Journal of Engineering and Development, Vol. 14, No. 4, December (2010) ISSN 1813-7822

Comparing Of Standard And Recent Extreme Probable Distributions Corresponding To Iraqi Streams Flow

Dr. Sadiq S. M.
Al-Mustansiriya University
Collage of Engineering
Environment Department

M.Sc. Safaa N. H.
Al-Mustansiriya
University Collage of
Environment Department

Abstract

Annual extreme flood distributions were statistically analyzed and studied to determine a best probability distribution that describes the observed annual maximum peak flows for six streams of Tigris River in Iraq. This is done through fitting of three different probability distributions and estimating the goodness of the fit to confirm a hydrologic series of 39 annual maximum peak flows of the six streams. In this study, two standard probability distributions were used as well as a new one model. The two standard distributions are the Extreme Value Type I distribution (the Gumbel's distribution) and the log Pearson type 3 distribution. While the new model is the Demetris distribution which is a recent model developed by Demetris Koutsoyiannis depending on theoretical and empirical grounds for many hydrologic series in the world. The analyses showed that the Demetris distribution is significantly overestimating the actual flows for small return periods of about less than 10 years, where it seems approximately as the two standard distributions for a return period greater than 10 years. Also, the Chi-square test indicated that all the distributions represent a very good fitting for the observed hydrologic series of the six streams but Log pearson III distribution is found to be the best fit one.

الخلاصـــــــة

في هذا البحث, تم دراسة وتحليل التوزيعات الاحتمالية للقيم الهايدرولوجية المتطرفة لتحديد التوزيع الاحتمالي الافضل الذي يمكن استخدامه في تمثيل السلاسل الزمنية الهايدرولوجية للتصاريف السنوية العظمى المسجلة خلال 39 سنة و لستة روافد رئيسية من روافد نهر دجلة في العراق.

لقد تم استخدام توزيعين قياسيين من التوزيعات المتطرفة المشهورة وهما توزيع القيم المتطرفة من النوع الاول والمعروف بتوزيع كامبل والتوزيع الثاني هو توزيع بيرسن اللوغارتيمي الثالث. هذا بالاضافة الى احد التوزيعات الهايدرولوجية الحديثة وهو توزيع ديميترس والذي تم تطويره حديثا من قبل الباحث اليوناني كوتسويانس ديمترس بالاعتماد على الاسس النظرية والقيم التجريبية للعديد من السلاسل الهيدرولوجية في العالم. بينت الدراسة ان توزيع ديمترس الحديث يعطي نتائج اكثر من الواقع في فترات العودة التي لا تزيد على 10 سنين بينما تكون النتائج متقاربة مع التوزيعات القياسية لفترات العودة التي تزيد عن 10 سنين. كما بينت التحليلات الاحصائية وبالاعتماد على نتائج حسن المطابقة لاختبارات مربع كاي ان التوزيعات القياسية والحديثة يمكن استخدامها في تمثيل السلاسل الهيدرولوجية لروافد نهر دجلة وبشكل جيد جدا مع الاخذ بنظر الاعتبار ان توزيع بيرسن اللوغارتيمي الثالث هو الافضل من بينهن.

1- Introduction

Flood studies regularly require the estimation of the peak discharge for a specified return period that is substantially longer than the available gauged record. Typically the estimation of the peak for the 100-year return period event is based on a gauged annual maximum series less than 25 years in length [1].

Acreman and Horrocks [2], show how the use of information on historic floods, based on the methodology described in the Flood Studies Report (NERC, 1975), gave greater confidence in the assessment of the rarity of the two events, than that using the relatively short gauged record alone. Archer [3], provides a further example of how historical flood information has been used to improve flood frequency estimates. Historic discharges from 1771 for the River Wear at Durham were estimated and used to extend the gauged record that began in 1958. After fitting flood frequency curves to the gauged data, and a number of combinations of gauged and historical data for different periods, it was shown that the use of the gauged flood series alone was likely to lead to a serious underestimation of the risk of flooding.

