ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

CHARACTERIZATION AND FREQUENCY ANALYSIS OF ONE DAY ANNUAL MAXIMUM AND TWO TO FIVE CONSECUTIVE DAYS' MAXIMUM RAINFALL OF ACCRA, GHANA

Xeflide Seth Kwaku[1] and Ophori Duke[1]
[1]Department of Earth and Environmental Studies, Montclair State University, Upper Montclair, New Jersey, USA
E-mail: xeflides1@mail.montclair.edu

ABSTRACT

Annual one day maximum rainfall and two to five consecutive day's maximum rainfall corresponding to a return period of 2 to 100 years has been conducted for Accra, Ghana. Three commonly used probability distributions; normal, lognormal and gamma distribution have been tested to determine the best fit probability distribution that describes the annual one day maximum and two to five consecutive days' maximum rainfall series by comparing with the Chi-square value. The results revealed that the log-normal distribution was the best fit probability distribution for one day annual maximum as well as two to five consecutive days' maximum rainfall for the region. Based on the best fit probability distribution a maximum of 84.05 mm in 1 day, 91.60 mm in 2 days, 100.40 mm in 3 days, 105.67 mm in 4 days and 109.47 mm in 5 days is expected to occur at Accra every two years. Similarly a maximum rainfall of 230.97 mm, 240.49, 272.77 mm, 292.07 mm and 296.54 mm is expected to occur in 1 day, 2, 3, 4 and 5 days respectively every 100 years. The results from the study could be used as a rough guide by engineers and hydrologists during the design and construction of drainage systems in the Accra metropolis as poor drainage has been identified as one of the major factors causing flooding in Accra.

Keywords: return period, frequency, probability distribution, gamma, consecutive days maximum rainfall.

INTRODUCTION

The procedure for estimating the frequency of occurrence (return period) of a hydrological event such as flood is known as (flood) frequency analysis. Though the nature of most hydrological events (such as rainfall) is erratic and varies with time and space, it is commonly possible to predict return periods using various probability distributions (Upadhaya and Singh, 1998). Flood frequency analysis was developed as a statistical tool to help engineers, hydrologists, and watershed managers to deal with this uncertainty. Flood frequency is utilized to determine how often a storm of a given magnitude would occur. It is an important tool for the building and design of the safest possible structures (e.g. dams, bridges, culverts, drainage systems etc.) because the design of such structures demands knowledge of the likely floods which the structure would have to withstand during its estimated economic useful life (Bruce and Clark, 1966).

In particular, analysis of annual one day maximum rainfall and consecutive maximum days rainfall of different return periods (typically 2 to 100 years) is a basic tool for safe and economic planning and design of small dams, bridges, culverts, irrigation and drainage work as well as for determining drainage coefficients (Bhakar *et al.*, 2006).

The devastating flooding in Accra in recent times among other things has been attributed to poor drainage (Twumasi and Asomani-Boateng, 2002). Even though some studies have been carried out to map out areas within Accra that are susceptible to flooding (e.g. Nyarko, 2002; Twumasi and Asomani-Boateng, 2002) nothing is known of the frequency of occurrence of such events. In this present study, annual 1 day and 2 to 5 consecutive days maximum rainfall data of Accra (Accra airport station) have been analytically fitted with three theoretical probability distribution functions, viz., normal, lognormal, and gamma distribution. Frequency analysis (corresponding to 2 to 100 years) of annual 1 day and 2 to 5 consecutive days maximum rainfall has also been carried out using the best fitted probability distribution.

MATERIALS AND METHODS

The daily rainfall data recorded at the Accra airport station (05°36°N latitude, 00°10°W longitude and 66.7 m above sea level) for a period of 30 years (1975-2004 inclusive) were used for this analysis. Annual 2 to 5 days consecutive days rainfall were computed using the method described by Bhakar *et al.* (2006), by summing up rainfall of corresponding previous days. Maximum amount of annual 1 day to 2 to 5 days consecutive days rainfall for each year was used for the analysis. Statistical parameters of annual 1 day as well as consecutive days maximum rainfall have been computed and are shown in Table-1.One day to five days maximum rainfall data were fitted with three main probability distributions (Table-2).

