At-site flood frequency analysis for Thailand

HN Phien[1]* and N Laungwattanapong[2]
[1]*Asian Institute of Technology, PO Box 2754, Bangkok 10501, Thailand*
[2]*Planning and Technical Division, Department of Agriculture, Bangkok, Thailand*

Abstract

At-site flood frequency analysis was carried out for 64 unregulated streamflow stations located in various regions of Thailand with each data set having at least 20 observations. Using the extreme value type 1 (EV1), general extreme value (GEV) and log-logistic (LLG) distributions as the parent models, along with commonly used methods of parameter estimation, it was found that the GEV is the most suitable distribution, and the method of probability weighted moments is the most desirable method for parameter estimation in this case.

Introduction

Many attempts have been made to search for a statistical distribution which best represents actual flood records. As numerous studies have demonstrated, there is no general agreement among statistical hydrologists as to which distribution best describes these annual series. Among those distributions which have been proposed, the general extreme value (GEV), the log Pearson type 3 (LP3), and log-normal (LN3) distributions are most commonly used with the extreme value type 1 (EV1) or Gumbel distribution being considered as a particular case of the GEV. Although extensive experience with the LP3 distribution (US Water Resources Council, 1967) has been obtained (Phien and Hira, 1983; Phien and Hsu, 1985; Phien and Yang, 1988), this distribution has been found not to be robust. On the other hand, difficulties have been experienced with the LN3 distribution when the desirable method of maximum likelihood is used for parameter estimation. This leaves the GEV as the only obvious choice. This distribution was also favored for use in the UK (NERC, 1975). Recently, Ahmad *et al.* (1988) proposed the log-logistic (LLG) distribution for flood frequency analysis. By application to flood data in Scotland, they found that the LLG possesses many desirable properties.

In Thailand, there has been no systematic study on the determination of the most suitable distribution for flood frequency analysis. It is therefore desirable to consider the GEV, EV1 and LLG distributions for this purpose. The present study will focus on the at-site analysis as a required step towards a more comprehensive flood study for Thailand, which should involve regionalisation as well.

The general extreme value (GEV) distribution

Definitions

The GEV distributions with three parameters, denoted a, b and c for simple notation, has distribution function:

$$F(x) = \exp\{-[1 - b(x - c)/a]^{1/b}\} \tag{1}$$

where:

x is bounded by $c + a/b$ from above for $b > 0$ and from below for $b < 0$. Here $a(>0)$ and c are respectively the scale and location parameters, and the shape parameter b determines which extreme value is represented. Fisher-Tippett types I, II and III correspond to $b = 0$ (Gumbel), $b < 0$ (Frechett) and $b > 0$ (Weibull), respectively.

The inverse distribution function for $b \neq 0$ is:

$$x(F) = c + a\{1 - (-\ln F)^b\}/b \tag{2}$$

By differentiation of Eq. 1, the density function of the GEV distribution is obtained:

$$f(x) = \exp\{-y - \exp(-y)\}/[a(1 - t)] \tag{3}$$

in which y is the reduced variate:

$$y = -\ln(1 - t)/b \tag{4}$$

and t is given by:

$$t = b(x - c)/a \tag{5}$$

The T-year event X_T (i.e the value with a return period of T years) is defined as:

$$Prob(X > X_T) = 1/T$$

In view of Eq. 1:

$$X_T = c + (a/b)\{1 - [-\ln(1 - 1/T)]^b\} \tag{6}$$

The mean, variance and skewness of the GEV distribution are:

$$E(X) = c + a\{1 - \Gamma(1 + b)\}/b \tag{7}$$

$$Var(X) = \{\Gamma(1 + 2b) - \Gamma^2(1 + b)\}(a^2/b^2) \tag{8}$$

$$Skew(X) = -(b/|b|) * \frac{\Gamma(1 + 3b) - 3\Gamma(1 + 2b)\Gamma(1 + b) + 2\Gamma^3(1 + b)}{[\Gamma(1 + 2b) - \Gamma^2(1 + b)]^{3/2}} \tag{9}$$

Methods of parameter estimation

Method of probability weighted moments (PWM)

The probability weighted moments of a random variable X with distribution function ($F(x) = (P(X \leq x)$) are the quantities:

$$M_{p,r,s} = E[X^p\{F(X)\}^r\{1 - F(X)\}^s] \tag{10}$$

*To whom all correspondence should be addressed

Received 14 September 1990; accepted in revised form 24 October 1990

where:

p, r and s are real numbers.

