JOURNAL OF
ENVIRONMENTAL HYDROLOGY

The Electronic Journal of the International Association for Environmental Hydrology
On the World Wide Web at http://www.hydroweb.com

VOLUME 13 2005

EVALUATION OF THE SYNTHETIC ANNUAL MAXIMUM STORMS

Betül Saf

Pamukkale University
Hydraulics & Hydrology & Water Resources Department
Denizli, Turkey

Long term historical records of hydrological information such as rainfall and runoff data form the basis of planning and design of major water resources projects. However, in most cases such historical records are unavailable, and in situations where they are available, the records are too short to have any statistically significant meaning. One approach that has been adopted to overcome this difficulty is to generate long-term data synthetically. In this study, the outcome of an attempt to generate synthetic rainfall data is presented. The Monte Carlo method is used, which is an experimental statistical method in generating samples for solving some probability problems as old as probability theory itself. In this study, synthetic annual maximum storms distributed as Type-I Extremal (or Gumbel) with random effective durations and specific time distribution for given population mean and variance are generated using the method. Effective durations of the synthetic annual maximum storms are related to the basin characteristics, length (L) and harmonic slope (S) of the main course, by Kirpich's time of concentration relationship. After synthetic annual maximum storms are generated, sample statistics and frequency distributions of the generated annual maximum storms of random effective durations are investigated. Eight well-known probability distribution models, (Normal (N), LogNormal with two and three parameters (LN2 and LN3), Gumbel (GUM), LogGumbel (LGUM), Gamma with two and three parameters (G2 and G3), and LogPearson 3 (LP3), with moment and maximum likelihood parameters are used for synthetic storm series by chi-square and probability plot correlation goodness of fit tests. The results of the study reveal that the probability distribution of the rainfall input may even diverge from their parent (Type-I Extremal) distributions because of the sampling, and since the generated input series is a mixture of rainfall events of variable durations.

INTRODUCTION

For appropriate design and management of water resources structures, hydrologists must deal with large storm and flood events. This type of information is needed for a wide variety of design applications, including dams spillways, and culverts. Unfortunately, many small watersheds are ungauged and do not have rainfall and flow records, and the historical records are often too short to accurately predict extreme events (Ajward, 1996).

However, because of the uncertainty in rainfall-runoff estimates, it may be more appropriate to compute “the probability distribution of the criterion variable” by considering the probabilistic structure of rainfall input and the climatologic and physiographic factors of the watershed.

The main objective of this study is to see if the type of the parent distribution (Type-I Extremal) is preserved or the distribution law of the mixed-duration storm series converges to another type of frequency distribution. That is, this study aimed to find a satisfactory answer to this question: Does the type of the probability distribution of mixed storm events remain unchanged provided that the storm events of different durations have the same type of probability distribution (say, Type-I Extremal)?

In order to find answers to the above given questions, an experimental statistical procedure is followed in the study. Forty-two sets of synthetic storms of different durations each of size N=100 are generated. Kirpich’s empirical relationship between the time of concentration of the small watershed and the length and slope of the main course is used for deciding the critical rainfall durations over the hypothetical watersheds.

CONJUNCTIVE GENERATION OF UNIFORMLY DISTRIBUTED RANDOM STORM DURATIONS

In this study, synthetic storms of random durations (D) conforming with the basin lag (or time of concentration,t_c) and recurrence intervals (T) are generated from a given probability distribution. The Extreme Value Type I (or Gumbel) probability distribution is selected since it is commonly used in modeling observed rainfall frequencies and is easy to use.

Durations of synthetic storms (D) conforming to basin lag are generated by considering that the length (L) and harmonic slope (S) of the main channel are dominating factors on the time of concentration. The effective storm Duration, D_e, is given as (Chow, 1964; SCS, 1972):

$$D_e=2(t_c)^{0.5} \quad (1)$$

$$t_c=0.00032(L)^{0.77}/S^{0.385} \quad (2)$$

where L is the length of the main course in meters, S is the harmonic slope, t_c is the time of concentration of the watershed and D_e is the effective storm duration, both in hours.

Time of concentration (t_c) according to Kirpich’s formula Equation 2 for various lengths of main course (L) and harmonic slopes (S) are given in Table 1, while the effective rainfall durations (D_e) are given in Table 2. For $t_c \geq 4$ hour, effective storm durations are assumed to be equal to time of concentration ($D_e=t_c$).

