

## **Probability Analysis for Prediction of Consecutive Days Rainfall of Allahabad Region**

DHARAMABIR DHAN<sup>1</sup>, M. IMTIYAZ<sup>1</sup>, A. K. BAXLA<sup>2</sup>, ENO RAI<sup>1</sup> AND N. BIWALKAR<sup>3</sup>

<sup>1</sup>*Department of Soil Water Land Engineering and Management Sam Higginbottom  
Institute of Agriculture, Technology & Sciences, Allahabad 211007, India*

<sup>2</sup>*Agromet Service Cell, IMD, N. Delhi, India*

<sup>3</sup>*Department of Soil Water Engineering, PAU, Ludhiana, India*

### **Abstract**

A study was undertaken to predict the consecutive days rainfall for Allahabad. It comes under the influence of Ganga and Yamuna rivers. Three common probability models (Gumbel, Log Pearson Type III and Log Normal) were tested and compared to chi-square value. The return periods was obtained by weibull's formula. The rainfall data were fitted in the above distributions and their corresponding rainfall events were estimated at 10, 20, 25, 50 and 95% probability levels. The comparison between the measured and predicted maximum rainfall, resulted that the developed model can be efficiently used for the prediction of maximum rainfall of one to five consecutive days. The statistical comparison by chi-square test indicates that the Log normal was found to be best model for predicting one and four consecutive days maximum rainfall while Log Pearson Type III model shows fairly close to observed values for three and five consecutive days maximum rainfall. Rainfall prediction by Gumbel distribution shows close relation to the observed rainfall for two consecutive days maximum rainfall.

**Key words :** Consecutive days, Allahabad, Rainfall prediction, Probability analysis.

Rain plays a major role in hydrology that finds its greatest application in the design and operation of water resources, engineering works and agricultural systems (1). In an agricultural country like India, the success or failure of the crops and water scarcity in any year is always viewed with greatest concern. These problems are closely linked with the behavior of the monsoons in India. Therefore, daily rainfall data (1950—2003) analysis is more useful for planning crop patterns, water management practices and periods of droughts. Three wide probability models (Gumbel, Log Pearson Type III, and Log Normal) were tested and compared to chi-square value. As the growth rate of its population in India is increasing day by day, therefore it becomes important to use the available natural resources for the sustainable crop production. The agriculture production can be increased significantly with proper management of rainwater, application of optimum irrigation in time and drainage of agricultural fields. Since the occurrence and distribution of rainfall is erratic and not evenly distributed overtime and space. Therefore, probability analysis can be used for prediction of occurrence of future

events from available records of rainfall with the help of statistical methods (2) for solving various water management problems and to assess the crop failure due to deficiency or excess of rainfall. According to Kite (3) the most important criteria in selection of distribution functions are : It should be theoretically based function; and it should extract the maximum information from the data available. The primary need of water resource development in any area depends on the estimation of rainfall at different probabilities for efficient planning and design of irrigation and drainage systems, command area development, soil and water conservation programmers and the optimum utilization of water resources in various agricultural production systems. Most of the watershed planning activities include the estimation of runoff volume, design of water storage structures and erosion control structures and efficient utilization of runoff for irrigation of different crops. Recently few researchers have been done on consecutive days rainfall prediction for different places in India (4—8). As pattern varies from place to place the most suitable distribution may also vary from place to place (9). Therefore,

**Table 1.** Comparison between observed and prediction maximum rainfall by various prediction models. P (%) = Probability percentage, LPIII = Log Pearson Type III, LN = Log Normal.