Testing the goodness of fit of probability distribution

For the purpose of prediction, it is usually required to understand the shape of the underlying distribution of the population. To determine the underlying distribution, it is a common practice to fit the observed distribution to a theoretical distribution. This is done by comparing the observed frequencies in the data to the expected frequencies of the theoretical distribution since certain types of variables follow specific distribution (Tilahun, 2006).

One of the most commonly used tests for testing the goodness of fit of empirical data to specify theoretical frequency distribution is the chi-square test (Haan, 1977). In applying the Chi-square goodness-of-fit test, the data are grouped into suitable frequency classes. The test compares the actual number of observations and the expected number of observations (expected values are calculated based on the distribution under consideration) that fall in the class intervals. The expected numbers are calculated by multiplying the expected relative frequency by the total number of observation. The sample value of the relative frequency of interval i is computed in accordance with equation 1 (Bhakar *et al.*, 2006) as:

$$f_s(x_i) = \frac{n_i}{n} \quad \ldots\ldots(1)$$

The expected relative frequency in a class interval i can also be approximated using equation 2

$$f_{x_i} = \Delta x_i \, p(x_i) \quad \ldots\ldots(2)$$

The Chi-square test statistic is computed from the relationship

$$\chi_c^{\,2} = \sum_{i=1}^{m} (O_i - E_i)^2 / E_i \quad \ldots\ldots(3)$$

O_i is the observed and E_i is the expected (based on the probability distribution being tested) number of observation in the i^th^ class interval. The observed number of observation in the i^th^ interval is computed from equation 1 as:

$nf(x_i) = n_i$. Similarly, nf_{x_i} is the corresponding expected numbers of occurrences in interval i.

The distribution of $\chi_c^{\,2}$ is the chi-square distribution with $\nu = k - l - 1$ degrees of freedom. In conducting the goodness of fit test using the chi-square test , a confidence level, often expressed as $1-\alpha$, is chosen (where α is referred to as the significance level). Typically, 95% is chosen as the confidence limit. The test poses the null hypothesis (H_o) that the proposed probability distribution is from the specified distribution. H_o is rejected if $\chi_c^{\,2} > \chi^2_{1-\alpha,\nu}$. The value of $\chi^2_{1-\alpha,\nu}$ is determined from published χ^2 tables with v degrees of freedom at the 5% level of significance.

In this study three commonly used probability distributions were fitted with 1 day and 2 to 5 days consecutive days' maximum rainfall. The three distributions are briefly discussed below.

Normal distribution

The normal distribution, a two parameter distribution, has been identified as the most important distribution of continuous variables applied to symmetrically distributed data (Tilahun, 2006). The probability density function is given by

$$n(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] \quad \text{for}$$
$$-\infty \le x \le \infty \quad \ldots\ldots(4)$$

Log normal distribution

Most hydrological continuous random variables are noted to be asymmetrically distributed hence it is computationally advantageous to transform the distribution to a lognormal distribution (Tilahun, 2006). The transformation is usually achieved by taking the logs of the variables in question. A random variable x is said to follow a lognormal distribution if the logarithm (usually natural logarithm) of x is normally distributed. The probability density function of such a variable $y = \ln x$:

$$y = \frac{1}{\sigma_y\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{y-\mu_y}{\sigma_y}\right)\right] \quad \text{for } 0 \le x \le \infty \quad ..(5)$$

Gamma distribution

In hydrology the gamma distribution has an advantage of having only positive values (Tilahun, 2006). The probability density function of a gamma distributed random variable x is given by

$$f(x) = \frac{1}{\beta^{\alpha}\Gamma_{\alpha}} \alpha^{x-1} e^{-x/\beta} \quad \text{for } 0 \le x \le \infty \quad \ldots\ldots(6)$$

The gamma function is defined such that the total area under the density function is unity as:

$$\Gamma_{\alpha} = \int_0^{\infty} x^{\alpha-1} e^{-x}\,dx \quad \ldots\ldots(7)$$

