The Extremes of the Extremes: Extraordinary Floods (Proceedings of a symposium held at Reykjavik, Iceland, July 2000). IAHS Publ. no. 271, 2002.

Study of parameter estimation methods for Pearson-III distribution in flood frequency analysis

CHEN YUANFANG, HOU YU
Department of Hydrology, Hohai University, Nanjing 210024, China
e-mail: yf.chen@jlonline.com

PIETER VAN GELDER
Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft 2600 GA, The Netherlands
e-mail: p.vangelder@ct.tudelft.nl

SHA ZHIGUI
Yangtze Water Resources Commission, Wuhan 430010, China

Abstract Based on the Monte-Carlo method, the main parameter estimation methods for the Pearson-III (P-III) distribution are comprehensively compared, including traditional moments (MOM), curve-fitting, probability weighted moments (PWM) and weighted function moments (WF) which has been developed recently in China. In the comparison, series with historical floods are generated not only in the Chinese given-number model, but also in the Stedinger–Hirsch model (threshold model). Two criteria for evaluation of an estimation method are considered. One criterion is the bias and efficiency of the parameters and design flood values (quantiles) with a fixed probability, the other is the bias and efficiency of the probability of failure of a design flood. Many calculations show that the PWM and a curve-fitting method with absolute norm are the two best estimation methods of the P-III distribution in flood frequency analysis.
Key words Pearson-III; flood frequency analysis; Monte-Carlo; probability weighted moments; historical flood; bias and efficiency; flood sampling models; curve-fitting; comparison; quantile; probability of failure

INTRODUCTION

The curve-fitting method by eye estimation is widely used in practical flood frequency analysis in China at present because of its flexibility, however, the weakness of this method is obvious and serious—different results of the parameter estimation are obtained by different persons and at different times. It is thus not an objective method. Other objective parameter estimation methods with good statistical performance in design flood calculation should therefore be considered. Considerable research on other parameter estimation techniques has been carried out in the past 20 years in China. Cong (1980) applied the Monte-Carlo method to study the statistical performance of the method of moments (MOM), maximum likelihood (ML) method and several curve-fitting methods by optimization. Results show that the curve-fitting method (FIT) with the Weibull plotting position formulae and absolute norm (for its goal function see Cong (1980)) is the best of all the estimation methods considered. The ML method often shows (in the case of the P-III distribution) a slow convergence (or even divergence) in the parameter estimation. Cohn & Stedinger (1987) and Cohn *et al.* (1997) investigated the ML method in a framework with historical information and could therefore improve the performance of the ML estimation procedure. The evaluation standard in van Gelder & Vrijling (1997) is the bias and efficiency of the parameters and quantile estimates. Ding (1988) has used the same evaluation standard to compare the PWM and the FIT method. Bobée & Ashkar (1991) give a complete overview of the gamma distributions (in particular the P-III distribution) including many parameter estimation methods for flood frequency analysis. Ding's (1988) results show that PWM is a good estimation method for P-III. Ma (1984) has developed an alternative estimation method called the weighted function moments method (WF) for flood series with a single sample. Chen (1999) has developed effective parameter estimation formulae, which may consider historical flood information.

A new evaluation standard, expected probability, has been presented in recent years. Several researchers have investigated it in China. The results show that it is reasonable to use expected probability to evaluate the quantile statistical performances.

As for the historical flood sampling, Hirsch & Stedinger (1985) put forward a new model, the so-called threshold sampling model. Cong (1980) suspected that the new sampling model might be better than the Chinese given-number model.

So it is very interesting to evaluate the main estimation methods for the P-III distribution with the new evaluation standard and with the new historical flood sampling model.

