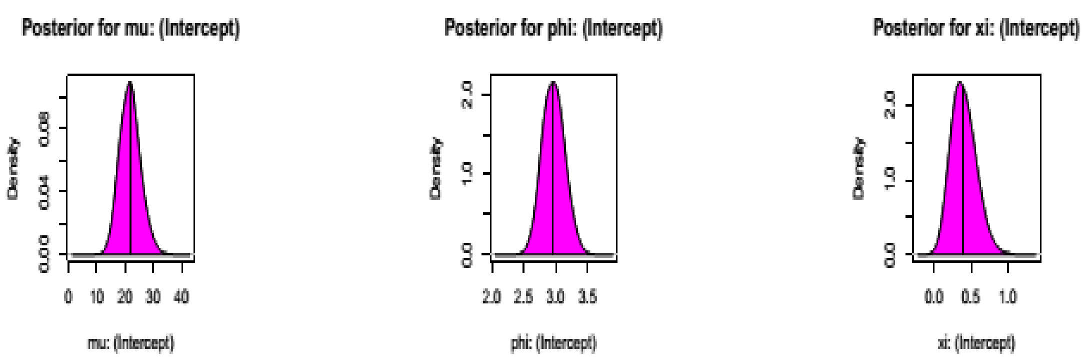
Karachi



Badin



Chhor



Rohri

Fig. 3. Posterior density plots for different cities of Sindh.

Karachi

Badin

Chhor

Rohri

Fig. 4. Probability plot and Quantile plot of maximum likelihood method for Sindh.

5. CONCLUSIONS

The main purpose of carrying out this analysis is to apply, and compare the results of Bayesian approach and Maximum Likelihood method, and also to estimate the return levels against different return periods. We utilize yearly maximum rainfall values of Karachi, Badin, Chhor and Rohri, for the period of January 1961 to December 2010. It was observed that maximum likelihood and Bayesian MCMC technique are strongly related because both methods involve likelihood function in their initial steps. In this study we used Gibbs sampling in combination of Metropolis-Hasting scheme to get the desired posterior distribution. Our calculation also shows that The acceptance rates in getting high proficiency of MCMC simulation for Karachi, Badin, Chhor and Rohriare 31.2%, 30.9%, 31.5% and 31.1%, respectively. Because of non-informative priors, no any big difference has been found between Bayesian and Maximum Likelihood method in calculating the values of GEV parameters. But the deeply analysis of forecasting errors i.e. Relative Root Mean Squared Error (RRMSE), Relative Absolute Squared Error (RASE)and Probability Plot Correlation Coefficient(PPCC),shows that ML method is slightly better than Bayesian method for the calculation of return levels of observed meteorological stations. Furthermore, Fig (4) also showed that the model values and empirical values of the above said meteorological stations are approximately overlapping the actual line in ML method. Therefore we can conclude that ML can explain more accurately the yearly maximum rainfall behavior of these stations. The values of parameters and their standard deviations shown in Table (1), also indicated that the ML method gives slightly better results than Bayesian, so we can use it as the best fitted model. We also observe that for short term forecast, the results obtained by Bayesian and ML methods are very close to each other but for long range forecast they differ from one another.

Hence the estimated return levels for 50 years return period by ML method are 222 mm, 209 mm, 213 mm and 122 mm while those for 100 years return period are 271 mm, 245 mm, 258 mm and 243 mm for above mentioned meteorological stations in 24 hours, respectively. These return levels are very useful for planning division, civil engineering, forecasting sections of Pakistan Meteorological Department, ministry of climate change, etc. So our work suggests to upgrade the flood forecasting system by using modern technology i.e. GIS based technology, and to get remarkable improvement in the river structures of Sindh province of Pakistan.

6. ACKNOWLEDGEMENT

We are very thankful to Pakistan Meteorological Department (PMD) for providing the required daily rainfall data.