Frequency analysis and frequency factors

Values of 1 day, 2, 3, 4 and 5 consecutive days maximum rainfall can be estimated statistically through the use of the Chow (1951) general frequency formula. The formula expresses the frequency of occurrence of an event in terms of a frequency factor, K_T, which depends upon the distribution of particular event investigated. Chow (1951) has shown that many frequency analyses can be reduced to the form

$$X_T = \overline{X}\,(1 + C_V K_T) \quad \ldots\ldots(8)$$

For the normal distribution the frequency factor can be expressed by equation 10 (which is the same as the standard normal variable z).

$$K_T = \left(\frac{x_T - \mu}{\sigma}\right) \quad \ldots\ldots(9)$$

The value of z could be found from standard normal density function tables or could be calculated from equation 10 as below:

ARPN Journal of Engineering and Applied Sciences

z = w-

$$\left[\frac{2.515517+0.802853w+0.0100.010328w^2}{1+1.432788+0.189269w^2+0.001308w^3}\right] \text{.........(10)}$$

where

$$w=\left[In\left(\frac{1}{p^2}\right)\right]^2 \text{.........(11)}$$

$1-p$ is substituted for p in equation 11 when $p>0.5$ .The value of z in this case is given a negative sign (Bhakar *et al.*, 2006) .

The equation for the parameters in terms of the sample moments for the normal distribution is given by

$$\mu=\bar{x}\,,\sigma=S_x \text{.........(12)}$$

For the lognormal distribution it is assumed that Y=In X is normally distributed. The magnitude of an event having a return period T , X_T is obtained from the relation $X_T=\exp(Y_T)$ where

$$Y_T=\overline{Y}(1+C_{vY}K_T) \text{.........(13)}$$

and

$$K_T=\left(\frac{y_T-\mu_y}{\sigma_y}\right) \text{.........(14)}$$

The value of K_T can be computed using equation 10 or found from the standard normal distribution table.

The equation for the parameters in terms of the sample moment for the lognormal distribution is given by

$$\mu_y=y\,,\,\sigma_y=S_y \text{.........(15)}$$

In the case of the gamma distribution frequency analysis can be done using the method of moments as described by Haan (1977). The equation for the parameters in terms of the sample moment is given by

$$\alpha=\frac{\bar{x}}{{s_x}^2}\,,\;\beta=\left(\frac{\bar{x}}{s_x}\right) \text{.........(16)}$$

RESULTS AND DISCUSSION

The data as presented in Table-2 revealed that the computed Chi-square values for the three probability distributions were less than the critical Chi-square value at the 95% confidence level for 1 day as well as consecutive days maximum rainfall series (except the 2 consecutive days maximum rainfall for the normal distribution). Results of the Chi-square values for the different distribution as presented in Table-2 indicated that the lognormal distribution gave the minimum χ^2 value for all daily rainfall series analyzed. The lognormal distribution function is deemed the best fit function for 1day and 2 to 5 consecutive days maximum rainfall in the study region. Table-3 gives the annual 1 day and consecutive days maximum rainfall for different return periods as determined by the selected best fit distribution. The result show that a maximum of 84.05 mm in 1 day, 91.60 mm in 2 days, 100.40 mm in 3 days, 105.67 mm in 4 days and 109.47 mm in 5 days is expected to occur at Accra (Airport and surrounding areas) every two years. Similarly a maximum rainfall of 230.97 mm, 240.49mm, 272.77 mm, 292.07 mm and 296.54 mm expected to occur in 1 day, 2, 3, 4 and 5 days respectively every 100 years.

It was recommended by Bhakar *et al.* (2006) that 2 to 100 years is a sufficient return period for soil and water conservation measures, construction of dams, irrigation and drainage works. The 2 to 100 years return period obtained in this study could be used as a rough guide during the construction of such similar structures. In particular, these values could be very beneficial during the construction of drainage systems in the Accra metropolis as poor drainage has been identified as one of the major factors causing flooding in the area.

Table-1. Summary statistics of annual 1 day as well as consecutive days maximum rainfall.