PARAMETER ESTIMATION METHODS TO BE STUDIED

In the WF method let X represent a series with historical floods in which the maximum return period is N; the length of recorded series (systematic) is n, the number of historical floods is a, the number of historical floods in the recorded series is l, $\{x_m, m = 1, 2, \ldots, (n + 1 - a)\}$ is the series in descending order (from highest to lowest). The formulae of the mean EX and the coefficient of variations, Cv with the above historical information is the same as MOM (Chen, 1999), and that of the coefficient of skewness Cs is as follows:

$$Cs = 4\sigma\, E/G \tag{1}$$

where:

$$E = \frac{1}{N}\left[\sum_{m=1}^{a}(EX - x_m)\,\varphi(x_m) + \frac{N-a}{n-l}\sum_{m=a+1}^{a+n-l}(EX - x_m)\,\varphi(x_m)\right] \tag{2}$$

$$G = \frac{1}{N}\left[\sum_{m=1}^{a}(EX - x_m)^2\,\varphi(x_m) + \frac{N-a}{n-l}\sum_{m=a+1}^{a+n-l}(EX - x_m)^2\,\varphi(x_m)\right] \tag{3}$$

$\varphi(x)$ is the probability density function of the normal distribution with mean EX and standard deviation σ, $\sigma = EX\,Cv$.

The curve-fitting method (FIT) is widely used for the study of flood frequency analysis in China. Its detailed procedure of parameter estimation is treated in Cong (1980). Its plotting position formula is the expected value (Weibull) formula, which is recommended by Cong (1980). The goal function of optimization:

$$\Delta = \sum_{m=1}^{n-l+a} \left| \hat{x}_{p_m} - x_{p_m}^0 \right| \tag{4}$$

where $\hat{x}_{p_m}$ is the quantile estimated of p_m. $x_{p_m}^0$ is the quantile of the population of p_m. The goal of optimization is to minimize Δ. *EX* is still estimated by the MOM.

For the formulae of PWM the reader is referred to Ding (1988).

EVALUATION STANDARD OF PARAMETER ESTIMATION METHOD

The performance of a parameter estimation is evaluated by calculating the bias and efficiency of the parameter, the quantile and the probability of failure. The mean of the probability of failure is the so-called expected probability.

The bias and efficiency of the parameters and the probability of failure are indicated by the mean and root mean squared error (r.m.s.e.) of *Ns* Monte-Carlo generating samples respectively. When the mean of a parameter is equal to its population value, the estimation of the parameter is considered as unbiased, and the smaller the r.m.s.e. of a parameter is, the more efficient is the estimation of the parameter. The same is valid for the probability of failure.

For clearness, the efficiency of the quantile S_{X_p} and the bias B_{X_p} are expressed by the following formulae:

$$S_{X_p} = \sqrt{\sum_{i=1}^{N_S} \left(\hat{x}_p^i - x_p^0\right)^2 / \left(N_S \cdot x_p^{0\,2}\right)} \times 100\% \qquad B_{X_p} = \frac{\sum_{i=1}^{Ns} \left(x_p^i - x_p^0\right)}{N_s x_p^0} \times 100\% \tag{5}$$

In which p is the design probability (p = 0.01%, 0.1%, 1%), x_p^0 is the population quantile with p, where $P(X \geq x_p^0) = p$, x_p^i is the quantile estimated by the *i*th generating sample (i = 1, 2, …, *Ns*), *Ns* is the number of generating samples, in this paper $N_s = 500$.

MONTE-CARLO EXPERIMENT SCHEMES DESIGN

Schemes design for the Chinese given-number sampling model

(a) *A simple sample*:
n = 30, 50, $EX_0 = 1$, Cv_0 = 0.5, 1, Cs_0/Cv_0 = 3, 4. So there are in total 2 × 2 × 2 = 8 schemes.

(b) *A sample with a historical flood*:
N = 100, n = 30, 50, a = 1, 5, EX_0 = 1, Cv_0 = 0.5, 1, Cs_0/Cv_0 = 3, 4. So there are 2^4 = 16 schemes in total.

Schemes design for threshold sampling model

The main difference of the two sampling models is expressed as follows: for the given-number sampling, the number of historical floods, a, must be given before sampling a historical flood series; for the threshold sampling, it is not necessary to give the number a, but it needs a threshold value X_0. If a sampling flood value exceeds X_0, it is treated as a historical flood. In other words, in the threshold model, it is necessary to supply the sampling parameters (N, n and X_0), the number a, is not necessary. In the paper N = 100, n = 30, 50, X_0 takes x_p^0, p = 1/80, x_p^0 is the population quantile, EX_0 = 1, Cv_0 = 0.5, 1, Cs_0/Cv_0 = 3. There are only 2 × 2 = 4 schemes.

For clearness, the statistical performance of the MOM is also calculated and presented in this paper.