For the GEV distribution, Hosking *et al.* (1985) used Eq. 10 with p and s taking values 1 and 0 respectively, and r taking the values 0,1,2.... Then we may write:

$$M_r = M_{1,r,0} = E[X\{F(x)\}^r] \qquad (11)$$

When this is applied to the GEV distribution for b = O, the PWM estimators are obtained as (Hosking *et al.*, 1985):

$$M_r = (r+1)^{-1}[c + a\{1-(r+1)^{-b}\Gamma(1+b)\}/b], \quad b > -1 \qquad (12)$$

When $b \leq -1$, M_0 (the mean of the distribution) and the rest of the M_r do not exist. Given a random sample of size n from the distribution F, estimation of M is more conveniently based on the ordered sample $x_1 \leq x_2 \leq \ldots \leq x_n$. The statistic:

$$m_r = n^{-1} \sum_{j=1}^{n} \frac{(j-1)(j-2)\ldots(j-r)}{(n-1)(n-2)\ldots(n-r)} x_j \qquad (13)$$

is an unbiased estimator of M_r (Landwehr *et al.*, 1979).

For the GEV distribution, only three estimators of PWM are needed, which are obviously as follows:

$$m_0 = n^{-1} \sum_{j=1}^{n} x_j$$

$$m_1 = n^{-1}(n-1)^{-1} \sum_{j=1}^{n} (j-1)x_j$$

$$m_2 = n^{-1}(n-1)^{-1}(n-2)^{-1} \sum_{j=1}^{n} (j-1)(j-2)x_j$$

Hosking *et al.* (1985) proposed the approximators of the GEV parameters as below:

$$\begin{aligned} b &= 7.8590d + 2.9554d^2 \\ a &= (2m_1 - m_0)b/\Gamma(1+b)(1-2^{-b}) \\ c &= m_0 + d\{\Gamma(1+b) - 1\}/b \end{aligned} \qquad (14)$$

where:

$d = (2m_1 - m_0)/(3m_2 - m_0) - [ln2/ln3]$

Method of maximum likelihood (MML)

The log likelihood function of the GEV distribution is

$$L = -nlna - (1-b)\sum_{i=1}^{n} y_i - \sum_{i=1}^{n} e^{-y_i} \qquad (15)$$

where:

y_i is the reduced variate corresponding to x_i.

The first partial derivative of L with respect to each of the three parameters to be estimated is equated to zero. This yields three non-linear equations which when solved produce the ML estimates. These equations can be solved iteratively. Hosking (1985) provided an algorithm to actually maximise, L, which is used in this study.

Variances-covariances of estimators

The efficiencies of the various methods of parameter estimation can be evaluated by the variances of the estimators obtained by these methods.

Method of probability weighted moments

Although the expressions for the asymptotic variances-covariances of the parameter estimators are available (Hosking *et al.*, 1985), these are quite complicated. Consequently, the Jack-knife technique is used to obtain the variance of the PMM estimators and the variance of X_T. The detail of this technique was provided by Yang and Robinson (1986), pp. 150-160).

Method of maximum likelihood

The asymptotic variances-covariances of the ML estimators can be computed in two different ways. The first method is obtained by inversion of the Fisher information matrix:

$$M = \begin{pmatrix} -E\frac{\partial^2 L}{\partial a^2} & -E\frac{\partial^2 L}{\partial a \partial b} & -E\frac{\partial^2 L}{\partial a \partial c} \\ -E\frac{\partial^2 L}{\partial b \partial a} & -E\frac{\partial^2 L}{\partial b^2} & -E\frac{\partial^2 L}{\partial b \partial c} \\ -E\frac{\partial^2 L}{\partial c \partial a} & -E\frac{\partial^2 L}{\partial c \partial b} & -E\frac{\partial^2 L}{\partial c^2} \end{pmatrix}$$

where:

E donotes the expected value operator.

The explicit formulas for the elements of M were given by Prescott and Walden (1980). The second method is obtained by the inverse of the observed information matrix:

$$H = \begin{bmatrix} -\frac{\partial^2 L}{\partial a^2} & -\frac{\partial^2 L}{\partial a \partial b} & -\frac{\partial^2 L}{\partial a \partial c} \\ -\frac{\partial^2 L}{\partial b \partial a} & -\frac{\partial^2 L}{\partial b^2} & -\frac{\partial^2 L}{\partial b \partial c} \\ -\frac{\partial^2 L}{\partial c \partial a} & -\frac{\partial^2 L}{\partial c \partial b} & -\frac{\partial^2 L}{\partial c^2} \end{bmatrix}$$

Once the variances-covariances of the parameter estimators have been obtained, the variance of the T-year event can easily be calculated.

