Journal of Development and Agricultural Economics Vol. 1(5), pp. 121-126, August, 2009
Available online at http://www.academicjournals.org/JDAE

Full Length Research Paper

Characterization and frequency analysis of consecutive days maximum rainfall at Boalia, Rajshahi and Bangladesh

M. A. B. Barkotulla[1]*, M. S. Rahman[2] and M. M. Rahman[3]

[1]Department of Crop Science and Technology, University of Rajshahi, India.
[2]Department of Statistics, University of Rajshahi, India.
[3]Department of Civil Engineering, Rajshahi University of Engineering and Technology, India.

Accepted 21 July, 2009

Probability distributions were to predict rainfall status of various return period estimating one day and two to seven consecutive days annual maximum rainfall of Boalia, Rajshahi, Bangladesh. Three commonly used probability distributions (viz: Normal, Log Normal and Gamma distribution) were tested to determine the best fit probability distribution using the comparison of chi-square values. Results showed that the log-normal distribution was the best fit probability distribution for one day and two to seven consecutive days annual maximum rainfall for the region. Based on the best fit probability distribution the maximum rainfall of 116.15, 161.09, 190.14, 205.96, 220.37, 234.66 and 245.21 mm was expected to occur 1, 2, 3, 4, 5, 6 and 7 days respectively at Boalia (Rajshahi City Corporation and surrounding areas) every two years. For a recurrence interval of 100 years, the maximum rainfall expected in 1, 2, 3, 4, 5, 6 and 7 days were 290.24, 406.49, 544.08, 558.56, 600.33, 631.28 and 633.89 mm respectively. The results of this study would be useful for agricultural scientists, decision makers, policy planners and researchers in order to identify the areas where agricultural development and construction of drainage systems in Boalia as one of the major factors causing flooding should be focused as a long-term environmental strategy for Bangladesh.

Key words: Return period, frequency, probability distribution, consecutive days' maximum rainfall.

INTRODUCTION

The rainfall in Bangladesh varies, depending upon sea-son and location. Rainfall is comparatively less in this area than the other parts of the country. The average annual rainfall is about 1500 mm, which mainly occurs during the monsoon season. Thus this region has already been known as drought prone area of the country. In some places in Rajshahi districts the maximum temperature in summer season rises up to 40°C or more. Again in the winter the temperature even falls at 5°C at the area.

Most of the hydrological events occurring as natural phenomena are observed only once. One of the important problem in hydrology deals with interpreting past re-records of hydrologic events in terms of future probabilities of occurrence. The procedure for estimating frequency of occurrence of a hydrological event is known as frequency analysis. Though the rainfall is erratic and varies with time and space, it is commonly possible to predict return periods using various probability distribu-tions (Upadhaya and Singh, 1998).

Rainfall is one of the important hydrologic variable for which historical data are available. This helps in the probability based analysis of various aspects of the rainfall data. The different aspects of the rainfall are its intensity, daily, seasonal or annual totals, onset of mon-soon, occurrence of the consecutive non-rainy days etc. Each of these is relevant to different activities in the agricultural production process such as sowing, irrigation, drainage etc. Rainfall intensity is needed to estimate the peak rate of run-off (Schwab et al., 1981); knowledge of onset of effective monsoon is useful for planning land preparation and sowing and daily or moving total rainfall

*Corresponding author. E-mail: barkotru@yahoo.com. Tel: 9997140721-750041-49, 088-01716961119.

is required for drainage system design (Ashok Raj, 1979). Bhattacharya and Sarkar (1982) reported that estimating the drainage coefficient of agricultural crops, one needs to know the total rainfall over duration of crop tolerance period. Normally, the tolerance period of commercially grown crops vary from one day (for pulses) to six days (for rice). If the crops remain waterlogged for more days, these show signs of irreversible damage, resulting in low yield. The analysis becomes comparatively much easier with one day rainfall data. Therefore, it is required to use two or more consecutive days of rainfall, which can be done expeditiously if the rainfall for the desired consecutive day could be predicted with a reasonable accuracy from one day rainfall values. In particular, analysis of annual one day maximum rainfall and consecutive maximum days rainfall of different return periods (typically 2 - 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).

There is no widely accepted procedure to forecast the one day maximum rainfall. However, a hydrological frequency analysis has an application for predicting the future events on probability basis/return period. Frequency analysis of rainfall data has been attempted for different places in India (Dabralr and Pandey, 2008; Rizvi et al., 2001 and Mohanty et al., 1999). In this present study, the objectives determine the statistical parameters and prediction of annual one day and two to seven consecutive days maximum rainfall data of Boalia at various return periods using three probability distribution functions, viz., normal, lognormal, and gamma distribution.