RESULTS AND ANALYSIS

The results show that MOM is not a good estimation method of the P-III distribution because it is seriously negatively biased in the quantile estimation. This was also the finding of Cong (1980). The other three estimation methods are compared in the following by considering different historical flood sampling models. Partial results are listed in Tables 1 and 2.

Comparison of estimation methods in the given-number model

(a) *A simple sample*:
For the statistical performance of the parameters, ECv (mean of Cv estimated from the sample) by WF is smaller than the population value Cv_0, especially when Cv_0 is relative big, ECv is much smaller than Cv_0. ECs (the same meaning as ECv) by

Table 1 Results of bias and efficiency of probability of failure.

Population parameters:					Estimation method	EP_1	EP_2	EP_3	SP_1	SP_2	SP_3
Cv_0	Cs_0	N	n	a							
0.50	1.50	5	30	0	WF	0.018 284	0.004 467	0.001 488	0.020 036	0.008 267	0.003 815
0.50	1.50	30	30	0	PWM	0.016 746	0.003 900	0.001 234	0.007 246	0.003 187	0.003 187
0.50	1.50	30	30	0	FIT	0.015 227	0.003 392	0.001 097	0.016 954	0.007 278	0.003 739
0.50	1.50	100	50	1	WF	0.013 428	0.002 235	0.000 494	0.010 603	0.003 340	0.001 151
0.5	1.50	100	50	1	PWM	0.012 549	0.001 949	0.000 394	0.009 673	0.002 800	0.000 871
0.50	1.50	100	50	1	FIT	0.012 475	0.001 985	0.000 411	0.010 298	0.003 016	0.000 914

Note: EP and SP are the mean and the standard deviation of the probability of failure in N_s generating samples.
P_1 = 0.01%, P_2 = 1%, P_3 = 1%, EX_0 = 1.0.

WF is almost equal to Cs_0, so Cs may be considered as unbiased. In most cases, bias and efficiency of Cv and Cs by FIT and PWM is smaller than WF. In most schemes, the bias of Cv by PWM is better than FIT. The bias of Cs by FIT and PWM is almost the same. When Cv_0 is small, the efficiency of Cv by PWM is better than by FIT, when Cv_0 is relatively large, the results are the opposite. In most schemes, the efficiency of Cs by FIT is better than by PWM.

As far as the quantiles are concerned; when Cv_0 and Cs_0 are small, EX_p (the same meaning as ECv) by WF is slightly smaller than x_p^0; when Cv_0 and Cs_0 are relatively large, it is still smaller than x_p^0, (usually $|BX_p| < 5\%$). In most cases, the efficiency of X_p by WF is worse than FIT, and similar to PWM. EX_p by PWM is almost equal to x_p^0 i.e. unbiased, but EX_p by FIT is positively biased ($|BX_p| < 5\%$).

As far as the probability of failure is concerned, the quantiles by all three methods cannot reach the design standard. From the point of view of bias and efficiency of the probability of failure, the results of WF are the worst of the above three methods. In most cases FIT is better than PWM.

(b) *A sample with historical flood information*:

As for the parameters, the conclusion of WF is similar to that in the simple sample. Bias and efficiency of Cv and Cs by FIT and PWM are fairly good. If Cv_0, Cs_0 are relative small, results of Cv and Cs by FIT are slightly worse than PWM. When Cv_0 and Cs_0 are relative large, results of PWM are much better than FIT.

Table 2 Results of bias and efficiency of parameters and quantiles.