The incorporation of historical data into flood frequency estimates has been the subject of considerable debate in the literatures [4,5,6]. The use of pale flood techniques has also received considerable attention, particularly in the USA [7,8]. A review of the extensive literature available, written for the practitioner, should greatly assist those; wishing to assess the preferred flood frequency curve produced by conventional analysis in the light of historical information.

In consultation with the Upper Thames River basin southwestern Ontario, Gumbel and the Log Pearson T-3 statistical distributions were employed to fit the hydrologic output data. The results show similar outcomes for both distributions up to the 100-year return period; Log Pearson T-3 distribution shows higher magnitudes for larger recurrence interval[9].

2-Methodology

2-1 Flood Frequency Analysis

The flood frequency analysis is one of the important studies of river hydrology. It is essential to interpret the past record of flood events in order to evaluate future possibilities of such occurrences. The estimation of the frequencies of flood is essential for the quantitative assessment of the flood problem. The knowledge of magnitude and probable frequency of such recurrence is also required for proper design and location of hydraulic structures and for other allied studies. The gauge data, which are random variable, follow the law of statistical distribution.

After a detailed study of the distribution of the random variables and its parameters such as standard deviation, skewness, etc. and applying probability theory, one can reasonably predict the probability of occurrence of any major flood events in terms of discharge or water level for a specified return period [10].

In this present work, flood frequency analysis have been done by selecting annual maximum discharge of the major six streams of Tigris river in Iraq (Musul, Upper Zab, Lower Zab, Adhaim, Derbendi-khan and Hemreen), for a period of 39 years of historical monthly inflow from October 1962 to September 2001. Two standard methods of statistical distribution, Gumbel's extreme value distribution and Log Pearson T-3 distribution, as well as a new model that was suggested by Demetris are attempted here to fit the hydrologic series of the annual peak inflow of 39 years for each stream. In the following sections, the above methods will be explained breviary.

2-1-1 Plotting Position

This method represents one of the popular methods of studying the stream flow variability. It is a plot of discharge against the percentage probability of equaled or exceeded (***P***). In many cases the corresponding returned period ***T*** is also used instead of the ***P***. the value of ***P*** may be estimated using Weibull formula as follows [11]:

$$p = \frac{m}{N+1} \qquad \text{With} \quad T = 1/P \qquad \text{------------------------- (1)}$$

Where:

P = probability of flow magnitude being equaled or exceeded.

T = Return period in years.

In deed, this is an empirical formula and there are many other formulas that could be used to estimate the value of ***P***, Table (1).

Table (1): Some of Famous Plotting position formulas [12].

No.	Formula	Method
1	m / N	California
2	(m – 0.5) / N	Hazon
3	(m – 0.3) / (N + 0.4)	Mecodeaf
4	(m – 0.44) / (N + 0.12)	Blum
5	(m – 3/8) / (N + 1/4)	Krtkortn

2-1-2 Gumbel's Method

This extreme value distribution was introduced by Gumbel (1941) and is commonly known as Gumbel's distribution [11]. It is one of the most widely used probability analysis for extreme values in hydrologic and meteorological studies for prediction of flood, rainfall etc.

Gumbel defined a flood as the largest of the 365 daily flows and the annul series of flood flows constitute a series of largest values of flows. To provide the inflow at different return periods, Gumbel suggested the following form to be used:

$$Q = \bar{Q} + K \cdot S \qquad \text{------------------------- (2)}$$

Where:

Q = Value of variate with a return period 'T'

$\bar{Q}$ = Mean of the variate.

S = Standard deviation of the sample.

K = Frequency factor expressed as:

$$K = \frac{Y_T - Y_n}{S_n} \qquad \text{------------------------- (3)}$$

Y_T = Gumbel reduced variate expressed by:

$$Y_T = -Ln\left[Ln\left[\frac{T}{T-1}\right]\right] \qquad \text{------------------------- (4)}$$

Where:

T = Return period which is $1/P$
P = probability of flow magnitude being equaled or exceeded
Y_n = Reduced mean from table
S_n =Reduced standard deviation from table.