From Eq. 6:

$$\begin{aligned} &Var(X_T) = \\ &(\partial X_T/\partial a)^2 Var(a) + (\partial X_T/\partial b)^2 Var(b) \\ &+ (\partial X_T/\partial c)^2 Var(c) + 2(\partial X_T/\partial a)(\partial X_T/\partial b)Cov(a,b) \\ &+ 2(\partial X_T/\partial b)(\partial X_T/\partial c)Cov(a,c) \\ &+ 2(\partial X_T/\partial c)(\partial X_T/\partial a)Cov(c,a) \end{aligned}$$

where:

$$\begin{aligned} \partial X_T/\partial a &= [1 - (-lnT_1)^b]/b \\ \partial X_T/\partial b &= -(a/b^2)[1 - (lnT_1)^b] - (a/b)(-lnT_1)^b ln(-lnT_1) \\ \partial X_T/\partial c &= 1 \\ T_1 &= (T-1)/T \end{aligned} \qquad (16)$$

The extreme value type 1 distribution (EV1)

Definitions

The EV1 distribution or Gumbel distribution is one type of the GEV distribution corresponding to $b=0$. The EV1 is commonly defined by its distribution function:

$$F(x) = \exp\{-\exp[-(x-u)/a]\} = \exp[-\exp(-y)] \quad (17)$$

where:

a and u are respectively the scale and location parameters with $a > 0$, and y is the reduced variate defined by:

$$y = (x-u)/a$$

The density function is obtained by differentiating Eq. 17:

$$f(x) = (1/a)\exp[-y - \exp(-y)] \quad (18)$$

For a given return period T, the magnitude X_T of the T-year event is obtained from Eq. 17 as:

$$X_T = u + aY_T \quad (19)$$

where:

Y_T is the value of the reduced variate corresponding to T:

$$Y_T = -ln\{-ln[(T-1)/T]\} \quad (20)$$

Properties

The mean, variance and skewness of the EV1 distribution are:

$E(X) = u + \gamma a$, $\gamma = 0.5772$, is Euler's constant

$Var(X) = \pi^2 a^2/6$

$Skew(X) = 1.14$

Methods of parameter estimation

Following Phien (1987), only four methods of parameter estimation are considered. A brief description of these methods and solution techniques follows.

Method of moments (MMM)

In the MMM, μ and σ are estimated by the sample mean $\bar{x}$ and the sample standard deviation S, respectively; hence:

$$a = 0.7797S \quad , \quad u = \bar{x} - 0.4500S \quad (21)$$

$$X_T = x + SK_T$$

where:

$$\bar{x} = n^{-1}\sum_{i-1}^{n} x_i; \quad S^2 = (n-1)^{-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 \quad (22)$$

n being the sample size, and $K_T = -(\sqrt{6}/\pi)(Y-\gamma)$, is the frequency factor.

Method of maximum likelihood (MML)

The log-likelihood function of a sample $\{X_1, X_2, ..., X_n\}$ is:

$$L = -nlna - \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} \exp e^{-y_i} \quad (23)$$

The maximum likelihood equations are obtained by equating the partial derivatives of L with respect to a and u to zero. As evidenced from the work of Lowery and Nash (1970), the MML estimators are biased. A good correction for the bias has been proposed by Fiorentio and Gabriele (1984) and the resulting corrected maximum likelihood (CML) estimators are as follows:

$$a' = \hat{a}/(1 - 0{,}8/n) = n\hat{a}/(n - 0{,}8)$$
$$u' = a'ln[n/\sum \exp(-x/a')] - 0{,}7a'/n \quad (24)$$

where:

a is the MML estimator of a.

Method of maximum entropy (MME)

In this method, the parameters a and u must be chosen to yield:

$$E[(x-u)/a] = \gamma \text{ and } E\{\exp[-(x-u)/a]\} = 1 \quad (25)$$

In actual situations, the expectations are replaced by the corresponding unbiased estimators to give:

$$(1/n)\sum_{i=1}^{n} y_i = \bar{y} = \gamma \text{ and } (1/n)\sum_{i=1}^{n} V_i = \bar{V} = 1 \quad (26)$$

where:

$$V = \exp(-y) \quad (27)$$

The MME estimates of a and u are obtained by solving Eq. 26.

Method of probability weighted moments (PWM)

The PWM estimators of a and u can be expressed as:

$$a = (2b_1 - b_0)/ln2 \; ; \; u = b_0 - \gamma a \quad (28)$$

where:

b_0 is the sample mean, $\bar{x}$ and b_1 is proposed by Landwehr *et al.* (1979) as:

$$b_1 = (1/n)\sum^{n}(i-1)x_i/(n-1) \quad (29)$$

where:

i is the rank of x_i *in the sequence* $x_1, x_2, ..., x_n$ arranged in ascending order of magnitudes.