When a random duration D is generated, the time interval to be used for the hyetograph can be decided from

$$\Delta t = D/M \quad (3)$$

Table 1. Time of Concentration (t_c; hour) According to Kirpich's Formula for Various Lengths of Main Course (L) and Harmonic Slopes (S)

		S	(Harmonic	Slope)		
L(m)	0.001	0.005	0.007	0.01	0.015	0.020
8000	4.63	2.49	2.20	1.91	1.63	1.46
10000	5.50	2.96	2.60	2.26	1.94	1.74
12000	6.32	3.40	3.00	2.60	2.23	2.00
14000	7.12	3.83	3.40	2.94	2.51	2.25
16000	7.89	4.25	3.73	3.25	2.78	2.49
18000	8.64	4.65	4.10	3.56	3.00	2.73
20000	9.37	5.00	4.44	3.86	3.30	2.95

Table 2. Effective Rainfall Durations (D_e; hour) for Various Lengths of the Main Course (L) and Harmonic Slopes (S)

		S	(Harmonic	Slope)		
L(m)	0.001	0.005	0.007	0.01	0.015	0.020
8000	4.63	3.16	2.96	2.76	2.55	2.42
10000	5.50	3.44	3.22	3.01	2.78	2.63
12000	6.32	3.69	3.46	3.23	2.98	2.83
14000	7.12	3.92	3.67	3.43	3.17	2.99
16000	7.89	4.25	3.86	3.61	3.34	3.15
18000	8.64	4.65	4.10	3.77	3.49	3.30
20000	9.37	5.00	4.44	3.93	3.64	3.44

where M is the number of Δt intervals in the storm duration.

Δt computed from Equation 3 is rounded off as multiple of 15 min., 30 min., 60 min for durations shorter than 5 or 6 hours ($D \leq 6$ hour), and multiples of 2, 3, or 4 hour for larger rainfall durations.

TYPE-I EXTREMAL DISTRIBUTION OF ANNUAL MAXIMUM STORMS WITH ASYMPTOTIC MOMENT PARAMETERS

In this study, synthetic rainfall depths, Y_D, distributed as the Extreme Value Type-I (Gumbel) with asymptotic parameters given by Equations 6 and 7 are generated through Equation 4.

$$Y_{D,T} = M_D + S_D . K_T \tag{4}$$

where

$$K_T = - \{0.45 + 0.7797 \ln[-\ln(1-1/T)]\} \tag{5}$$

is the frequency factor corresponding to the return period T or probability of nonexceedance P_T which can be calculated from generated uniform random numbers $0 < P_T < 1$, that is $T=1/(1-P_T)$.

Type I Extremal (or Gumbel) and LogGumbel Distributions

The Type I Extremal (or Gumbel) distribution is often used for maximum type events, such as the annual maximum storms and annual peak flows.

The probability density function (pdf) of the Gumbel distribution is given by Equation 6.

$$f(x) = \frac{1}{\alpha} \exp\left[-\left(\frac{\chi - \beta}{\alpha} \right) - e^{-\left(\frac{\chi - \beta}{\alpha} \right)} \right] \tag{6}$$

The cumulative distribution function of x is given by Equation 7, where α is the scale parameter and β is the mode of the pdf (Yevjevich, 1972a; Kite, 1977).

$$\Pr(x < X) = F(x) = \exp\left[-e^{-\alpha(x-\beta)}\right] \tag{7}$$

The reduced variate of the distribution u, can be calculated from a double-logarithmic transformation of the non-exceedence probabilities:

$$u = \alpha(x - \beta) = -\ln[-\ln F(x)] \tag{8}$$

Chow (1954) has considered the Type-I Extremal distribution as a special case of the lognormal distribution for which the coefficient of skew is γ_I=1.1396.

The frequency factor for the Type-I Extremal distribution is given by (Kite, 1977),

$$k_T = \left(u_T - \bar{u}\right)/ S_u \tag{9}$$

where

$$u_T = -\ln\left[-\ln\left(1 - \frac{1}{T}\right)\right] \tag{10}$$

and

$$\bar{u} = \frac{1}{N}\sum_{m=1}^{N} U_m \tag{11}$$

$$S_u = \sqrt{\sum_{m=1}^{N}\left(u_m - \bar{u}\right)^2 / N} \tag{12}$$

where $\bar{u}$ and S_u depends only on the sample size N and the plotting position formula used in the analysis.

In the last two equations, u_m is the reduced variate, u_m= -ln[-lnP_m], corresponding to the empirical probability of the ordered event, x_m.

