International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2015): 78.96 | Impact Factor (2015): 6.391

Bayesian Modelling of Extreme Rainfall Data: Construction of Priors Using WRF Model Outputs, A Case of Dar es Salaam Tanzania

Edwin C. Rutalebwa

Dar es Salaam Institute of Technology (DIT), Department of General Studies, P.O. Box 2958, Dar es Salaam, Tanzania

Abstract: ***A 54 years dataset (1961 -2014) of recordings of the maximum daily (24 h) rainfall in the Dar Es Salaam area, Tanzania is analyzed using Extreme Value Theory with Bayesian framework. Prior distributions are constructed using quantiles approach. Experts in the field are normally used to elicit distribution of quantiles. With scarcity of data, these experts may not exist. In this paper a different approach in eliciting prior distribution is proposed. In this proposal the quantiles for prior construction are obtained from the Weather Research Forecasting (WRF) model. For the case of Dar es Salaam, WRF outputs are generated based on the physical conditions around 20th December 2011, a day when Dar es Salaam experienced the extreme rainfall which has never been experienced for more than 50 years. A combination of these two data sets, through Bayesian framework, has improved the reliability of forecasting of extreme events. The point estimates for both Maximum Likelihood and Bayesian estimation methods are almost the same, but the associated 95% confidence intervals for Bayesian method are narrower highlighting the reliability of the estimator. It is demonstrated in this paper that WRF outputs can be used to construct prior distributions, and hence improve reliability of extreme rainfall forecasting.***

Keywords: Bayesian modeling, WRF, Generalized Extreme Value distribution, Prior distribution, Extreme rainfall, Dar es Salaam

1. Introduction

Extreme rainfall events cause significant damage to many sectors. These include agriculture, infrastructure and transport, human settlements, heath sector, insurance, to mention few. In Tanzania many areas are usually affected by heavy rainfall. The floods that occurred in Dar es Salaamin the year 2011 are example of such destructive heavy rainfall. More examples of such destructive heavy rainfall in Tanzania can be found, for example, in Ngailo et al (2016).

Most climate models indicate that in many places global warming is likely to increase the frequency of extremeweather events [14]. The United Nations Intergovernmental Panel on Climate Change (IPCC) issued a report inFebruary 2007 stating that “It is very likely that hot extremes, heat waves, and heavy precipitation eventswill continue to become more frequent”. Therefore, understanding the patterns of extreme rainfall and their future behaviour is very important to policy makers in Tanzania. To achieve this,a long historical dataset on rainfall is needed. In Tanzania, like other developing (even developed) countries, such data are not available. Classical statistical techniques for the analysis of extreme rainfall, that requires long historical data, may not produce the reliable results for scarcity data. In such case, a Bayesian framework for extreme values analysis is usually preferred [14], [20]. Using Bayesian inference in such situation would allow any additional information about the processes to be incorporated as prior information. The benefits of using any information available are likely to be great- due to lack of data; however, there is some concern that it may not be possible to formulate such prior information. Coles and Powell (1996) comment that if data on the extremes are so scarce, then it may not be possible for an expert to independently formulate prior beliefs about the process. Alternatively, one would opt to use non-informative prior when no such informative prior is possible.This paper aims at using the output from Weather Research and Forecasting (WRF) model as a source (external) of an informative prior. Furthermore, this will link two data sets: gauging (or automatic) station records and WRF outputs in order to enhance forecasting of extreme events.

Recently there has been an increasing interest in Bayesian methods applied to extreme value problems, and there have been a number of studies on extreme value problems (see for example Coles and Powell (1996) and Coles and Tawn (1996, 1991)). In most cases atrivariate normal prior distribution of parameters is used. The means of marginal distributions are set to zero (except few cases for shape parameter [14]) and variances are set to high values to reflect the absence of external information. Another approach is to construct a prior distribution in terms of quantiles. Coles andTawn (1996) argue that eliciting prior information in terms of extreme quantiles is sensible because this is a scale on which an expert is most likely to be able to accurately quantify their prior beliefs about extremal behaviour. Likewise, the author of this paper argues that since the outputs from WRF on extremal rainfall are in quantiles forms, then it is sensible to construct prior distribution from these outputs using quantiles approach.