Population parameter:					Method	*EEX*	*ECv*	*ECs*	*SEX*	*SCv*	*SCs*	BX_{p1}	BX_{p2}	BX_{p3}	SX_{p3}	SX_{p2}	SX_{p1}
Cv_0	Cs_0	*N*	*n*	*a*													
0.50	1.50	30	30	0	MOM	1.00	0.49	1.15	0.09	0.08	0.64	–8.96	–7.32	–4.91	17.66	22.26	25.59
0.50	1.50	30	30	0	WF	1.00	0.49	1.49	0.09	0.08	0.65	–1.15	–0.70	–0.32	18.64	24.35	28.50
0.50	1.50	30	30	0	PWM	1.00	0.50	1.45	0.09	0.08	0.54	0.51	0.82	1.05	18.87	23.95	27.56
0.50	1.50	30	30	0	FIT	1.00	0.52	1.48	0.09	0.08	0.53	2.18	2.94	3.46	18.70	23.56	27.05
1.00	3.00	30	30	0	MOM	1.00	0.93	2.13	0.19	0.21	1.16	–18.97	–15.97	–11.15	32.18	37.24	40.29
1.00	3.00	30	30	0	WF	1.00	0.93	3.05	0.19	0.21	1.07	–4.28	–2.98	–2.10	32.23	39.29	42.65
1.00	3.00	30	30	0	PWM	1.00	0.99	2.97	0.19	0.21	0.72	1.26	2.26	2.88	33.60	39.21	43.65
1.00	3.00	30	30	0	FIT	1.00	1.00	2.99	0.19	0.19	0.63	–0.29	0.53	1.04	29.85	34.93	37.91
0.50	1.50	100	50	1	MOM	1.00	0.49	1.36	0.07	0.05	0.49	–3.45	–2.82	–1.89	12.56	16.34	19.21
0.50	1.50	100	50	1	WF	1.00	0.49	1.53	0.07	0.05	0.50	0.01	0.50	0.89	13.34	17.81	21.15
0.50	1.50	100	50	1	PWM	1.00	0.50	1.50	0.07	0.06	0.37	0.91	1.24	1.48	12.88	16.17	18.57
0.50	1.50	100	50	1	FIT	1.00	0.51	1.49	0.07	0.06	0.36	1.31	1.64	1.87	13.03	16.11	18.34
1.00	3.00	100	50	1	MOM	1.00	0.97	2.64	0.1	0.14	0.92	–8.36	–7.32	–5.34	21.52	25.96	28.83
1.00	3.00	100	50	1	WF	1.00	0.97	3.01	0.13	0.14	0.82	–2.24	–1.64	–1.22	21.94	26.83	29.96
1.00	3.00	100	50	1	PWM	1.00	1.01	3.04	0.13	0.14	0.49	1.21	2.02	2.52	22.16	25.46	27.44
1.00	3.00	100	50	1	FIT	1.00	1.02	3.06	0.13	0.14	0.51	2.33	3.29	3.87	21.42	24.94	27.10
0.50	1.50	100	50	*	WF	1.00	0.50	1.52	0.07	0.05	0.48	–0.01	0.41	0.74	12.51	16.41	19.32
0.50	1.50	100	50	*	PWM	1.00	0.50	1.50	0.07	0.05	0.37	0.89	1.21	1.44	12.11	15.21	17.47
0.50	1.50	100	50	*	FIT	1.00	0.51	1.51	0.07	0.05	0.37	1.77	2.32	2.69	13.18	16.09	18.22
1.00	3.00	100	50	*	WF	1.00	0.98	2.99	0.13	0.14	0.80	–2.22	–1.84	–1.55	20.81	24.96	27.65
1.00	3.00	100	50	*	PWM	1.00	1.01	3.04	0.13	0.14	0.49	1.25	2.06	2.55	21.07	24.17	26.04
1.00	3.00	100	50	*	FIT	1.00	1.03	3.07	0.13	0.14	0.50	4.02	5.09	5.72	19.62	22.99	25.04

Note: * represents the results of the threshold model; *EEX*, *SEX* are the mean and the standard deviation of the mathematical expectation in N_s generating samples; the meaning of SX_p and BX_p refers to the equation (5).

With regard to the quantiles, when Cv_0 and Cs_0 are relative small, EX_p by WF is almost equal to x_p^0, when Cv_0 and Cs_0 are relative large, X_p by WF is still negatively biased. The efficiency of quantile estimations by WF is slightly worse than PWM. The quantiles by PWM are slightly positively biased but $|BX_p| < 5\%$. Bias of the quantiles by FIT is worse than PWM especially when Cv_0 and Cs_0 and the number *a* are relatively large. When *n* is relative large, the efficiency of quantiles estimated by PWM is almost the same as FIT, but when *n* is small, PWM is better than FIT.

As with respect to the probability of failure, the expected probabilities of the above three methods are closer to the design frequencies, in comparison with the case of a simple sample. However, all expected probabilities cannot attain the design standard. FIT is the best one and WF is the worst according to the bias and efficiency of the probability of failure.