Short sample moment estimators of a and β are (Kite, 1977)

$$\left.\begin{aligned} \alpha &= S_u / S \\ \beta &= \bar{x} - \bar{u}/\alpha \end{aligned}\right] \tag{13}$$

If N is large enough, the asymptotic moment estimators of a and β calculated by

$$\left.\begin{aligned} \alpha &= 1.2825/S \\ \beta &= \bar{x} - 0.45S \end{aligned}\right] \tag{14}$$

can be used in Equation 8, and the frequency factor equation takes the form (Kite, 1977)

$$K_T = -\left[0.45 + 0.7797\ln\{-\ln(1 - 1/T)\}\right] \tag{15}$$

Recalling Equation 8, and knowing that $F(x)=P_T$, the T-year event magnitude can be calculated from Equation 16 as,

$$Y_{D,T} = \beta + \left[-\ln(-\ln P_T)\right]/\alpha \tag{16}$$

where a and β, are the asymptotic moment estimators of the Gumbel distribution and are given in terms of the mean (M_D) and standard deviation (S_D) of the D-hour storms as:

$$\alpha = 1.2825/S_D \tag{16a}$$

$$\beta = M_D - 0.45 S_D \tag{16b}$$

Time distribution of the generated storms

Cumulative time distribution of annual maximum storms in Turkey has been put into three broad categories as shown in Figure 1 (Kizilkaya, 1988).

Time distribution curves in Figure 1 proposed by DSI are commonly used in Turkey in deriving rainfall hyetographs of given design storms. In this study DSI's-A and SCS 6-hour time distributions are used in order to calculate rainfall hyetographs of given synthetic storms. The Soil Conservation Service dimensionless cumulative rainfall curves were developed for various storm types, storm durations and regions in the United States (SCS, 1972).

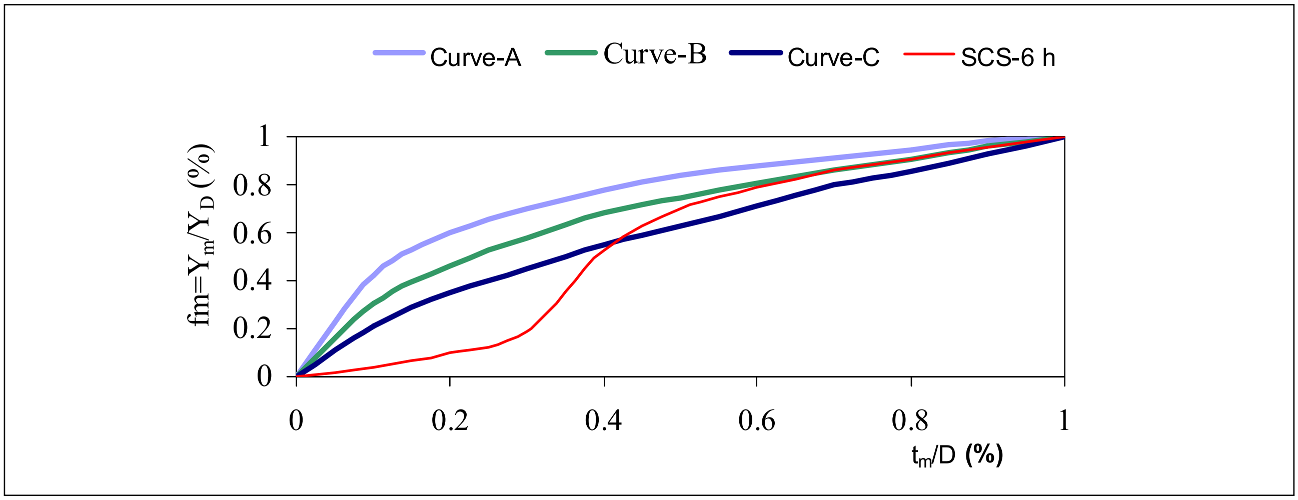

Figure 1. Time distribution of cumulative rainfall (Kizilkaya, 1988; Wanieliesta et al., 1997; SCS, 1972).

A design storm, Y_D, is divided into increments ΔY_m using an appropriate time distribution curve for the project site. A time distribution curve represents the cumulative percentage of the precipitated rainfall $f_m = Y_m / Y_D$, during the percentage time $X_m = t_m / D$, where $D = M\Delta t$ is the rainfall duration and $t_m = m\Delta t$ (m=1,2,...,M). Having f_m values, the cumulative rainfall amount (Y_m) precipitated during period 0 to t_m, and incremental rainfall amount (ΔY_m) can be computed as:

$$Y_m = f_m D, \qquad m=1,2,...,M \tag{17}$$

$$\Delta Y_m = Y_m - Y_{m-1}, \qquad m=1,2,...,M \tag{18}$$

where y_0=0 for m=1.

Assumptions for the computational algorithm for generation of synthetic storms

Rarely is there a constant rainfall excess over a single time increment. Usually, the rainfall excess varies with time. Consider a unit hydrograph method that divides a rainfall into successive shorter time events, each of constant rainfall excess and equal times.