Thepaper is organizedasfollows. Section2, describes study area, data and methodology for Bayesian methods, likelihood of extremes and prior distribution construction. Finally, the results and conclusion are presented insection 3.

2. Materials and Methods

2.1 Study Area

Dar es Salaam is located in the eastern part of Tanzania mainland between latitudes 60S and 70S and longitudes 33.330E and 390E. To the east, it borders Indian Ocean. It stretch about100 km between the Mpiji River to the north and beyond the Mzinga River in the south, enclosing a land of 1,350 km^2.

Tanzania has a tropical equatorial ltype of climate. Dar es Salaam region, which is Northern Coast of Tanzania, has an average temperature of 25.9^o C and about 1074.5 mm of precipitation falls annually [6]. The driest month is August with average of 23mm of rain and the most of precipitation falls in April with average 251mm. February is the warmest month of the year with average temperature of 27.9^o C and July is the coldest month with average temperature of 23.8^o C.

2.2 Data

Two types of data sets are used in this study: the daily rainfall and WRF outputs. The dailyrainfall data set recorded atDar es Salaam Airport, Tanzania (latitude: 6.87 S; longitude: 39.20 E; 53 m above mean sea level) between 1st January 1961 - 31st December 2014wasobtained from the Tanzania Meteorological Agency (TMA), while the WRF outputs are the outputs of the experiment run for 15 days (14th-28th December 2011).WRF experiments were run using optimal parameterization for Dar es Salaam region. For more details on these experiments refer toNgailo (2017).

Figure 1 represents the daily rainfall at Dar es Salaam over the period 1961-2014 and we can see some extreme rainfall (>50mm) over the periods. Figure 2 represents the WRF outputs of the experiment on rainfall for Dar es Salam in December 2011. In this month Dar es Salaam experienced heavy rainfalls for three consecutive days (19th to 21st, with the highest on 20th) which have not been experienced for more than 50 years according to TMA [7]. The experiments were run over the period from 14th to 28th December 2011. As it can be seen, the experiment satisfactorily captured the event (WRF outputs are indicated by "+") although the highest quantile (156.4 mm) was not captured adequately.

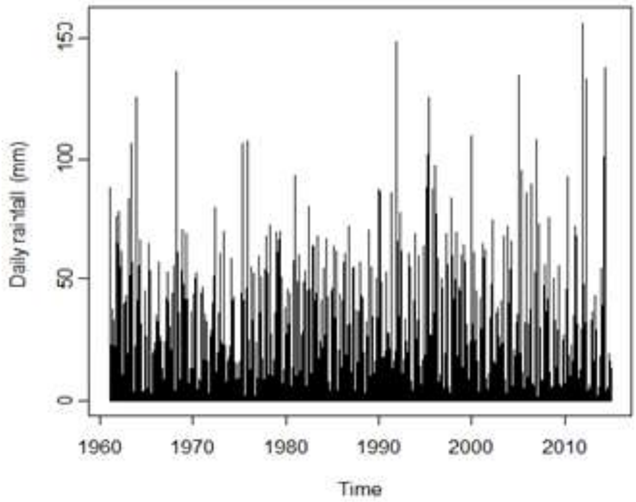

Figure 1: Original Daily rainfall data for Dares Salam

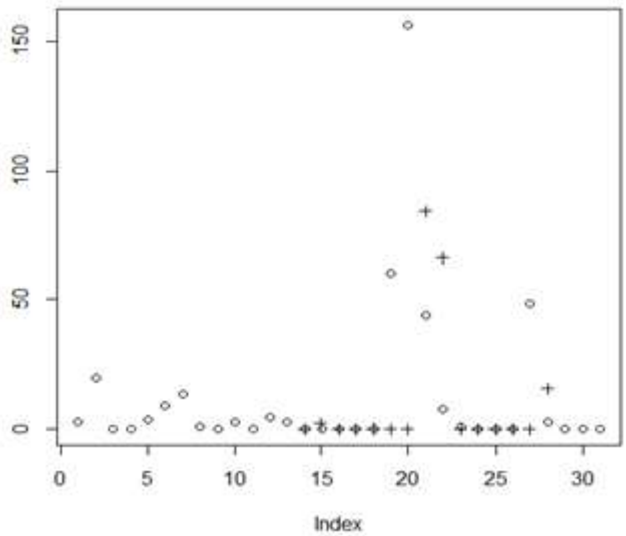

Figure 2: WRF outputs of the experiment of rainfall (in mm) for Dar es Salam for 15 days (14th to 28th)in December 2011 indicated by "+". Values indicated by "o" correspond to the observed daily (24h) rainfall for the entire month.