	P (%)	Return periods (T)	Observed rainfall (O)	Predicted rainfall (E)			Chi-square		
				Gumbel	LPIII	Ln	Gumbel	LPIII	LN
One day maximum	95	1.05	61.7	64.256	54.828	56.623	0.102	0.861	0.455
	50	2	95	92.725	94.406	93.43	0.056	0.004	0.023
	25	4	113	109.638	106.169	105.68	0.122	0.44	0.507
	20	5	113.3	113.523	112.46	112.46	0.977	0.006	0.006
	10	10	125.1	128.955	123.56	123.59	0.063	0.636	0.018
	Sum						1.320	1.347	1.009
	Fitting condition						unfit	unfit	Best fit
Two days maximum	95	1.05	82.1	73.473	88.715	64.565	1.013	0.493	4.762
	50	2	35.8	122.939	120.22	123.59	1.345	2.019	1.205
	25	4	164.4	152.334	141.905	144.54	0.955	3.566	2.728
	20	5	165.1	160.826	154.525	156.31	0.144	0.724	0.494
	10	10	167.4	185.910	179.06	176.6	1.843	0.759	0.48
	Sum						5.270	7.561	9.669
	Fitting condition						Best fit	Unfit	Unfit
Three days maximum	95	1.05	89.7	84.014	82.224	77.446	0.385	0.68	1.939
	50	2	146.2	141.488	140.92	142.56	0.157	0.198	0.009
	25	4	169.9	175.641	164.058	165.2	0.187	0.208	0.134
	20	5	175.2	185.507	177.01	177.83	0.573	0.019	0.039
	10	10	187.2	214.652	200.9	199.53	3.511	0.997	0.761
	Sum						4.812	2.102	2.966
	Fitting condition						unfit	Best fit	Unfit
Four days maximum	95	1.05	95.9	93.693	95.279	88.307	0.052	0.004	0.653
	50	2	165.9	158.563	154.8	159.96	0.339	0.796	0.221
	25	4	184.4	197.112	182.81	184.93	0.820	0.014	0.001
	20	5	185.3	208.248	197.242	198.61	2.529	0.723	0.892
	10	10	203.2	241.143	224.3	222.33	5.970	1.985	1.646
	Sum						9.710	3.522	3.413
	Fitting condition						unfit	unfit	Best fit
Five days maximum	95	1.05	95.90	100.642	100.23	94.623	0.223	0.187	0.017
	50	2	173.6	171.731	171	173.38	0.020	0.040	0.000
	25	4	187.8	213.974	199.06	200.05	3.203	0.064	0.749
	20	5	193.6	226.178	214.78	215.77	4.692	2.089	2.279
	10	10	240.3	262.226	247.78	242.1	1.833	0.050	0.013
	sum						9.972	2.430	3.085
	Fitting condition						Unfit	Best fit	Unfit

commonly accepted theoretically based probability functions namely Gumbel, Log Pearson Type III and Log Normal are used for prediction of maximum annual rainfall of one to five consecutive days. The analysis of rainfall at different durations like one day, weekly, monthly and seasonal is important for better planning and management of water resources. Therefore attempt was made to find out possible forecasting of one to five consecutive days maximum rainfall.

(First author wishes to acknowledge the finan-

cial assistance obtained from UGC. We are thankful to all reviewers for their valuable suggestions and comments on the manuscript).

### Methods

Daily rainfall data for a period of 19 years (1985–2003) for Allahabad region were collected from India Meteorological Department, Pune and were analyzed to obtain maximum annual rainfall of one to five con-

secutive days by fitting different theoretical distributions viz., Gumbel, Log Pearson Type III and Log Normal. The study area was located at Sangam or the confluence of sacred river, Ganga, Yamuna and the mythical Saraswati. It is 98 m above sea level and comes under humid sub-tropical (warm summer) regions. Allahabad experiences all four seasons. The summer season is from April to mid June with the maximum temperatures ranging between 40—45 C. Monsoon set in early June and continues upto the end of September. The minimum temperature ranges to 2—5 C.

Return period T was computed

$$T = \frac{(N + 1)}{m} \quad \dots 1$$

Where, N is the total number of years of record and m is the rank of the observed rainfall values when arranged in descending order. The probability of rainfall values is the reciprocal of the return period. These statistical parameters were used to evaluate the probability analysis for prediction of maximum rainfall of one to five consecutive days period for Allahabad region.