Comparison of estimation methods in the threshold model

As far as the parameters are concerned, *ECv* estimated by WF is slightly smaller than Cv_0 in all schemes, however, *ECs* is very close to Cs_0. Efficiency of *Cs* by WF is worse than FIT and PWM. *ECs* by PWM is almost equal to Cs_0, so it can be considered as unbiased. The efficiency of *Cs* by PWM is better than FIT, especially when *n* is relative small. As for the quantiles, when Cv_0 and Cs_0 are relative small, EX_p by WF is very close to x_p^0, but when Cv_0 and Cs_0 are relative large, X_p is slightly negatively biased. The efficiency of X_p by WF is similar to FIT, but it is worse than PWM. X_p by PWM is slightly positively biased, but BX_p is less than 5%. If *n* is relatively large, the efficiency of X_p by PWM is almost the same as FIT, but when *n* is smaller, PWM performs better than FIT.

With regard to the probability of failure, all results are consistent with those of the given-number model.

In brief, WF performs better than MOM, but it is still worse than PWM and FIT. For PWM and FIT, each has its own advantages in different evaluation standards. In terms of bias and efficiency of quantiles, the bias of PWM is smaller than FIT. When *n* is relatively large, the efficiency of X_p by PWM and FIT is similar, however, when *n* is relatively small and Cv_0 and Cs_0 are relatively large, PWM is better than FIT in efficiency. When the number *a* is equal to 0, the efficiency of X_p by FIT is better than PWM. The results of the probability of failure by FIT are better than PWM.

CONCLUSIONS

According to the above results and the conclusion obtained by Chen (1990), X_p estimated by FIT by eye estimation is usually seriously positively biased. Our conclusions are expressed as follows:

– Both PWM and FIT are parameter estimation methods for the P-III distribution with good statistical performance and objectivity. We propose here to use them as the main parameter estimation methods in a practical flood frequency analysis.

- WF may overcome the shortcomings of MOM—estimation of *Cs* by MOM is seriously negatively biased, but estimation of *Cv* gives the same performance as MOM. It would be useful to make a further study for improving the estimation of *Cv*.
- The historical sampling model has a relative greater effect on FIT.
- When the sample size is small, all expected probabilities of the quantiles estimated by the above methods are larger than the design frequencies. To consider an expected probability is still a problem to be addressed.

REFERENCES

Bobée, B. & Ashkar, F. (1991) *The Gamma Family and Derived Distributions Applied in Hydrology*. Water Resources Publication. Littleton, Colorado, USA.

Cong, S. Z. (1980) Study on parameter estimation method in flood frequency analysis by using Monte-Carlo method. *J. Wat. Conservancy* **3**, 1–15.

Cong, S. Z. (1987) Present situation and prospect of flood frequency analysis (in Chinese). *J. Hydrol.* **96**, 52–58.

Chen, Y. F. (1999) A new weighted function moments method to estimate parameters of P-III distribution with historical floods In: *Proc. of International Symposium on Flood and Droughts*. Hohai University Press, Nanjing.

Chen, Y. F (1990) Statistic proof of several problems on curve-fitting methods In: *Proc. of Symposium on National Scientist of Water Science.* Water Power Press, Beijing.

Cohn, T. A., Lane, W. L. & Baier, W. G. (1997) An algorithm for computing moments-based flood quantile estimates when historical information is available. *Wat. Resour. Res.* **33**(9), 2089–2096.

Cohn, T. A. & Stedinger, J. R. (1987) Use of historical information in a maximum-likelihood framework. *J. Hydrol.* **96**, 215–223.

Ding, J. (1988) A new parameter estimation method of P-III distribution—probability weighted moments method. *J. Sichun Univ.* **2**, 15–22.

Hirsch, R. M. & Stedinger, J. R. (1985) Probability plotting position formulae for flood record with historical information. In: *Proc. Sino-American Symposium on Hydrological Extreme Values* (Nanjing, China).

Ma, X. F. (1984) Weighted function moments method to estimate parameters for P-III distribution (in Chinese). *J. Hydrol.* **4**, 1–9.

van Gelder, P. H. A. J. M. & Vrijling, J. K. (1997) A comparative study of different parameter estimation methods for statistical distribution functions in civil engineering applications. *Structural Safety and Reliability* **1**, 665–668.