2.3 Bayesian Modelling

Suppose that the data $\boldsymbol{x} = (x_1, x_2, \ldots x_n)$ are independent realizations of a random variable whose density $f(.)$ falls within parametric family$\{f(x|\theta): \theta \in \Theta\}$. The likelihood function is defined as
$L(\theta; \boldsymbol{x}) = \prod_{i=1}^{n} f(x_i|\theta)$.

Usually it is easier to work with log-likelihood function$l(\theta; \boldsymbol{x}) = log\{L(\theta; \boldsymbol{x})\}$. In classical framework (i.e. Frequentist paradigm) parameters are assumed to be constant. However, in Bayesian framework, parameters are assumed to be variable with a certain distribution. More details on Bayesian framework on extreme value analysis can be found, for instance, in [8], [9],[10].

In Bayes Theorem we assume that, without reference to the data, it is possible to formulate beliefs about θ that can be expressed as a probability distribution.

This distribution is known as prior distribution. Let $\pi(\theta)$ denote the density of the prior distribution of θ. Bayes' theorem can be used to combine likelihood and prior information to give posterior distribution as follows:

$$\pi(\theta|\boldsymbol{x}) = \frac{\pi(\theta)L(\theta;x)}{\int_{\Theta}\pi(\theta)L(\theta;x)d\theta} \propto \pi(\theta)L(\theta;\boldsymbol{x}), \quad (1)$$

where$\pi(\theta|x)$ is the density of posterior distribution.

The primary objective of an extreme value analysis is usually prediction. Let y denote a future observation with density function$f(y|\boldsymbol{\theta})$, where $\theta \in \Theta$. The posterior predictive density of y, given observed data $\boldsymbol{x}$, is

$$f(y|\boldsymbol{x}) = \int_{\Theta} f(y|\theta)\pi(\theta|\boldsymbol{x})d\theta$$

So, if a suitable prior can be specified Bayesian framework can be used for prediction purpose. The difficulty in using Bayes' procedure is the computation of the integral in (1). This problem can be overcome by using simulation based techniques such as Markov Chain Monte Carlo (MCMC) to simulate realizations of the posterior distribution.

Apart from prior distribution specification, equation (1) requires the likelihood function that can be obtained from the governing distribution of the observed processes, in this case the extreme behaviour. In the following section, the likelihoods based on different characterizations (that are used in this study) of extreme behaviour are outlined.

2.4 Likelihoods for Extremes

Let $\boldsymbol{\theta} = (\mu, \sigma, \xi)$where $(\mu, \sigma, \xi)$are the location, scale and shape parameters; suppose the data $\boldsymbol{x} = (x_1, x_2, \dots x_n)$ are independent realizations of a random variable from extreme value distribution. Then the following outlines the likelihood functions that are used in study.

2.4.1 Generalized Extreme Value Distribution

Generalized Extreme Value (GEV) distribution has distribution function given by

$$F(z) = exp\left\{-[1+\xi(z-\mu)/\sigma]_+^{-1/\xi}\right\}, \quad (2)$$

Where $\sigma > 0$ and $h_+ = \max(h, 0)$. The case of $\xi = 0$ the distribution converges to Gumbel distribution.

The log-likelihood for GEV($\boldsymbol{\theta}$) is given by

$$l(\boldsymbol{\theta};\boldsymbol{x}) = -n\log\sigma - \left(1+\frac{1}{\xi}\right)\sum_{i=1}^{n}\log\left\{1+\frac{\xi(x_i-\mu)}{\sigma}\right\} - \sum_{i=1}^{n}\left\{1+\frac{\xi(x_i-\mu)}{\sigma}\right\}^{-\frac{1}{\xi}},$$

Provided $1+\xi(x_i-\mu)/\sigma$ is positive for each i=1,..,n.