*Probability of Occurrence of Rainfall*

Probability of occurrence of rainfall after the estimated return period was calculated with the following formula.

$$\chi_t = \bar{X} + K \times \sigma_n \quad \dots 2$$

Where,  $\chi_t$  is the rainfall amount for return period of T years,  $\bar{X}$  is the mean of rainfall data,  $\sigma_n$  is the standard deviation, K is the frequency factor which depends upon the return period (T).

*Gumbel Distributions*

Gumbel equation was used to analyzed the one to five consecutive days rainfall probability distribution (10). The mean and standard deviation were also computed.

Reduced variate was calculated as

$$y_t = - \ln \ln \left( \frac{T}{T-1} \right) \quad \dots 3$$

Frequency factor of Gumbel model was estimated by

$$K = \frac{\gamma_t - \gamma_n}{S_n} \quad \dots 4$$

Where,  $y_t$  is the reduced variate and it is a function of T and K is the frequency factor of Gumbel distribution,  $\gamma_n$  is the Reduced mean and  $S_n$  is the reduced standard deviation. Both  $\gamma_n$  (reduced mean) and  $S_n$  (reduced standard deviation) are the function of sample size N, and its values are available in standard tables (11). Predicted rainfall was calculated by equation (2).

*Log Pearson Type III Distribution*

In this distribution, the sample (i. e.  $\bar{Z}$  in this case) is first transformed into logarithmic form and the mean and standard deviation of the log series were computed. For Log Pearson type III distribution,  $K_z$  is a function of both the return period and the coefficient of Skewness ( $C_s$ ).

Predicted rainfall was computed by

$$\chi_t = \text{antilog} (\bar{Z} + K_z \times \sigma_n) \quad \dots 5$$

*Log Normal Distribution*

The frequency factor for the Log Normal distribution was derived by Chow (12). In this method, the sample (i.e. X in this case) is first transformed into logarithmic form and the mean and standard deviation of the log transformed series were computed. At zero Skewness (i.e.  $C_s = 0$ ), the Log Pearson type III distribution was reduced to Log Normal Distribution. Predicted rainfall was estimated by equation (5). The value of frequency factor ( $K_x$ ) was taken from the statistical table corresponding to  $C_s$  to T (recurrence interval).

*Test of Goodness of Fit*

All three probability distribution function were tested for goodness of fit by chi-square ( $\chi^2$ ). The test is generally used to test the closeness of the expected

values obtained by fitting the theoretical distribution and the observed values (10).

$$\chi^2 = \frac{\sum (O - E)^2}{E} \quad \dots 6$$

Where O and E are observed and expected rainfall values, respectively. The distribution with the least sum of C values will be adjusted the best (13).

### Results and Discussion

Upadhyaya and Singh (14) fitted consecutive day rainfall maxima, up to 5 days. The expected values for one to five consecutive days maximum rainfall were estimated by three most commonly used probability distribution method viz. Gumbel, Log Pearson Type III and Log Normal. The one to five consecutive days maximum rainfall values were computed and the maximum rainfall at Allahabad was 136.5, 194.6, 267.4, 315.8 and 344.2 mm for one, two three, four and five consecutive days respectively. The statistical parameter i.e. chi-square test for goodness of fit was conducted for all proposed models. The least sum of the chi-square values fitted the model best (15) Table 1 shows the minimum (1.009) sum of chi-square value for Log Normal, which reveals that the model can be accurately used for predicting one day maximum rainfall. For two consecutive days maximum rainfall the sum of chi-square value was minimum (5.270) for Gumbel model. In case of three consecutive days maximum rainfall the sum of chi-square values was minimum (2.102) for Log Pearson Type III model and resulted that it can be used for predicting rainfall for three consecutive days maximum rainfall. The least sum of chi-square of four consecutive days maximum rainfall was observed by Log Normal method which reveals the overall accuracy of the model for predicting rainfall. For five consecutive days maximum rainfall the Log Pearson type III was accepted for rainfall prediction as minimum value of chi-square was estimated best (2.430).