Variances-covariances of estimators

The asymptotic variance of X_T is given by:

$$Var(X_T) = Var(u) + 2Cov(a,u)Y_T + Var(a)Y_T^2 \quad (30)$$

and the variances-covariances of the parameter estimators of the four methods were provided by Phien (1987).

Method of maximum likelihood

$$Var(X_T) = (a^2/n)(1.168 + 0.192Y_T + 1.100Y_T^2) \quad (31)$$

Method of maximum likelihood

The variance-covariances of the parameter estimators can be obtained by the inverse of the Fisher information matrix. Insertion of these expressions into Eq. 30 gives:

$$Var(X_T) = (a^2/n)(1.109 + 0.514Y_T + 0.608Y_T^2) \quad (32)$$

Method of maximum entropy

The variance of the estimator of the T-year event is:

$$Var(X_T) = (a^2/n)(1.115 + 0.546Y_T + 0.645Y_T^2) \quad (33)$$

Method of probability weighted moments

$$Var(X_T) = (a^2/n)[(1.1128n - 0.9066) - (0.4574n - 1.1722)\,Y_T + (0.8046n - 0.1855)Y_T^2\,]/(n-1) \quad (34)$$

The log-logistic distribution (LLG)

Definitions

The variable X is defined as being log-logistic if $Y = \ln(X-a)$ has the logistic distribution. The probability density function (pdf) is:

$$f(x) = \frac{[(x-a)/b]^{-1/c}}{c(x-a)\{1 + [(x-a)/b]^{-1/c}\}^2} \quad (35)$$

where:
c is the shape parameter, $c > 0$, b is the scale parameter, $b > 0$, a is the location parameter, $x > a$.

The cumulative distribution function (cdf) is obtained as:

$$F(x) = \frac{1}{1 + [(x-a)/b]^{-1/c}} \quad (36)$$

and then:

$$x = a + b[F/(1-F)]^{-c}$$

The T-year event X_T, in view of Eq. 36, is obtained as: (37)

$$X_T = a + b(T-1)^c \quad (38)$$

Properties

The mean, variance and skewness of the LLG are:

$$E(X) = a + bA(1,c) \quad (39)$$

$$Var(X) = b^2[A(2,c) - A^2(1,c)] \quad (40)$$

$$Skew(X) = \frac{A(3,c) - 3A(2,c)A(1,c) + 2A^3(1,c)}{\sqrt{[A(2,c) - A^2(1,c)]^3}} \quad (41)$$

In the above equations:

$$A(j,c) = \Gamma(1 + jc)\Gamma(1 - jc) \quad (42)$$

Methods of parameter estimation

Method of probability weighted moments (PWM)

There are two PWM methods of estimating the parameters of the LLG distribution by using Eq. 10 with different values of r and s.

The first method (denoted PWM1) is obtained by assigning 1 and 0 to p and r respectively, following Greenwood *et al.* (1979). Then:

$$M_{1,o,s} = E\{X(1 - F(x))^s\} \quad (43)$$

When this is applied to the three parameters of the LLG distribution with s taking the values 0,1,2 (Ahmad *et al.*, 1988) we have:

$$M_1 = aB(1, 1+t) + bB(1+c, 1+t-c), \quad t = 0,1,2 \quad (44)$$

Solving these three equations for c, b, a in that order, we obtain:

$$\left.\begin{aligned} c &= 3 - 2(M_0 - 3M_2)/(M_0 - 2M_1) \\ b &= (M_0 - 2M_1)/cA(1,c) \\ a &= M_0 - bA(1,c) \end{aligned}\right\} \quad (45)$$

Sample values of the probability weighted moments are calculated from the data by Eq. 13, or by using a suitable plotting position, for example:

$$p_i = (i - 0.35)/n\ , \text{ to replace } F(x)$$

Then:

$$m_t = \sum_{i=1}^{n} (1 - p_i)^t x_i/n, \quad t = 0,1,2$$

The PWM estimates of the parameters are found by substituting m_t for $M_t (t=0,1,2)$ in Eq. 45. The gamma functions can readily be approximated accurately by using the algorithm developed by Phien (1988).