7. REFRENCES

1. Bansod, S.D., H.N. Singh, S.D. Patil, & N. Singh. Recent changes in the circulation parameters and their association with Indian summer monsoon rainfall. *Journal of Atmospheric and Solar-Terrestrial Physics* 77: 248-253 (2012).
2. Zin, W.Z.W., A.A. Jemain, & K. Ibrahim. The best fitting distribution of annual maximum rainfall in Peninsular Malaysia based on methods of L-moment and LQ-moment. *Theoretical and Applied Climatology* 96(3-4): 337-344 (2009).
3. Ali, M., & M.J. Iqbal. A probabilistic approach for estimating return periods of extreme annual rainfall in different cities of Khyber Pakhtunkhwa (KP), Pakistan. *Nucleus* 49(2): 107-114 (2012).
4. Jenkinson, A.F. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of the Royal Meteorological Society* 81(348): 158-171 (1955).
5. Coles, S.G., & J.A. Tawn. A Bayesian analysis of extreme rainfall data. *Applied Statistics* 45: 463-478 (1996).
6. Huang, Wenrui, S. Xu, & S. Nnaji. Evaluation of GEV model for frequency analysis of annual maximum water levels in the coast of United States. *Ocean Engineering* 35(11): 1132-1147 (2008).

7. Siliverstovs, B., R. Ötsch, C. Kemfert, C.C. Jaeger, A. Haas, & H. Kremers. Climate change and modeling of extreme temperatures in Switzerland. *Stochastic Environmental Research and Risk Assessment* 24(2): 311-326 (2009).
8. Cotton, J., & C. Ferro. *Bayesian Priors and Estimation of Extreme Rainfall in Camborne*. MAS3029 Mathematical Sciences Project 1, University of Exeter, Exeter EX4, Devon, UK (2008).
9. IPCC. IPCC Report "*Climate Change 2007: The Physical Science Basis, Summary for Policymakers"*. Contribution of Working Group I to the Fourth Assessment Report, Paris (2007).
10. Katz, R.W., G.S. Brush, & M.B. Parlange. Statistics of extremes: Modeling ecological disturbances. *Ecology* 86(5): 1124-1134 (2005).
11. Eli, A., M. Shaffie, & W.Z.W. Zin. Preliminary study on Bayesian extreme rainfall analysis: A case study of AlorSetar, Kedah, Malaysia. *Sains Malaysiana* 41(11): 1403-1410, (2012).
12. Smith, R.L., & J.C. Naylor. A comparison of maximum likelihood and Bayesian estimators for the three-parameter Weibull distribution. *Applied Statistics*: 358-369 (1987).
13. Pickands III, J. Bayes quantile estimation and threshold selection for the generalized Pareto family", In: *Extreme Value Theory and Applications*, Springer US, p. 123-138. (1994).
14. Gelfand, A.E., & A.F. Smith. Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association* 85(410): 398-409 (1990).
15. Madsen, H., F.F. Rasmussen, & D. Rosbjerg. Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events: 1. At-site modeling. *Water Resources Research* 33(4): 747-757 (1997).
16. Zalina, M., M. Desa, V. Nguyen, & A. Kassim. Selecting a probability distribution for extreme rainfall series in Malaysia. *Water Science & Technology* 45(2): 63-68 (2002).
17. Zin, W.Z.W., A.A. Jemain, K. Ibrahim, S. Jamaludin, & S.M. Deni. A comparative study of extreme rainfall in Peninsular Malaysia: with reference to partial duration and annual extreme series. *Sains Malaysiana* 38: 751-760 (2009).
18. Coles, S. & M.J. Dixon. Likelihood-based inference for extreme value models. *Extremes* 2(1): 5-23 (1999).
19. Coles, S.G., & A. Jonathan. Tawn. A Bayesian analysis of extreme rainfall data. *Applied Statistics* 45(4): 463-478 (1996).
20. Ailliot, P., C. Thompson, & P. Thomson. Mixed methods for fitting the GEV distribution. *Water Resources Research*, doi: 10.1029/2010WR009417 (2011).