MATERIALS AND METHODS

Study area and collection of data

Rajshahi municipality was established in 1876 and the whole municipal area was under Boalia thana. Rajshahi municipality was turned into City Corporation in 1987 and at present, it consists of four thanas, namely Boalia, Shah Makhdum, Motihar and Rajpara. Rajshahi town stands on the bank of the river Padma/Gangas (Figure 1).

The daily rainfall data recorded for the period of 15 years (1994 - 2008) at the Boalia station (24°21' - 24°24' N latitude, 88°35' - 88°39'E longitude) were obtained from Bangladesh Water Development Board, Rajshahi, Bangladesh for the purpose of this analysis. The daily data, in a particular year, 2 - 7 days consecutive days rainfall were computed by summing up rainfall of corresponding previous days. Maximum amount of annual 1 day and 2 - 7 consecutive days rainfall for each year was used for the analysis.

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, 1994). 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:

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

Where, n_i is the number of observations in the ith interval and n is the total number of observations.

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

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

Where, f_{x_i} is the expected relative theoretical frequency of the ith class, $p(x)$ is the: probability density functions of a random variable (in the ith class interval) and Δx_i is the mid point of the ith class interval.

The Chi-square test statistic is computed from the relationship

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

O_i is the observed and E_i is the expected (based on the probability distribution being tested) number of observation in the ith class interval. The observed number of observation in the ith interval is computed from equation 1 as: $nf_s(x_i) = n_i$

Similarly, nf_{x_i} is the corresponding expected numbers of occurrences in interval i.

The distribution of χ^2 is the chi-square distribution with $v = m - l - 1$ degrees of freedom. Where, m is the number of class intervals and l is the number of parameters used in fitting the proposed distribution.

In conducting the goodness of fit using the chi-square test, a confidence level, often expressed as 1 - α , is chosen (where α is referred to as the significance level). Typically, 95% is chosen as the confidence limit. Test the null hypothesis (Ho) that the proposed probability distribution is from the specified distribution. Ho is rejected if $\chi_c^2 > \chi^2_{1-\alpha, v}$. The value of $\chi^2_{1-\alpha, v}$ 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 7 days consecutive days' maximum rainfall. The three distributions