The second method (denoted PWM2) is obtained by assigning 1 and 0 to p and s respectively, then:

$$M_{1,r,o} = E[X\{F(x)\}^r] \quad (46)$$

When this is applied to the LLG distribution with r taking the values 0,1,2 we have:

$$M_t = \frac{a}{t+1} + bB(c+1+t, 1-c)\ , \quad t = 0,1,2 \quad (47)$$

Solving these three equations for c, b, a we obtain:

$$\left.\begin{aligned} c &= \frac{2(3M_2 - 2M_1)}{2M_1 - M_0} - 1 \\ b &= \frac{2M_1 - M_0}{c^2\Gamma(1-c)\Gamma(c)} \\ a &= M_0 - \frac{2M_1 - M_0}{c} \end{aligned}\right\} \quad (48)$$

By substituting the value of m_t for M_t in Eq. 48 as above we can find the PWM estimators.

Method of maximum likelihood (MML)

From the pdf (Eq. 35), with $y = (x-a)/b$ and $z = y^{1/c}$ then:

$$f(x) = \frac{z^{c+1}}{bc(1+z)^2} \quad (49)$$

The log likelihood function of the LLG distribution is obtained from Eq. 49, as:

$$L = (1+c)\sum_{i=1}^{n} ln(z_i) - nln(b) - nln(c) - 2\sum_{i=1}^{n} ln(1+z_i) \quad (50)$$

In this study, the direct search with systematic reduction of the size of search region (DSSRSSR) algorithm (Ong and Lee, 1986) was modified for use in solving for the values of a, b and c from Eq. 50.

Variances-covariances of estimators

Method of probability weighted moments

For the LLG distribution, it is too complicated to derive the formulas for finding the variances-covariances of the estimators. Because of this reason, the Jack-knife technique is used to estimate these variances and covariances.

Method of maximum likelihood

The variances-covariances of the MML estimators can be obtained by inversion of the observed information matrix.

The variance of X_T can readily be computed by Eq. 16, with:

$$\partial X_T / \partial a = 1$$
$$\partial X_T / \partial b = (T - 1)^c$$
$$\partial X_T / \partial c = b(T - 1)^c ln(T - 1)$$

TABLE 1
DESCRIPTIVE STATISTICS OF ANNUAL MAXIMUM SERIES

No.	Station	N	Mean (m^3/s)	SD (m^3/s)	Skewness
1	B1A	26	136.87	52.75	-1.0450
2	B5	24	165.95	281.85	3.5253
3	B6	26	247.09	227.59	0.8324
4	C2	32	2644.22	1064.95	0.4519
5	C7	34	1174.94	157.80	0.1816
6	C13	31	2348.97	1138.65	0.2612
7	E1	31	692.35	425.28	0.5084
8	E2	31	1255.03	948.95	3.5287
9	E5A	30	315.10	154.36	1.9826
10	E8A	33	589.15	274.49	0.1416
11	E9	21	545.18	672.75	2.3659
12	E16A	23	454.64	518.65	1.9458
13	E18	34	890.15	568.67	3.4979
14	E32A	27	414.74	400.69	2.0285
15	E33A	24	338.63	175.83	1.7232
16	E49	21	32.86	30.33	1.6939
17	G2A	27	463.63	307.26	-0.2838
18	K10	23	1615.30	839.53	0.2903
19	K11	23	1827.17	737.61	0.3992
20	K17	22	274.92	348.94	2.1596
21	K22A	22	170.93	101.33	1.3511
22	KGT3	44	775.48	173.05	-0.1409
23	KGT10	21	308.62	188.49	2.0845
24	KGT12	21	187.48	77.94	2.5541
25	KGT19	22	124.45	261.84	3.7057
26	KH18	30	238.23	247.40	3.2482
27	M2	37	288.49	321.29	1.6296
28	M5	31	1612.19	1586.53	3.5071
29	M7	37	3232.35	2014.67	2.0830
30	M26	33	256.79	266.58	1.9255
31	M32	20	352.90	135.48	0.8465
32	M66	23	112.36	73.42	1.3634
33	M80	23	474.74	379.33	2.5365
34	N1	52	1307.92	587.65	0.4416
35	N5A	38	1261.32	370.95	-0.3377
36	N7	42	1216.95	271.27	-0.9551
37	N17	22	417.68	429.56	1.9430
38	N22	23	429.17	221.72	0.5969
39	N24	23	339.30	169.98	0.5722
40	N35	21	2006.86	1032.06	2.1772
41	P1	33	441.22	115.76	0.7787
42	P2A	35	1199.23	859.61	2.3679
43	P4A	32	206.13	129.26	2.1949
44	P5	37	209.20	66.48	0.3918
45	P7A	27	1231.63	492.90	0.4318
46	P14	33	429.45	211.14	0.8438
47	P19A	29	761.24	298.22	1.6889
48	P21	34	55.54	16.25	0.3361
49	P23	32	195.68	78.55	0.4817
50	PR3A	20	266.61	367.34	2.0189
51	S2	59	557.03	241.04	1.0869
52	TL1	20	121.92	101.85	1.7291
53	W1	36	439.33	220.69	1.0612
54	W1A	20	272.55	152.88	1.1902
55	W3A	21	652.10	338.80	0.5551
56	X40	20	951.70	798.14	1.9556
57	X44	20	286.21	121.28	-0.4645
58	X56	21	394.81	324.48	2.3424
59	X67	21	82.15	21.82	0.0993
60	Y3A	20	1111.25	325.81	-0.4338
61	Y6	35	1376.14	638.59	0.5386
62	Y13	30	167.95	187.71	2.0292
63	Y14	23	1390.09	799.29	1.5515
64	Z5	20	186.61	171.37	2.5582