As there are a great number of physiographic, morphologic, climatologic, soils and vegetal cover factors affecting watershed response, it is almost impossible to consider them in the simulation model. Therefore only the principal factors in the following are taken into consideration in the generation algorithm of the synthetic storms:

(1) Principal physiographic factors:

L : length of the main course (m) $(8000 \leq L \leq 20000)$

S : harmonic mean slope of the main course $(0.005 = S = 0.02)$

(2) Time distribution of synthetic storms:

DSI's A-curve (Kizilkaya, 1988) and SCS 6-hour time distributions (Wanieliesta et al., 1997) are assumed to represent the role of the time distribution of storms on the probability distribution of peak flows.

Before calculating synthetic storm hyetographs, total amount of precipitation Y_D^t is reduced by multiplying the areal reduction factor, $F(A,D)$.

$$F(A,D) = \exp(-0.213\ D^{-0.4149} A^{0.3825}) \tag{19}$$

This equation is reduced according to a relation of areal reduction factor with duration (D) and area (A) for the U.K. (Wanielista, et al., 1997).

(3) Critical storm duration and unit graph duration

In this study, it is assumed that the critical storm duration, D, which creates peak flows, is equal to or greater than the effective storm duration, D_e, at a given basin with a concentration time, t_c,

$$D_e \leq D \leq 2\ D_e \tag{20}$$

where D_e is given by Equation 1 for $t_c < 4h$, and is $D_e = t_c$ for $t_c \geq 4\ h$.

POPULATION MEANS AND STANDARD DEVIATIONS OF ANNUAL MAXIMUM STORMS AS FUNCTIONS OF STORM DURATION

As was mentioned, durations (D) of the critical storms, which yield annual peak flows in a given basin, may vary within the range $D_e \leq D \leq 2D_e$. Therefore, elements of the synthetic storms series that annual peak flows are not necessarily to be drawn from the same population.

In this study, the type of the parent distributions is assumed to be permanent (Type-I Extremal), although the population statistics (M_D, S_D or α_D, β_D) of any storm event in a sample series may vary from one event to another because of the natural behavior of an extreme rainfall process (M_D and S_D are rather smooth functions of small duration).

The mean depth-duration and standard deviation-duration relationships of the Usak Meteorological Station developed by Benzeden (2001) are used as population statistics of the synthetic storms. These relationships are as follows:

Mean depth-duration relationship:

$$M_D = 4.154[\ln(D/1.1763]^{1.0263} \tag{21}$$

Standard deviation-duration relationship:

$$S_D = \exp[-1.46047+1.79839\ln(D)-0.3155\ln^2(D)+0.01944\ln^3(D)] \tag{22}$$

where D is in minutes and M_D and S_D are both in mm. Observed (DSI, 1990) and fitted (Benzeden, 2001) mean depths and standard deviations of the Usak meteorological station are shown in Figure 2.

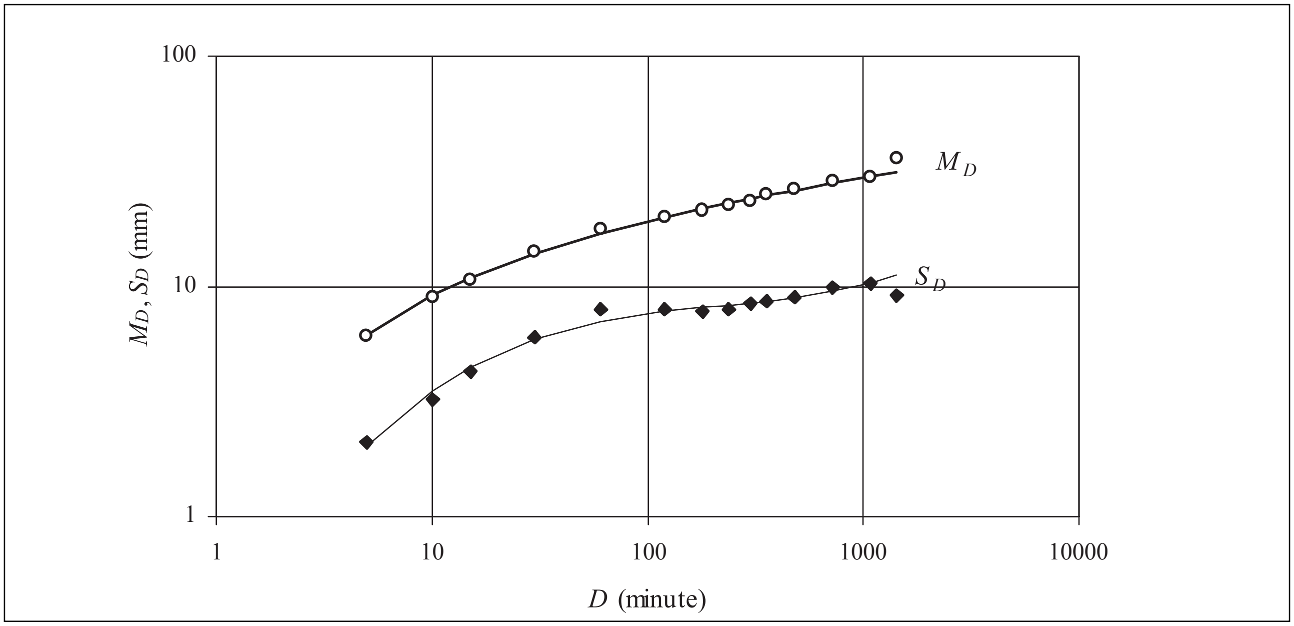

Figure 2. Mean depth-duration and standard deviation-duration relationships of the Usak Meteorological Station (Benzeden, 2001)