2.4.2 Point Process Characterization

Suppose $M_n = \max(X_1, \dots, X_n)$ can be approximatedby the GEV($\boldsymbol{\theta}$) with possibly end points z_- and z_+. For large $u > z_-$ the sequence $\{X_1, X_2, \dots, X_n\}$ viewed on the interval (u, z_+) is approximately a non-homogeneous Poisson process with intensity function

$$\lambda_{\boldsymbol{\theta}}(x) = \frac{1}{\sigma}\left\{1+\xi\left(\frac{x-\mu}{\sigma}\right)\right\}^{-\frac{(\xi+1)}{\xi}}, \quad u < x < z_+,$$

where$\sigma > 0$ and $\boldsymbol{\theta} = (\mu, \sigma, \xi)$. The intensity measure on (u, z_+) is therefore given by

$$\Lambda_{\boldsymbol{\theta}}(u, z_+) = \int_u^{z_+}\lambda_{\boldsymbol{\theta}}(x)dx = \left\{1+\xi\left(\frac{u-\mu}{\sigma}\right)\right\}^{-\frac{1}{\xi}}.$$

Suppose that n_u of n observations exceed the threshold u. Let $x_{(i)}$ denotes the ith exceedance. The log-likelihood is derived as

$$l(\boldsymbol{\theta};\boldsymbol{x}) = n_y\Lambda_{\boldsymbol{\theta}}(u, z_+) + \sum_{i=1}^{n_u} log\{\lambda_{\boldsymbol{\theta}}(x_{(i)})\},$$

provided$1+\xi(u-\mu)/\sigma$ and $1+\xi(x_{(i)}-\mu)/\sigma$ are positive for each i=1,..,n. $n_y$is the number of periods of observation, and the maxima over those periods are distributed as GEV($\boldsymbol{\theta}$). If n_y is the number of years of observation, the annual maxima are distributed as GEV($\boldsymbol{\theta}$). This approximation assumes that there are a large number of observations within each period. Mathematical details of these results can be obtained in [9],[11] ,[12].

2.5 Construction of Prior Distribution from WRF outputs

The likelihoods outlined in sections 2.4.1 and 2.4.2 are all function of the parameter vector $\boldsymbol{\theta} = (\mu, \sigma, \xi)$. The construction procedures of prior distribution on these parametersare the same and the quantiles approach used here is the same as in Coles and Tawn (1996). Since the aim of this paper is to construct prior from WRF outputs on rainfall, the construction of prior distribution in terms of quantiles is the only procedure illustrated here by using GEV distribution. Other procedures can be found, for instance, in [2], [3], [13].

Let $F(q_p) = 1 - p$where $F(.)$ is the GEV distribution function given in equation (2). It follows that

$$q_p = \mu + \frac{\sigma\left(x_p^{-\xi} - 1\right)}{\xi},$$

where $x_p = -\log(1-p)$. A prior distribution can be constructed in terms of quantiles $(q_{p_1}, q_{p_2}, q_{p_3})$ for specified probabilities $p_1 > p_2 > p_3$. It is easy to work with differences $(\tilde{q}_{p1}, \tilde{q}_{p2}, \tilde{q}_{p3})$, so that $\tilde{q}_{p_i} = q_{p_i} - q_{p_{i-1}}$ for $i = 1,2,3$, where q_{p_0} is the physical lower end point of the process variable. The quantile differences are assumed to be independent with gamma distribution i.e.