Applications

Data employed

The data sets used in this analysis comprise the annual maximum series from 64 streamflow stations located in different regions of Thailand. All these stations are under the responsibility of the Royal Irrigation Department and are selected according to the following two criteria:

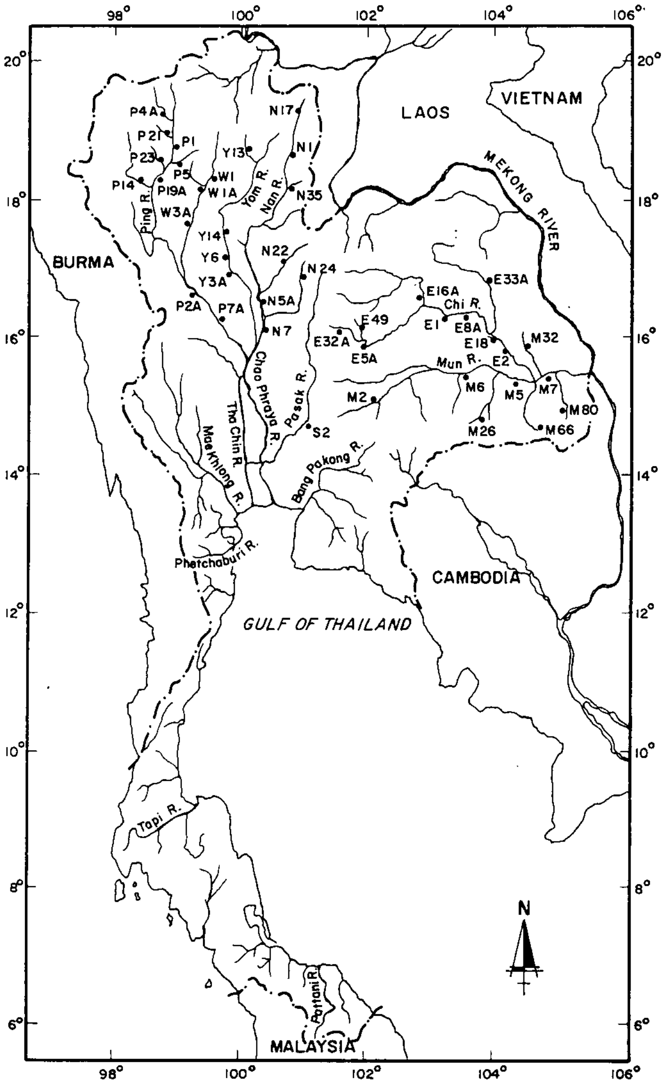

Figure 1
Locations of some selected stations

TABLE 2
ESTIMATED VALUES OF PARAMETERS AND TEST STATISTICS STATION S2)

Dist.	Methods	Estimated values			Test Statistics	
		a	b or u	c	Chi.	KS.
GEV	PWM	194,13	0,0361	451,700	2,10	0,048
	MML	197,79	0,0502	453,7600	1,63	0,054
EV1	MMM	187,94	488,5600	na	0,92	0,046
	MME	191,78	446,3300	na	1,15	0,046
	PWM	187,96	448,5400	na	0,92	0,046
	MML	198,63	447,2200	na	2,81	0,049
LLG	PWM1	-296,42	820,8100	0,1534	1,627	0,056
	PWM2	-329,50	855,3700	0,1469	2,339	0,056
	MML	-295,22	821,1400	0,1522	2,814	0,056

Note: na = not available
Chi = Chi-square statistic
KS = Kolmogorov-Smirnov statistic

- The flows at these stations are natural. They have not been regulated by a reservoir upstream.
- Each station must have at least 20 years of record.

Table 1 shows the sample statistics of these data sets.