Statistical parameters	1 day	2 day	3 day	4 days	5 days
Minimum (mm)	46.3	51	53.8	57.1	57.1
Maximum (mm)	243.9	248.6	248.6	249.6	252.8
Mean (mm)	92.35	99.81	110.07	116.62	119.96
Standard Deviation (mm)	42.03	43.19	49.47	53.22	53.74
Coefficient of Variation	45.51	43.27	44.94	45.64	44.80
Coefficient of skewness	1.93	1.57	1.13	1.10	1.22
Kurtoses	4.74	3.46	0.96	0.77	0.97

Table-2. Chi-square value for the three different distributions.

Consecutive days	Normal	Log normal	Gamma	Degrees of freedom	Critical Chi square value
One day	6.119128129	1.828190962	3.955516699	4	9.49
Two days	11.16695271	4.869751857	7.43802397	5	11.1
Three days	8.177292501	4.784323517	6.556892752	5	11.1
Four days	7.727966552	4.601362587	5.044697053	5	11.1
Five days	9.181824763	2.680704307	3.192649825	6	12.6

Table-3. 1 day as well as consecutive days' maximum rainfall for various return periods.

Return Period	1 day	2 days	3 days	4 days	5 days
2	84.05	91.60	100.40	105.67	109.47
5	121.54	130.27	144.57	153.12	157.47
10	147.10	156.31	174.60	185.52	190.07
20	171.60	181.08	203.32	216.61	221.23
50	205.44	215.04	242.94	259.61	264.20
100	230.97	240.49	272.77	292.07	296.54

ACKNOWLEDGEMENT

This study was funded by the College of Science and Mathematics at the Montclair State University, Upper Montclair New Jersey, USA. The authors are grateful to the Ghana Meteorological Services Department for providing the daily rainfall data.

REFERENCES

Bhakar, S.R., Bansal, A.N. Chhajed, N., and Purohit, R.C. 2006. Frequency analysis of consecutive days maximum rainfall at Banswara, Rajasthan, India. ARPN Journal of Engineering and Applied Sciences. 1(3): 64-67.

Bruce, J.P., and Clark, P.H. 1988. Introduction to hydrometeorology. Pergamon Press, Oxford.

Chow, V.T. 1951. A generalized formula for hydrologic frequency analysis. Transactions American Geophysical Union. 32(2): 231-237.

Chow, V.T. (ed). 1964. Handbook of applied hydrology. McGraw-hill, New York.

Nyarko, K.B. 2002. Application of a rational model in GIS for flood risk Assessment in Accra, Ghana. Journal of Spatial Hydrology. 2(1): 1-14.

Tilahun K. 2006. The characterization of rainfall in the arid and semi-arid regions of Ethiopia. Water S.A. 32(3): 429-436.

Twumasi, Y.A.; Asomani-Boateng, R. 2002. Mapping seasonal hazards for flood management in Accra, Ghana using GIS. Geoscience and Remote Sensing Symposium. IGARSS apos; 02. IEEE International. Vol. 5, pp. 2874-2876.

Upadhaya, A., and Singh, S.R. 1998. Estimation of consecutive days Maximum Rainfall by Various Methods and their Comparison. Indian Journal of S. Cons. 26(2): 193-2001.

ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

List of symbols

n_i : number of observation in the i^{th} interval
n : number of observation
f_{x_i} : expected relative theoretical frequency of the i^{th} class
$p(x_i)$: probability density function of a random variable (in the ith class interval).
Δx_i : mid point of the i^{th} class interval.
m : number of class intervals
z : standard normal variable
σ : standard deviation of the sample
μ_y : mean of ln x
σ_y : standard deviation of ln x
l : number of parameters used in fitting the proposed distribution
α : scale parameter of the gamma distribution
β : the shape parameter of the gamma distribution
Γ_α : the gamma function.
$\overline{X}$: mean value of X
$\overline{Y}$: mean of Y
C_V : coefficient of variation for original data , X
C_{vY} : coefficient of variation for ln X
X_T : event having a return period T
K_T : frequency magnitude of an factor.