For six harmonic slopes and seven main channel lengths, a total of 42 synthetic storm samples (each of size N=100) distributed as Gumbel are generated. Storm sample series are labeled as Y_k, where the integer k varying from 1 to 42 refers to the items of the (L_i, S_j) matrix given by Table 1 and k is computed in terms of i and j as k=6(i-1)+j. For example, the synthetic storm sample series that refers to the hypothetical basin with L_2=10 000 m, S_6=0.020, and t_c=1.74 hour is labeled as Y_{12}, since k=6(2-1)+6=12.

Statistical and distributional properties of the generated storm samples

The statistical and distributional parameters of the generated storm samples with mixed durations are calculated. Eight probability distribution models, namely, the Normal, the two- and three-parameter LogNormal, the extreme value Type-I (Gumbel), the LogGumbel, the two- and three parameter Gamma, and the LogPearson Type III, are chosen to fit to the synthetic samples.

These probability distribution models with moment and maximum likelihood parameters are tested by chi-square and probability plot correlation goodness of fit tests. Details of parameter estimation methods and the goodness of fit tests are presented in Kite, 1977.

Evaluation of the probability distribution of synthetic storm samples by goodness of fit tests

In order to compare the probability distribution models most frequently accepted for the synthetic storm samples, chi-square tests are performed on the 42 series. Relative acceptance frequencies (f_i) of each model at a significance level a=5% are calculated and presented in Figure 3 for MOM parameters and in Figure 4 for ML parameters. Relative acceptance frequency of a specific model (f_i) is defined as:

$$f_i = 100\ (\mathrm{TNCH})/42 \tag{23}$$

where TNCH is total number of series that passed the χ^2-test for a specific distribution.

As another alternative probability plot, correlations of each series for the eight types of frequency distribution models are calculated. A critical correlation r_c= 0.95 for the acceptance of any

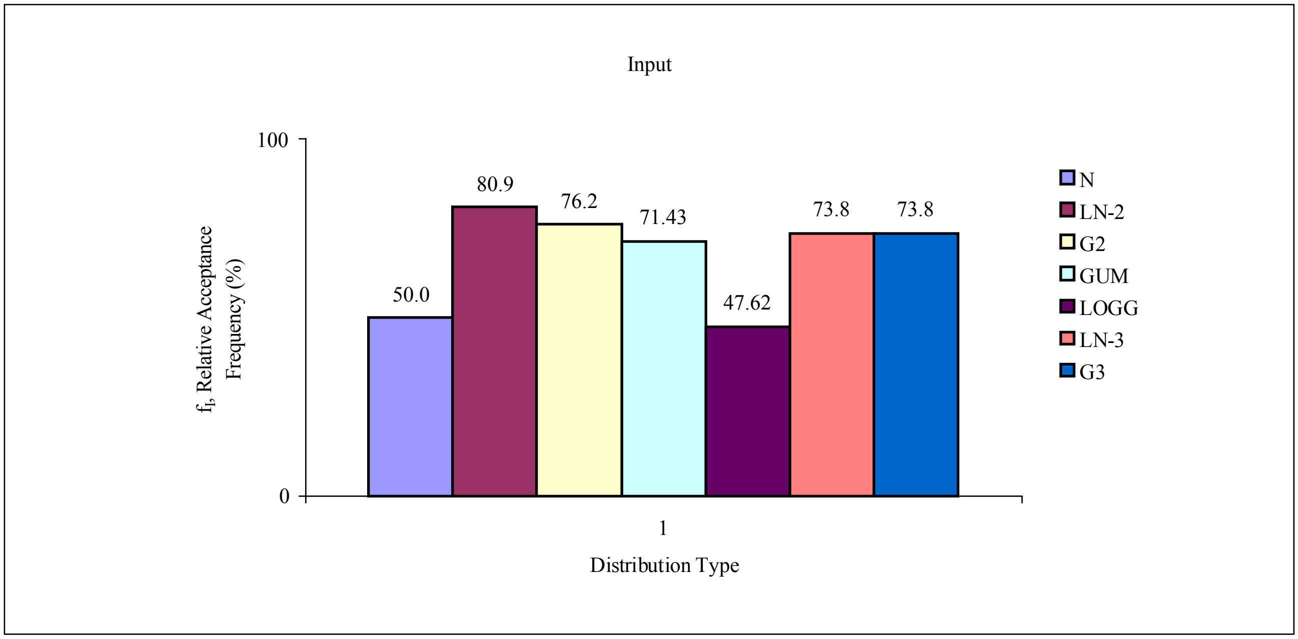

Figure 3. Percentage of total number of synthetic storm sample series (f_I) that passed the chi-square goodness of fit test at 5% level of significance (with MOM parameters) for 42 synthetic storm samples.