Data processing

Several computer programs were developed based on the formulas presented before. In parameter estimation, the estimates obtained by the PWM are used as initial values for the MML (and MME for the EV1 distribution). The magnitude and standard deviation of the T-year flood were computed for T = 100, 200, 500 and 1 000 years. Typical results are shown in Tables 2 to 5.

Results and conclusions

By repeating the same analysis that gave rise to the results collected in Tables 2-5, the results for all the stations were obtained. By examining the computed results the following observations could be made:

- Among the distributions considered, the GEV can be used to fit these flood data more frequently than the EVI and LLG. This is asserted by the summarised results in Table 6.
- For the GEV, the variance (or standard deviation) of the T-year flood obtained by the Jack-knife method has values that are closer to those obtained by inverting the observed information matrix than those obtained by inverting Fisher information matrix. This, to some extent, confirms the results obtained previously by Hinkey (1978) and Phien and Fang (1989).
- The PWM provides estimators with smaller bias when compared to other methods. Particularly, it gives unbiased estimators for the EV1 distribution.
- For the samples considered, the MML did not give the smallest standard deviation for the T-year flood, according to the results obtained by the Jack-knife method.

TABLE 3
T-YEAR FLOOD MAGNITUDE (STATION S2)

Dist.	Methods	100	200	500	1000
GEV	PWM	1274,45	1387,40	1532,13	1638,38
	MML	1266,25	1373,61	1509,62	1608,34
EV1	MMM	1313,12	1442,86	1616,36	1746,72
	MME	1328,57	1461,99	1638,01	1771,04
	PWM	1313,18	1443,93	1616,44	1746,82
	MML	1360,94	1499,12	1681,42	1891,19
LLG	PWM1	1364,47	1552,21	1832,16	2071,30
	PWM2	1350,97	1532,55	1801,89	2030,79
	MML	1357,02	1542,21	1818,08	2053,50

TABLE 4
T-YEAR EVENT VALUES FROM MML ESTIMATORS (STATION S2)

Dist.	Methods	100		200		500		1000	
		XT	**SD(XT)**	**XT**	**SD(XT)**	**XT**	**SD(XT)**	**XT**	**SD(XT)**
GEV	Obser.	1266,25	134,29	1373,61	168,46	1509,62	220,01	1608,34	263,37
	Fisher.	1266,25	26,16	1373,61	26,13	1509,62	26,08	1608,34	26,03
EV1	Fisher.	1360,94	104,53	1499,12	118,17	1681,42	136,29	1819,19	150,04
LLG	Obser.	1357,02	40,02	1542,21	44,71	1818,08	52,15	2053,51	58,80

TABLE 5
T-YEAR EVENT VALUES FROM JACK-KNIFE TECHNIQUE (STATION S2)

Dist.	**Methods**	100			200			500			1000		
		XT	**SD(XT)**	**Bias(XT)**	**XT**	**SD(XT)**	**Bias(XT)**	**XT**	**SD(XT)**	**Bias(XT)**	**XT**	**SD(XT)**	**Bias(XT)**
GEV	PWM	1274,70	198,90	13,00	1387,90	256,60	26,40	1533,00	342,80	51,60	1639,70	414,70	77,10
	MML	1264,80	245,30	-83,80	1371,90	308,20	-99,00	1507,60	397,20	-115,30	1606,20	468,00	-124,00
EV1	MMM	1312,96	136,21	-9,12	1443,68	156,16	-10,69	1616,14	182,59	-12,74	1746,48	202,64	-14,33
	MME	1328,02	104,17	-32,20	1461,34	118,66	-37,77	1637,23	137,93	-45,11	1770,17	152,58	-50,67
	PWM	1313,18	113,39	-0,03	1443,93	129,18	-0,02	1616,44	150,15	-0,02	1746,82	166,07	-0,02
	MML	1360,83	106,93	-6,36	1498,99	122,16	-7,38	1681,27	142,41	-8,68	1819,03	157,79	-9,66
LLG	PWM1	1364,81	195,23	19,45	1552,81	264,83	34,54	1833,28	383,80	64,91	2072,99	497,86	98,27
	PWM2	1351,10	198,59	7,90	1532,86	269,15	18,27	1802,59	389,16	40,69	2031,94	503,63	66,55
	MML	1357,59	190,87	33,30	1543,16	258,66	55,18	1819,74	374,69	96,19	2055,90	486,02	139,04

TABLE 6 SUMMARIZED RESULTS FOR THE PARAMETER ESTIMATES			
Dist.	Methods	No. of rejected cases	
		Chi.	KS
GEV	PWM	5	2
	MML	4	-
EV1	MMM	15	6
	MME	14	7
	PWM	16	6
	MML	15	3
LLG	PWM1	8	-
	PWM2	7	-
	MML	9	1

Note: There are 4 data sets for which PWM1 gives a>x(1)
There are 8 data sets for which PWM2 gives a>x(1)
x(1) is the minimum value of the data set

- When estimated by the MML, the standard deviation of the T-year flood is smaller for the LLG than for the GEV. This is asserted by the Jack-knife results as well as the results based on the information matrix. This indicates that when the LLG is applicable, more efficient estimators are expected to result from the LLG than from the GEV when the MML is used.