$$\tilde{q}_{p_i} \sim gamma(\alpha_i, \beta_i), \quad \alpha_i, \beta_i > 0,$$

for $i = 1,2,3$. This construction leads to the prior density

$$\pi(\boldsymbol{\theta}) \propto J\prod_{i=1}^{3}\tilde{q}_{p_i}^{\alpha_i-1}exp\left\{-\frac{\tilde{q}_{p_i}}{\beta_i}\right\},$$

provided that$q_{p_1} < q_{p_2} < q_{p_3}$. $J$is the Jacobian of the transformation from $(q_{p_1}, q_{p_2}, q_{p_3})$ to $\boldsymbol{\theta} = (\mu, \sigma, \xi)$, namely

$$J = \frac{\sigma}{\xi^2}\left|\sum_{\substack{i,j\in\{1,2,3\}\\ i<j}} (-1)^{i+j}\,(x_i x_j)^{-\xi}\log\left(\frac{x_j}{x_i}\right)\right|,$$

where $x_i = \log(1-p_i)$ for $i = 1,2,3$.
At $\xi = 0$ the prior distribution is defined by continuity, using

$$\lim_{\xi\to 0} q_{p_i} = \mu - \sigma\log x_i, \qquad i = 1,2,3,$$

and

$$\lim_{\xi\to 0} J = \frac{\sigma}{2}\left|\sum_{\substack{i,j\in\{1,2,3\}\\ i<j}} (-1)^{i+j}\log x_i\log x_j\log\left(\frac{x_j}{x_i}\right)\right|.$$

The hyperparameters α_i, β_i and p_i must be specified. The p_i's values are usually set at $p_i = 10^{-i}$ for $i = 1,2,3$ (It is also by default in the evdbayes package). In this study, the default values (p_1, p_2, p_3)= (0.1, 0.01, 0.001) are used.

To specify α_i, β_i for $i = 1,2,3$ from WRF outputs, three upper quantiles from the WRF outputs are used (from which the differences are obtained). For gamma random variable, if the mean and variance are known, the shape α_i and scale β_i parameters can be obtained (e.g. using igamma function in evdbayes R package). Let $\bar{q}_{p_i}$ and $v_{\tilde{q}_{p_i}}$ be respectively the mean and variance of a random variable $\tilde{q}_{p_i}$ where $\tilde{q}_{p_i}\sim gamma(\alpha_i,\beta_i)$. It is easily shown that $\alpha_i = \left(\bar{q}_{p_i}\right)^2/\left(v_{\tilde{q}_{p_i}}\right)$ and $\beta_i = \left(v_{\tilde{q}_{p_i}}\right)/\left(\bar{q}_{p_i}\right)$ for $i = 1,2,3$.

The Earth physical conditions of a given place for a given time interval are fixed. Simulation results based on these fixed conditions will be the same for a particular simulation method. Variations in the simulation output will depend on the variations on WRF model specifications. Since different specifications in the WRF model will possibly result into different outputs, in this study it is assumed that the simulated quantiles are the means of sample quantiles space for that WRF model. This assumption can be used for any simulation model. The remaining information required is the variances of the quantiles. To allow possible different specifications of the WRF model, as with informative priors, variances are set at high values. In this study, variances are set equal to 1000 for each quantile.

3. Results and Discussion

The first analysis is based on the observed annual maxima derived from daily (24h) rainfall data set described in section 2.2. It is assumed that these annual maxima are independent observations from the GEV distribution. Using the Maximum Likelihood Estimation (MLE) method [with the packages eXtremes (Gilleland and Katz, 2011) in R] we obtain the following estimates for the parameters $(\hat{\mu},\hat{\sigma},\hat{\xi}) = (68.52,\ 18.70, 0.111)$ with standard errors (2.807, 2.108, 0.092) for $\hat{\mu}$, $\hat{\sigma}$ and $\hat{\xi}$ respectively. Estimated 95% confidence intervals for each parameter are [63.014, 74.017] for μ, [14.57, 22.832] for σ and [-0.068, 0.291] for ξ (Table 1). Although the estimated shape parameter is positive, the estimated 95% confidence interval extends below zero showing uncertainty in estimating shape parameter. The return level plot in Figure 3 shows the curve to be approximately linear due to the estimate of ξ being close to zero.

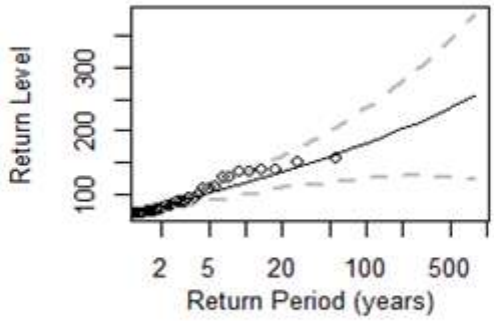

Figure 3: Return Level (in mm) plot using maximum likelihood estimates.