### Conclusion

The data were analyzed for fitting one to five day consecutive maximum rainfall through various theoretical distribution, Gumbel, Log Pearson Type III and

Log Normal. The most suitable distribution may also vary at different places. The statistical comparison was done at 10, 20, 25, 50 and 95% probabilities by chi-square test for goodness of fit. The following distribution is identified for the present rainfall analysis of the study area for reliable estimates of rainfall at different duration considered. Log Normal distribution was found to be best models for predicting one and four consecutive days maximum rainfall (mm). Log Pearson type III are fairly close to observed values of rainfalls for three and five consecutive days maximum rainfall (mm). Rainfall prediction by Gumbel distribution is showing very near to the observed rainfall for two consecutive days maximum rainfall (mm). The results may be utilized by the decision maker and farmer for adopting appropriate strategies for management of drought and floods.

### References

1. Srikanthan R. and T. A. McMahon. 2005a. Automatic evaluation of stochastically generated rainfall data. *Aust. J. Water Resour.* 2005, 8 : 195—201.
2. Kumar D. and S. Kumar. 1989. Rainfall distribution pattern using frequency analysis. *J. Agric. Engg.* 26 : 33—38.
3. Kite G. W. 1977. *Frequency and risk analysis in hydrology.* Water Resource Publ., Fort Collins, Colorado, USA.
4. Kumar V. 2003. Frequency analysis of consecutive day maximum rainfall at Srinagar (J & K) *Ind. J. Soil Cons.* 31 : 295—298.
5. Sethy B. K., Shakir Ali and S. N. Prasad. 2005. Frequency analysis for one day to five consecutive days annual maximum rainfall South-Eastern Rajasthan. *Ind. J. Soil Cons.* 33 : 22—26.
6. Bhakar S. R., A. K. Bansal, N. Chhajed and R. C. Purohit. 2006. Frequency analysis of consecutive days maximum rainfall at Banswara, Rajasthan, India. *ARPN J. Engg. and Appl. Sci.* 1 : 64—67.
7. Dingre S. and N. C. Shahi. 2006. Consecutive days maximum rainfall prediction from one day maximum rainfall for Srinagar in Kashmir Valleys. *Ind. J. Soil Cons.* 34 : 153—156.
8. Bhakar S. R., A. K. Bansal and N. Chhajed. 2009. Frequency analysis of consecutive days of maximum rainfall at Udaipur. *IE (I) J. AG* 14—16 pp.
9. Singh R. K. 2001. Probability analysis for prediction of annual maximum daily rainfall of eastern Himalaya (Sikkim mid Hills). *Ind. J. Soil Cons.* 29 : 263—265.
10. Hogg R. V. and E. A. Tanis. 1977. *Probability and statistical inference.* Macmillan Publ.Co. Inc., New York, USA.
11. Subramanya K. 1984. *Engineering hydrology.* Tata Mc Graw-Hill Publ. Co. Ltd., New Delhi, India.

- 242—253 pp.
12. Chow Ven Te., R. Maidment David and W. Mays Larry. 1988. *Applied hydrology*. Mc Graw Hill. Int. Ed. Civil Engg. Series. 250—405 pp.
  13. Agarwal M. C., V. S. Katiyar and R. Babu. 1988. Probability analysis of annual maximum daily rainfall of U.P. Himalaya. *Ind. J. Soil Cons.* 16 : 35—43.
  14. Upadhyaya A. and S. R. Singh. 1998. Estimation of consecutive days maximum rainfall by various method and their comparison. *Ind. J. Soil Cons.* 26 : 193—201.
  15. Bhatt V. K., A. K. Tiwari and A. K. Sharma. 1996. Probability models for prediction of annual Maximum daily rainfall for Datia. *Ind. J. Soil Cons.* 24 : 25—27.