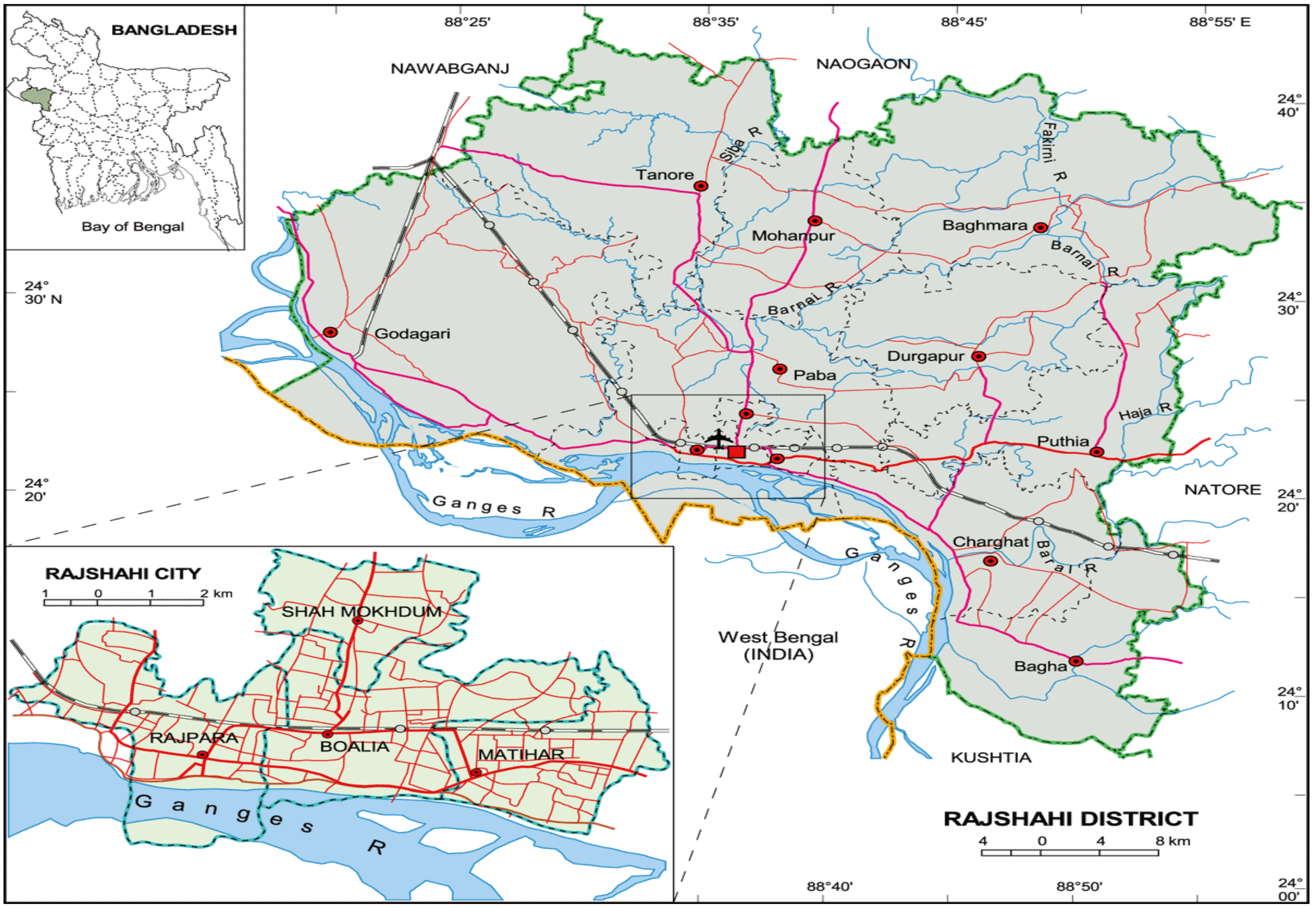

Figure 1. Map of study area of Rajshahi district in Bangaladesh.

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. The probability density function is given by:

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

Where, σ is the standard deviation and μ is the mean of the sample.

Log normal distribution

A random variable *x* is said to follow a lognormal distribution if the logarithm (usually natural logarithm) of is normally distributed. The probability density functions of such a variable $y = \ln x$:

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

Where, σ_y is the standard deviation and μ_y is the mean of $y = \ln x$

Gamma distribution

The probability density function of a gamma distributed random variable is given by *x*

$$f(x) = \frac{1}{\beta^{\alpha}\Gamma\alpha} x^{\alpha-1} e^{-x/\beta} \quad \text{for} \quad 0 \leq x \leq \infty \qquad (6)$$

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

$$\Gamma\alpha = \int_0^{\infty} x^{\alpha-1} e^{-x} dx \qquad (7)$$

Where, β is the scale parameter, α is the shape parameter of the gamma distribution and $\Gamma\alpha$ is the gamma function.

Frequency analysis and frequency factors

Values of 1 day, 2, 3, 4, 5, 6 and 7 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

$X_T = \overline{X}(1 + C_V K_T)$ Where, $\overline{X}$ is the mean, C_V is the coefficient of variation, is the frequency magnitude of a factor and X_T is the event having a return period T.

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

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

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

$$Z = W - \left[\frac{2.515517 + 0.802853W + 0.010328W^2}{1 + 1.432788W + 0.189269W^2 + 0.001308W^3}\right] \quad (10)$$

Where $W = \left[\ln\left(\frac{1}{p^2}\right)\right]^{1/2}$ (11)

$1 - p$ is substituted for p in equation (11) when $p \geq 0.5$. 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

$$\mu = \overline{x}, \sigma = S_x \quad (12)$$

For the lognormal distribution it is assumed that Y = ln X is normally distributed. The magnitude of an event having a return period T, X_T is obtained from the relation

$$X_T = \exp(Y_T) \quad (13)$$

$$Y_T = \overline{Y}(1 + C_{VY} K_T) \quad (14)$$

Where, $\overline{Y}$ is the mean and C_{VY} is the coefficient of variation of Y.

and $$K_T = \left(\frac{y_T - \mu_y}{\sigma_y}\right) \quad (15)$$

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:

$$\mu_y = \overline{y}, \quad \sigma_y = S_y \quad (16)$$

In the case of the gamma distribution frequency analysis can be done using the method of moments as described by Haan (1994). The equation for the parameters in terms of the sample moment is given by:

$$\beta = \frac{s_x^2}{\overline{x}}, \alpha = \left(\frac{\overline{x}}{s_x}\right)^2 \quad (17)$$

RESULTS AND DISCUSSION

Statistical parameters of annual 1 day as well as consecutive days maximum rainfall were computed (Table 1). The maximum rainfall found in monsoon season (June - October). One day to seven days maximum rainfall data were fitted with three main probability distributions.

The data presented in Table 2 found that the computed Chi-square values for three probability distribution that is normal, log-normal and gamma were found to be less than the critical value of Chi-square at 95% confidence level for 1 day and 2 - 7 consecutive days maximum rainfall series. Results in Table 3 indicated that the Chi-square values of lognormal distribution were minimum than normal and gamma distribution. So the log-normal distribution function was found to be the best fitted function for 1 day and 2 - 7 consecutive days maximum rainfall in the study region.