We now seek more reliable estimates based on the prior information using a Bayesian approach, through the package evdbayes under R (Stephenson A. and Ribatet M., 2012). As mentioned earlier, in this study we apply quantiles approach in eliciting the prior distribution but instead of using experts knowledge (which in most cases do not exists), quantiles from WRF outputs are used.

The WRF experiments were set to run at an interval of 3 minutes for 15 days. The outputs are converted to cumulative rainfall using an interval of 24 hours. This leads to have 15 records of cumulative daily rainfall. The resulting records are depicted together with gauge records on Figure 2. The three higher quantiles from the WRF outputs are 15.52084, 66.05064 and 84.49259. The associated differences are 15.52084, 50.52980, 18.44195. These differences are assumed to be the means of random variables associated with quantiles of the differences. To allow uncertainties associated with this assumption, large variances are set to these quantiles and are set to be 1000 in this study. Using these in *igamma* of *evdbayes* package gives $(\alpha_1,\alpha_2,\alpha_3)$=(0.241, 2.552, 0.339) and $(\beta_1,\beta_2,\beta_3)$=(64.433, 19.794, 54.348).

By using point process characterization and setting threshold equal to 50mm, these $(\alpha_1,\alpha_2,\alpha_3)$ and $(\beta_1,\beta_2,\beta_3)$ were applied in prior.quant and posterior functions of evdbayes package. Using MLEs as the initial vector $\theta = \left(\hat{\mu},\hat{\sigma},\hat{\xi}\right) = (68.52,\ 18.70, 0.111)$, and using proposal standard deviations psd = (2.8, 2.1, 0.09), a Markov chain Monte Carlo (MCMC) method was applied to generate samples from the posterior distribution (10000 iterations).The sample means of each marginal component of the chain are $\left(\hat{\mu},\hat{\sigma},\hat{\xi}\right)$=(72.731, 19.361, 0.052)with standard deviations 2.285, 1.406, 0.063 for $\hat{\mu}$, $\hat{\sigma}$ and $\hat{\xi}$ respectively. Estimated 95% confidence intervals for each parameter are [68.253, 77.210] for $\hat{\mu}$, [16.605, 22.117] for $\hat{\sigma}$ and [-0.071, 0.174] for $\hat{\xi}$ (These figures are rounded to 3 decimal places). The posterior return level plot in Figure 3 shows the curve to be approximately linear due to the estimate of ξ being close to zero.

Comparing the two methods (MLE and Bayesian), the point estimates are almost the same, but Bayesian method gives relatively small standard error which results into narrower 95% confidence intervals. The associated intervals are narrower by (19%, 33%, 32%) for (μ, σ, ξ) respectively. This highlights that Bayesian estimation method results into more reliable results compared to MLE for scarcity data. Similar result is obtained by Giugni(2013) who analyzed a 53 years dataset from the year 1958 to 2010. Giugni applied Bayesian approach with priors of parameters estimated using the normal distributions which are set at $\mu \sim N(0, 10^4)$, $\sigma \sim LN(0, 10^4)$ and $\xi \sim N(0.15, 0.2)$. A shape (ξ) parameter of 0.15 used by Giugni (also used in [17], [18], [19]) is the empirical evidence based on Europe and North America zones.

Comparing the quantile approach applied in this paper and non-informative prior with normal distribution set at at$\mu \sim N(0, 10^4)$, $\sigma \sim LN(0, 10^4)$ and $\xi \sim N(0, 10^4)$, the point estimates are almost the same, but the quantile approach with prior derived from WRF outputs gives relatively small standard errors. This results into narrower empirical 95% confidence intervals. The empirical 95% confidence intervals are narrower by (21%, 38%, 34%) for (μ, σ, ξ) respectively. This shows that by including informative priors (from WRF outputs) has improved reliability of estimates. Furthermore, comparing informative prior with normal distribution [the one used by Giugni (2013) applied to current dataset] set at $\mu \sim N(0, 10^4)$, $\sigma \sim LN(0, 10^4)$ and $\xi \sim N(0.15, 0.2)$ with quantile approach based on WRF outputs, the later gives almost the same point estimates but with relatively small standard errors. The informative prior $\xi \sim N(0.15, 0.2)$ gives the means (standard errors) estimates of 68.506(2.849) for$\hat{\mu}$, 19.675(2.334) for$\hat{\sigma}$ and 0.113 (0.093) for $\hat{\xi}$.