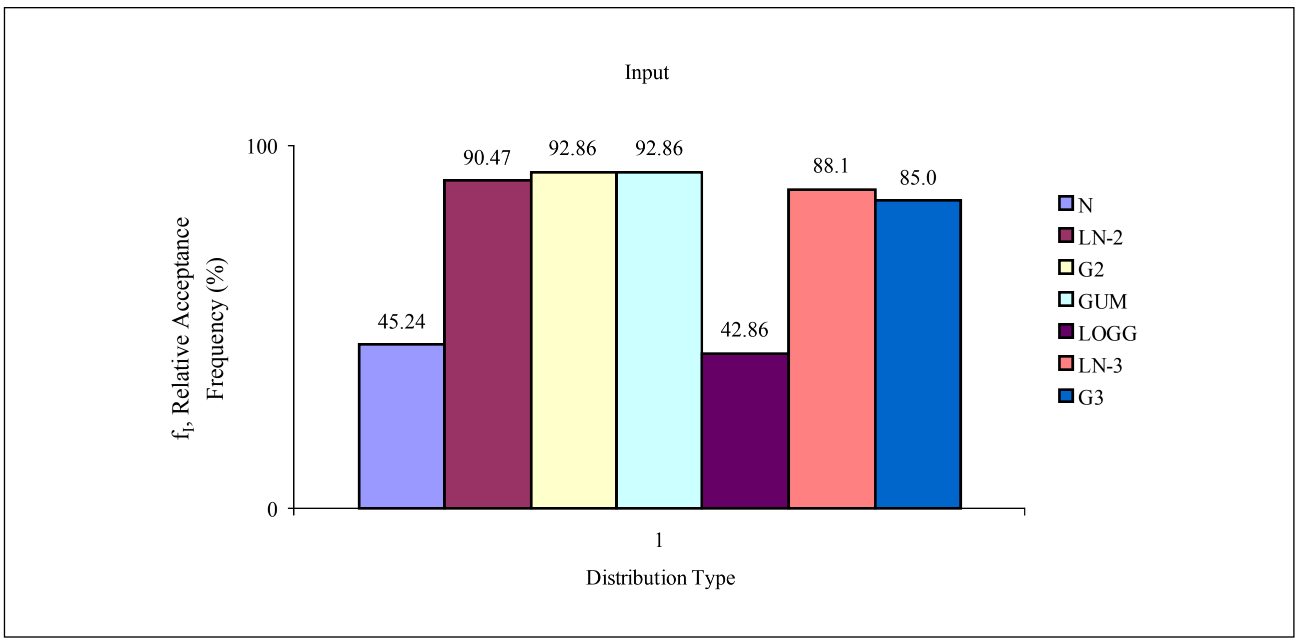

Figure 4. Percentage of total number of synthetic storm sample series (f_I) that passed the chi-square goodness of fit test at 5% level of significance (with ML parameters) for 42 synthetic storm samples.

distribution model is assumed, and relative acceptance frequencies of the alternative models are calculated as in the follows:

$$f_i = 100\ (\text{TNPP})/42 \qquad (24)$$

where TNPP is total number of storm series that have a PPCC coefficient greater than r_c=0.95. The PPCC results are shown in Figures 5 and 6.

Most of the log-transformed synthetic sample series are negatively skewed and therefore the LP3 model is not applicable for those series. Furthermore, the maximum likelihood parameters of many