From the above results, the following conclusions could be drawn:

- For at-site analysis of flood data in Thailand, the GEV distribution should be used. This distribution can provide a good fit in many cases. Moreover, for this distribution, the PWM should be used because it is less biased and more efficient (expressed by smaller values of the bias and standard deviation of the T-year flood).
- The LLG can also be used for flood frequency analysis in Thailand. It gives a good fit to many cases as well. In terms of the bias incurred, the PWM should be used for parameter estimation.
- The EV1, having only two parameters, cannot be used to satisfactorily represent the flood data in a larger number of cases. As such, it is not suitable for flood frequency analysis in Thailand.

References

AHMAD, MI, SINCLAIR, CD and WERRITTY, A (1988) Log-logistic flood frequency analysis. *Journal of Hydrology* **98** 205-224.

EFRON, B (1982) The Jackknife, the Bootstrap and other Resampling Plans. *SIAM,* Philadelphia.

FIORENTIO, M and GABRIELE, S (1984) A correction for the bias of maximum likelihood estimators of Gumbel parameters. *Journal of Hydrology* **73** 39-49.

GREENWOOD, JA, LANDWEHR, JM, MATALAS, NC and WALLIS, JR (1979) Probability weighted moments: Definition and relation to parameters of several distributions expressable in inverse form. *Water Resources Research* **15**(5) 1049-1054.

HINKEY, DV (1978) Likelihood inference about location and scale parameters. *Biometrika* **65** 253-261.

HOSKING, JRM (1985) Algorithm AS 215. Maximum-likelihood estimation of the parameters of the generalized extreme-value distribution. *Applied Statistics* **34** 301-310.

HOSKING, JRM, WALLIS, JR and WOOD, EF (1985) Estimation of the generalized extreme value distribution by the method of probability-weighted moments. *Technometrics* **27**(3) 251-261.

JENKINSON, AF (1955) The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of the Royal Meteorological Society* **18** 158-171.

LANDWEHR, JM, MATALAS, NC and WALLIS, JR (1979) Probability weighted moments compared with some traditional techniques in estimating Gumbel parameter and quantiles. *Water Resources Research* **15**(5) 1055-1064.

LOWERY, MD and NASH, JE (1970) A comparison of methods of fitting the double exponential distribution. *Journal of Hydrology* **10** 259-275.

NERC (National Environment Research Council) (1975) *Flood Studies Report* **4**. London, Authon.

ONG, SL and LEE, GS (1986) Program DSSRSSR for the solution of nonlinear optimization models. *Adv. Eng. Software* **8** 134-140.

PHIEN, HN (1987) A review of methods of parameter estimation for the extreme value type 1 distribution. *Journal of Hydrology* **90** 251-268.

PHIEN, HN (1988) A Fortran routine for the computation of gamma percentiles. *Adv. Eng. Software* **10**(3) 159-164.

PHIEN, HN and FANG, TE (1989) Maximum likelihood estimation of the parameters and quantiles of the general extreme-value distribution from censored samples. *Journal of Hydrology* **105** 139-155.

PHIEN, HN and HIRA, MA (1983) Log Pearson type 3 distribution: parameter estimation. *Journal of Hydrology* **64** 25-27.

PHIEN, HN and HSU, LC (1985) Variance of the T-year event in the Log Pearson Type 3 distribution. *Journal of Hydrology* **77** 141-158.

PHIEN, HN and YANG, JH (1988) Application of LP3 distribution to censored samples. *Proc. Int. Seminar on Hydrology of Extremes.* Roorkee, December 1-3. 77-78.

PRESCOTT, P and WALDEN, AT (1980) Maximum likelihood estimation of the parameters of the generalized extreme-value distribution. *Biometrika* **67**(3) 723-724.

US Water Resources Council (1967) A uniform technique for determining flood flow frequencies. Washington DC, Bulletin No. 15.

YANG, MCK and ROBINSON, DH (1986) *Understanding and Learning Statistics by Computer.* World Scientific Publishing Co. (Pty) Ltd., Singapore.