2-1-3 Log Pearson Type - III Method

This method is extensively used in USA for project sponsored by US Government. In this method, the variate is first transformed into logarithmic form (base10) and the transformed data is then analyzed. If 'X' is the variate of a random hydrologic series, then the series of *(Z)* variates, where $Z = log\ x$, are first obtained. For this 'z' series, for any recurrence interval "T', the equation is:

$$Z_Q = Z_m + K_z \cdot S_z$$ ------------------------ (5)

Where:

Z_m = Mean of the (Z_Q) values.

S_z = Standard deviation of the (Z_Q) variate sample.

K_z = Frequency factor usually taken from table with values of coefficient of skew "C_s" and recurrence interval (*return period T*).

C_s = Coefficient of skew of variate, Z_Q.

$$C_S = \frac{N \sum (Z_Q - Zm)^3}{(N-1)(N-2)\ (S_Z)^3}$$ ------------------------ (6)

N = Sample size = Number of years of record.

After finding (Z_Q) with the equation above, the corresponding value of Q is obtained by:

$$Q = 10 \wedge Z_Q$$ ------------------------- (7)

2-1-4 Demetris Model (New Model)

Demetris Koutsoyiannis suggested a new model based on theoretical arguments and several empirical analyses using long rainfall records to be an alternative one instead of Gumble distribution [13]. The new model is:

$$Q = \lambda\ [Ln(T) + \psi]$$ ------------------------- (8)

Where:

$$\lambda = 0.78 * S$$ ------------------------- (9)

$$\psi = (Q_m / \lambda) - 0.577$$ ------------------------- (10)

Where: Q_m is the mean value of Q.

All other variables are already defined above.

After obtaining the required inflow by the above methods for different return periods flood, Chi Square test should be carried out for "goodness of fit".

3- Results and Discussion

The results analysis for the annual maximum discharge of the Iraqi six streams by the three flood frequency distributions under the study are represented in Figures (1 to 6). The results indicate that in case of low return periods (less than 10 years), Demetris model allows the overestimation values comparing with Gumbel and log Pearson type 3 over all of the streams. This could be explained as follows:

The Gumbel model in forms of equations (2 to 4) could be expressed by the following form:

$$Q = \frac{S}{S_n} Y_T + \bar{Q} - \frac{S}{S_n} Y_n \qquad \text{------------------------- (11)}$$

For a very large sample size, the values of S_n and Y_n will be reduced to (1.2825 and 0.577) respectively [14]. Consequently, Eq.(11) will take the following form:

$$Q = (0.78\ S) Y_T + \bar{Q} - (0.78\ S)\ 0.577 \qquad \text{------------------------- (12)}$$

Considering the value of (λ) in Eq.(9) and comparing them with the above form, the following equation will be resulted after some algebraic steps:

$$Q = \lambda Y_T + \lambda \left(\frac{\bar{Q}}{\lambda} - 0.577 \right) \qquad \text{------------------------- (13)}$$

Now, substitute the form of ψ in Eq.(10), the above equation will be reduced to:

$$Q = \lambda\ (Y_T + \psi) \qquad \text{-------------------------- (14)}$$

Which is the form of Demetris model but with a reduced variate of Gumbel expressing by YT of Eq.(4) instead of *Ln(T)* as it is considered by Demetris. It means, that Demetris model is a special case of Gumbel model considering a large sample size and a modified reduced variate of $Y_T = Ln(T)$, considering high return periods.

The difference between the modified reduced variate (Demetris variate) and the standard variate (Gumbel variate) are shown in fig.(7). The figure explains that Demetris variate is always provides large values comparison to Gumbel variate a long low return periods while they tend to be similar for a large *T*.