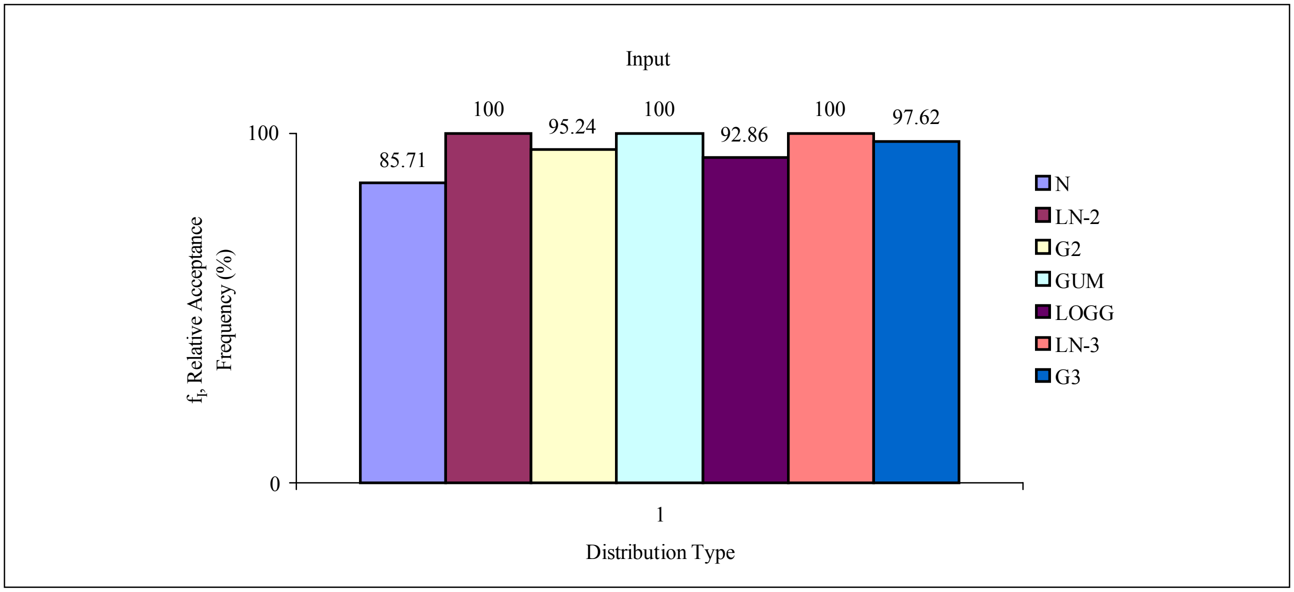

Figure 5. Relative acceptance frequencies of probability plot correlation greater than 0.95 for the alternative frequency distribution models fitted for the synthetic storms (with MOM parameters).

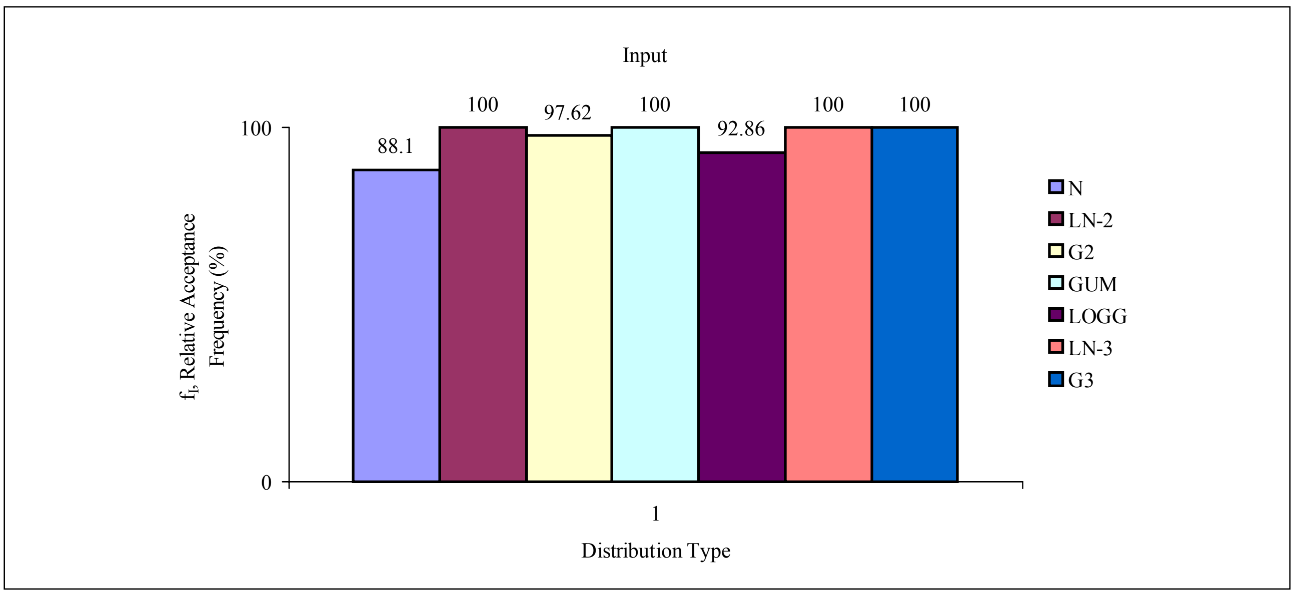

Figure 6. Relative acceptance frequencies of probability plot correlations greater than 0.95 for the alternative frequency distribution models fitted for the synthetic storms (with ML parameters).

series cannot be found since the solution of the likelihood equations did not converge. Therefore, during the application of both chi-square and probability plot correlation tests, the LP3 results were kept out of the evaluation. On the other hand, the maximum likelihood parameters of a few sample series cannot be found for the G3 distribution, but it is included in the evaluations. The acceptance frequencies for the chi-square and PPCC test are calculated.