Table 4 showed that the annual 1day and 2 to 7 consecutive day's maximum rainfall for different return periods as determined by the selected best fitted distribution. The result showed that a maximum of 116.15 mm in 1 day, 161.09 mm in 2 days, 190.14 mm in 3 days, 205.96 mm in 4 days, 220.37 mm in 5 days 234.66 mm in 6 days and 245.21 mm in 7 days was expected to occur at Boalia (Rajshahi City Corporation and surrounding areas) every two years. For a recurrence interval of 100 years, the maximum rainfall expected in 1, 2, 3, 4, 5, 6 and 7 days was 290.24, 406.49, 544.08, 558.56, 600.33, 631.28 and 633.89 mm respectively.

Table 1. Statistical parameters of annual one to seven consecutive days maximum rainfall.

Parameters	1 day	2 days	3 days	4 days	5 days	6 days	7 days
Minimum (mm)	70	97.5	100	107	120.2	132	143.5
Maximum (mm)	245	356	418	443	459	466	469
Mean (mm)	125.22	174.1	209.76	225.14	241.04	255.82	265.69
Standard Deviation (mm)	52.62	75.39	99.59	102.72	109.55	112.60	113.56
Coefficient of Variation	42.02	43.30	47.48	45.62	45.45	44.02	42.74
α	5.66	5.33	4.44	4.80	4.84	5.16	5.47
β	22.11	32.64	47.29	46.86	49.79	49.56	48.54
Coefficient of Skewness	0.56	0.56	0.94	1.17	1.15	-1.92	-1.93
Coefficient of Kurtosis	0.59	0.59	-0.90	-0.70	-0.72	3.70	3.75

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

Rainfall (maximum)	Normal	Lognormal	Gamma
1 Day	1.2	0.1429	0.1429
2 Day	0.1714	0.02428	0.0268
3 Days	1.071	0.1721	0.32359
4 Days	2.336	0.965	1.1835
5 Days	1.877	0.6	0.771
6 Days	0.867	0.1709	0.2678
7 Days	0.8678	0.17094	0.32475

Table 3. One to seven consecutive days of annual maximum rainfall for various return periods.

Return Period	1 day	2 days	3 days	4 days	5 days	6 days	7 days
2	116.15	161.09	190.14	205.96	220.37	234.66	245.21
10	192.36	268.23	339.30	356.83	382.74	404.75	413.77
20	221.95	309.96	399.87	417.01	447.61	472.42	479.95
25	231.39	323.29	419.46	436.37	468.49	494.18	501.14
50	260.71	364.72	481.03	496.94	533.83	562.17	567.14
100	290.24	406.49	544.08	558.56	600.33	631.28	633.89

Table 4. Return periods and probability level.

T (Return period) years	2	10	20	25	50	100
P (Probability level) %	50	10	5	4	2	1

Bhakar et al. (2006), recommended that 2 - 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 Boalia (Rajshahi City Corporation) as poor drainage has been identified as one of the major factors causing flooding in the area.

REFERENCES

Ashok Raj PC (1979). Onset of Effective Monsoon and Critical Dry Spells. IARI Research Bulletin No.11, p 196, New Delhi.

Bhakar SR, Bansal AN, Chhajed N, Purohit RC (2006). Frequency analysis of consecutive days maximum rainfall at Banswara, Rajasthan India. ARPN J. Eng. Appl. Sci. 1(3): 64-67.

Bhattacharya AK and Sarkar TK (1982). Analysis of Rainfall Data for Agricultural Land Drainage Design. J. Agric. Eng., 19(1): 15.
Chow VT (1951). A generalized formula for hydrologic frequency analysis. Transac. Am. Geophy. Union. 32(2): 231-237.
Dabralr PP, Pandey A (2008). Frequency Analysis for One Day to Seven Consecutive Days of Annual Maximum Rainfall for the District of North Lakhimpur, Assam. IE(I) J. AG. 89:29-34.
Haan CT (1994). Statistical methods in hydrology. Affiliated East-West Press Pvt. Ltd., New Delhi.
Mohanty S, Marathe RA and Singh S (1999). Probability analysis of annual maximum daily rainfall for Amravati. Indian J. Soil Conser. 43(1): 15-17
Rizvi RH, Singh R, Yadav RS, Tewari RK, Dadhwal KS and Solanki KR (2001). Probability analysis of annual maximum daily rainfall for bundelkhand region of Uttar Pradesh. Indian J. Soil Conser. 29(3): 259-262.
Schwab GO, Frevert RK, Edminster TW and Barnes KK (1981). Soil and Water Conservation Engineering. John Wiley & Sons, New York p. 25.
Tilahun K (2006). The characterization of rainfall in the arid and semi-arid regions of Ethiopia. Water S.A. 32(3): 429-436.
Upadhaya A and Singh SR (1998). Estimation of consecutive days Maximum Rainfall by Various Methods and their Comparison. Indian J. Soil Conser. 26(2): 193-2001.