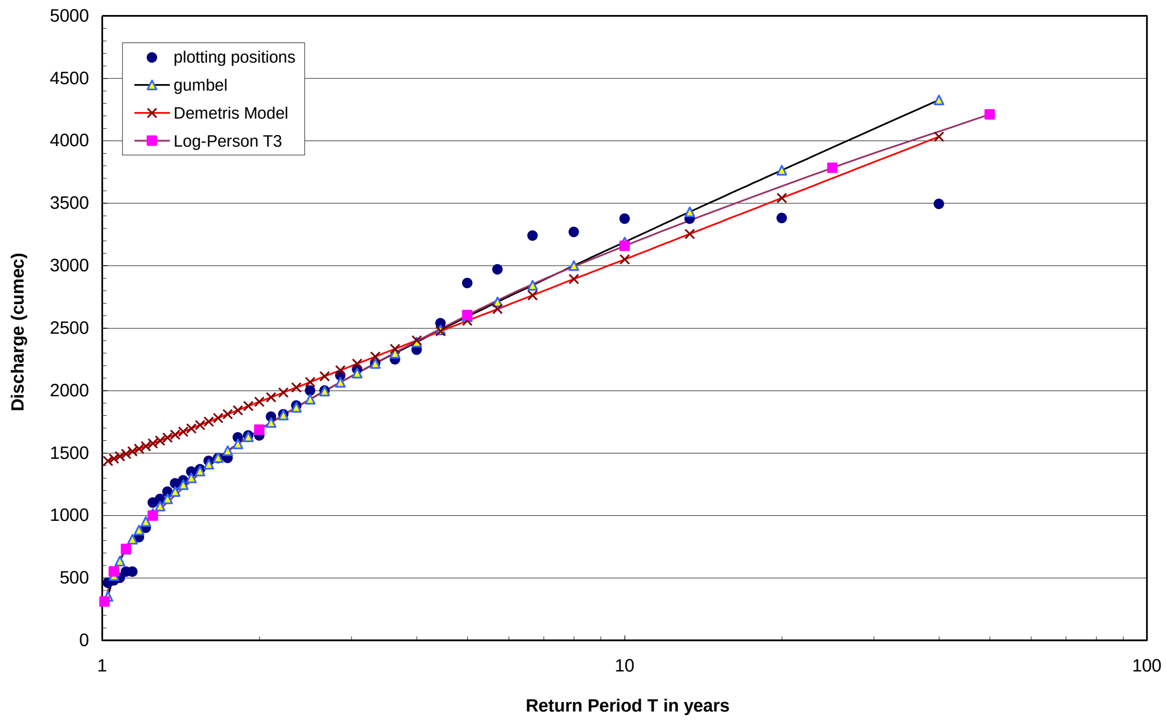


Fig.(1): Annual Maximum Yearly Inflow Series of Tigris River at Musul Stream, Fitted By Several Empirical and Theoretical Distributions.