CONCLUSIONS

As can be seen from Table 3, when the 42 synthetic storm series were evaluated with the chi-square goodness of fit test, the most suitable distribution with MOM parameters is LN2; and G2 and GUM with ML parameters. This means that when the generated synthetic storms distributed as Gumbel are put into a mixed duration series, the type of the appropriate distribution may change.

Table 3. Percentage of Total Number of Acceptances for χ^2 and PPCC tests with MOM and ML parameters

Percentage of Total Number of Acceptances (%)	Type of Goodness of Fit Test	Type of Parameter Estimation	Distribution Type							
			NOR	LN2	G2	GUM	LGUM	LN3	G3	LP3
f_i	χ^2 -test	MOM	21/42	**34/42**	32/42	30/42	20/42	31/42	31/42	--
		ML	19/42	38/42	**39/42**	**39/42**	18/42	37/42	34/40	--
	PPCC test	MOM	36/42	**42/42**	40/42	**42/42**	39/42	**42/42**	41/42	--
		ML	37/42	**42/42**	41/42	**42/42**	39/42	**42/42**	**40/40**	--

Similarly, according to the results of the probability plot correlation test given in Table 3, LN2, GUM and LN3 distributions are the most suitable distributions either with MOM or ML parameters. Under the acceptance criterion selected in this study, it is seen that some other distribution models such as G2 and G3 may approach the others as shown in Figure 5 and 6.

The results of the study reveal that the probability distributions of the rainfall input may diverge from their parent (Type-I Extremal) distributions because of the sampling, and because the generated input series is a mixture of rainfall events of variable durations.

Provided that the storm events of different durations have the same type of probability distribution, the probability distribution types of mixed storm events have not been observed to change significantly.

REFERENCES

Ajward, M.H. 1996. A Spatially Distributed Unit Hydrograph Model Using a Geographical Information System. Ph. D. Diss. Civil Engineering Dept., University of Calgary, Calgary.

Benzeden E. 2001. Standart Süreli Maksimum Yagislarin Frekans Analizinde Karsilasilan Sorunlar. DSI Teknoloji D. Bsk. Basim ve Foto-Film Sb. Md., III. Ulusal Hidroloji Kongresi Bildirileri, Izmir, p.11-18.

Chow, V.T. 1964. Handbook of Applied Hydrology. McGraw-Hill.

Kizilkaya, T. 1988. Sulama ve Drenaj. DSI Genel Müdürlügü.

DSI 1990 . Türkiye Maksimum Yagislari Frekans Atlasi, DSI Genel Müdürlügü, Ankara.

Kite, G.W. 1977. Frequency and Risk Analysis in Hydrology. Water Resources Publications, Fort Collins, Colorado, USA.

Lettenmaier, D.P., and S.J. Burges. 1982. Gumbel's Extreme Value I Distribution: A New Look. Journal of Hydraulic Division, 502-514.

Linsley, R.K., M.A. Kohler, and J. L.H. Paulhus. 1949. Applied Hydrology, McGraw-Hill, New York.

Marien, J.L., and G.L. Vandewiele. 1986. A Point Rainfall Generator with Internal Storm Structure. Water Resources Research, 22, 475-482.

Saf, B. 2003. Role of Rainfall Pattern and Basin Physiography on the Probability Distribution of Floods. Ph. D. Dissertation, Dokuz Eylül University, Turkey.

SCS. 1972. National Engineering Handbook, Section 4, Hydrology SoilConservation Services, U.S., Department of Agriculture, Washington, D. C.

Singh, V.P., and K. Singh. 1988. Parameter Estimation for Log-Pearson Type III Distribution by POME. Journal of Hydrulic Engineering, 112-122.

Vogel, R.M. 1986. The Probability Plot Correlation Coefficient Test for the Normal, Lognormal and Gumbel Distributional Hypthesis. Water Resources Research, 22, 587-590.

Vogel, R.M., and D.E. McMartin. 1991. Probability Plot Goodness-of-fit and Skewness Estimation Procedures for the Pearson Type 3 Distribution. Water Resources Research, 27, 3149- 3158.

Wanieliesta, M., R. Kersten, and R. Eaglin. 1997. Hydrology: Water Quantity & Quality Control. John Wiley & Sons.

ADDRESS FOR CORRESPONDENCE
Dr. Betül SAF
Pamukkale University, Engineering Faculty
Hydraulics & Hydrology & Water Resources Department
Denizli, Turkey

Email:bsaf@pamukkale.edu.tr