Table 1: A summary of results for the different methods of estimation in analysis of the Dar es Salaam data

Method		*MLE*	*Bayesian Quantiles from WRF*	*Bayesian Non-informative*
Estimates	μ (mm)	68.52	72.585	72.827
	σ(mm)	18.70	19.216	20.1373
	ξ(mm)	0.111	0.049	0.1149
95% intervals	μ(mm)	[63.014,74.017]	[67.920, 77.251]	[67.795, 77.859]
	σ(mm)	[14.57, 22.832]	[16.274, 22.158]	[16.385, 23.890]
	ξ(mm)	[-0.068, 0.291]	[-0.074, 0.172]	[-0.064, 0.294]

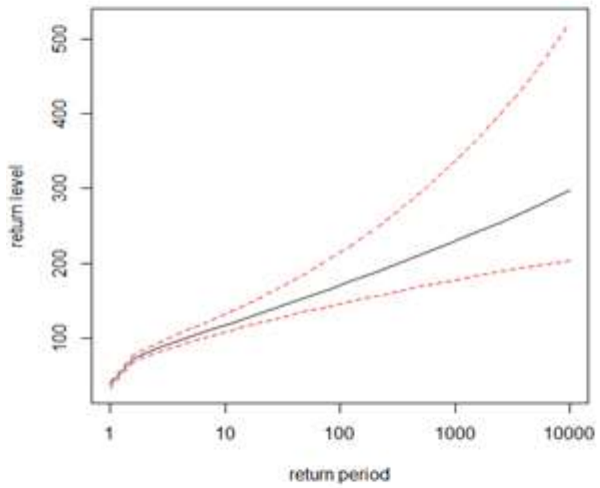

Figure 4: Posterior return level plot in Bayesian analysis of the Dar Es Salaam rain data. The curves represent medians (solid line) and intervals containing 95% of the posterior probability (dashed lines)

4. Conclusions

In this paper it is shown that WRF rainfall outputs can be used to construct informative priors in the Bayesian framework of extreme value rainfall analysis. This was achieved by considering high quantiles from WRF outputs as the means of the possible WRF outputs with different possible configurations for the given time interval and for the given place.

Distribution of annual maxima of rainfall is well modelled by GEV distribution [14]. The shape parameter is crucial in determining the characteristics of extreme value behaviour. Estimation of GEV parameters by methods such as maximum likelihood can be unreliable due to small length of rainfall records. Rainfall data of 54 years of Dar es Salaam, Tanzania was analyzed using both MLE and Bayesian methods. The shape parameter, based on MLE, was estimated as $\hat{\xi}$= 0.111 with a standard error of 0.092.

Bayesian method, constructing prior using WRF outputs, gives estimate of $\hat{\xi}$= 0.052 with a standard error of 0.063. The Bayesian based estimates with prior constructed using WRF rainfall outputs are in agreement with those obtained with prior information of $\xi \sim N(0.15, 0.2)$, but the later gives relatively higher standard error i.e.$\hat{\xi}$= 0.113 and standard error is 0.093.

5. Acknowledgements

I thank Tanzania Meteorological Agency for providing rainfall data through collaboration withDar es Salaam Institute of Technology (DIT). I also thank TriphoniaNgailo of DIT for availing data from the WRF experiments she did during her PhD studies.

References

[1] Ngailo T. J., Reuder J., Rutalebwa E., Nyimvua S. and Mesquita M. D (2016), "Modeling of Extreme Maximum Rainfall Using Extreme Value Theory for Tanzania", International Journal of Scientific and Innovative Mathematical Research (IJSIMR), Volume 4, No.3, pp. 34 – 45.
[2] Coles, S.G. and Powell, E.A. (1996). Bayesian methods in extreme value modelling: areview and new developments. Internat.Statist.Rev., 64, 119—136.
[3] Coles, S.G. and Tawn, J.A. (1996). A Bayesian Analysis of Extreme Rainfall Data.Appl.Statist., 45, 463—478.
[4] Coles, S.G. and Tawn, J.A. (1991). Modelling Extreme Multivariate Events. J. R. Statist.Soc., B, 53, 377—392.
[5] Ngailo T., Shaban N., Reuder J., Rutalebwa E. and Mugume I (2013), Non – homogeneous poison process modeling of seasonal extreme rainfall events in Tanzania, International journal of science and research (IJSR), Volume 5, No. 10, Pp. 1858 – 1868.
[6] Theodory John Kahwili (2017) Imperial Journal of Interdisciplinary Research (IJIR) Vol-3, Issue-6, 2017 ISSN: 2454-1362, [Online]: Availabehttp://www.onlinejournal.in(Accessed on October 3, 2017)
[7] https://reliefweb.int/report/united-republic-tanzania/tanzania-floods-dref-operation-n%C2%B0-mdrtz013-final-report(Accessed on October 4, 2017)
[8] Elizabeth Smith (2005). Bayesian Modelling of Extreme Rainfall Data.PhD Thesis, University of Newcastle upon Tyne.
[9] Coles,S.(2001).Anintroductiontostatisticalmodelingof extremevalues. Springer-Verlag
[10] Beirlant J., Goegebeur Y., Segers J., Teugels J. (2004). Statistics of extremes: Theory and Applications. Wiley.
[11] Pickands, J. (1971) The two–dimensional Poisson process and extremal processes. J. Appl. Probab., 8, 745–756.
[12] Smith, R. L. (1989) Extreme value analysis of environmental time series: an example based on ozone data (with discussion). Statist. Science, 4, 367–393.
[13] Stephenson, A. G. (2003) Multivariate Extreme Value Distributions and their Application. Ph.D. thesis, Lancaster University, Lancaster, UK.
[14] Giugni, Maurizio. (2013). Stationary Vs. Non-stationary of Extreme Rainfall in Dar Es Salaam (Tanzania). . 10.13140/2.1.3032.6724.
[15] IPCC (2007). Climate Change 2007: The Physical Science Basis, summary for policymakers. Contribution of Working Group I to the Fourth Assessment Report, Paris.
[16] Ngailo, J. Triphonia (2017). Statistical modelling of extreme rainfall variability in Tanzania. PhD thesis (unpublished), University of Dar es Salaam, Tanzania.
[17] Koutsoyiannis, D. (2004). Statistics of extremes and estimation of extreme rainfall: I. Theoretical investigation. Hydrological Sciences Journal, 49 (4), 575-590.
[18] Koutsoyiannis, D. (2004). Statistics of extremes and estimation of extreme rainfall: II. Empirical investigation of long rainfall records. Hydrological Sciences Journal, 49(4), 591-610.
[19] Garcia-Aristizabal, Alexander &Bucchignani, Edoardo& Palazzi, Elisa &D'Onofrio, Donatella &Gasparini, Paolo &marzocchi, Warner. (2015). Analysis of non-stationary climate-related extreme events considering climate change scenarios: an application for multi-hazard assessment in the Dar es Salaam region, Tanzania. Natural Hazards. 75. 289-320. 10.1007/s11069-014-1324-z.
[20] Chavez-Demoulin, V., and A. C. Davison (2012), Modelling time series extremes, REVSTAT Stat. J., 10, 109–133.

Author Profile

Edwin Rutalebwareceived the BSc(Ed) and MSc (Maths) degrees from university o Dar es Salaam (Tanzania) in 1996 and 2000, respectively. He also received the Msc(Statistics) and PhD degrees in 2004 and 2012 from Catholic University of Leuven (Belgium). Since 2000 up to date he is working at Dar es Salaam Institute of Technology (DIT), Tanzania as Lecturer of Engineering Mathematics and Statistics.