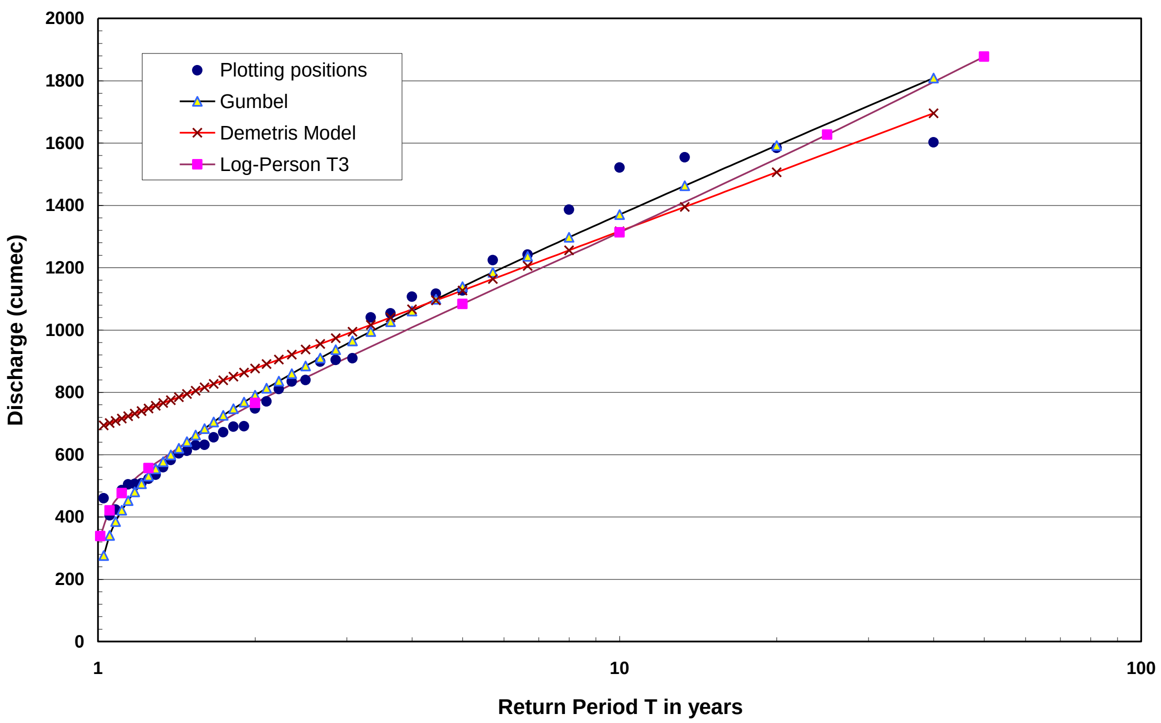


Fig.(2): Annual Maximum Yearly Inflow Series of Tigris River at Upper Zab Stream, Fitted By Several Empirical and Theoretical Distributions.

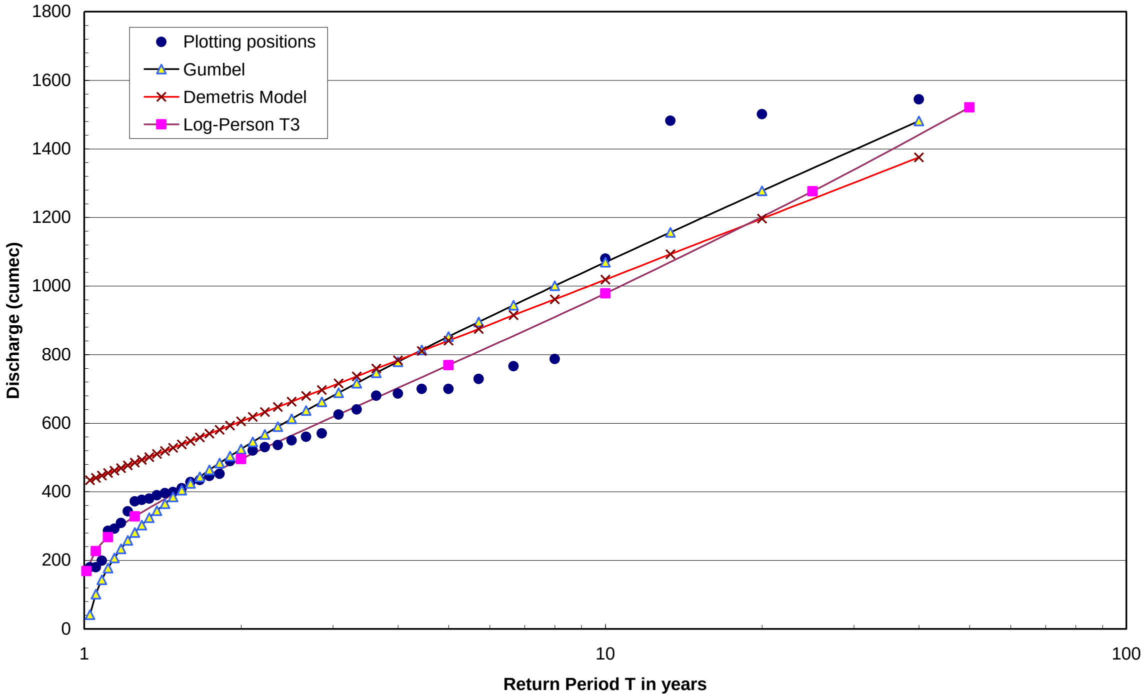


Fig.(3): Annual Maximum Yearly Inflow Series of Tigris River at Lower Zab Stream, Fitted By Several Empirical and Theoretical Distributions.

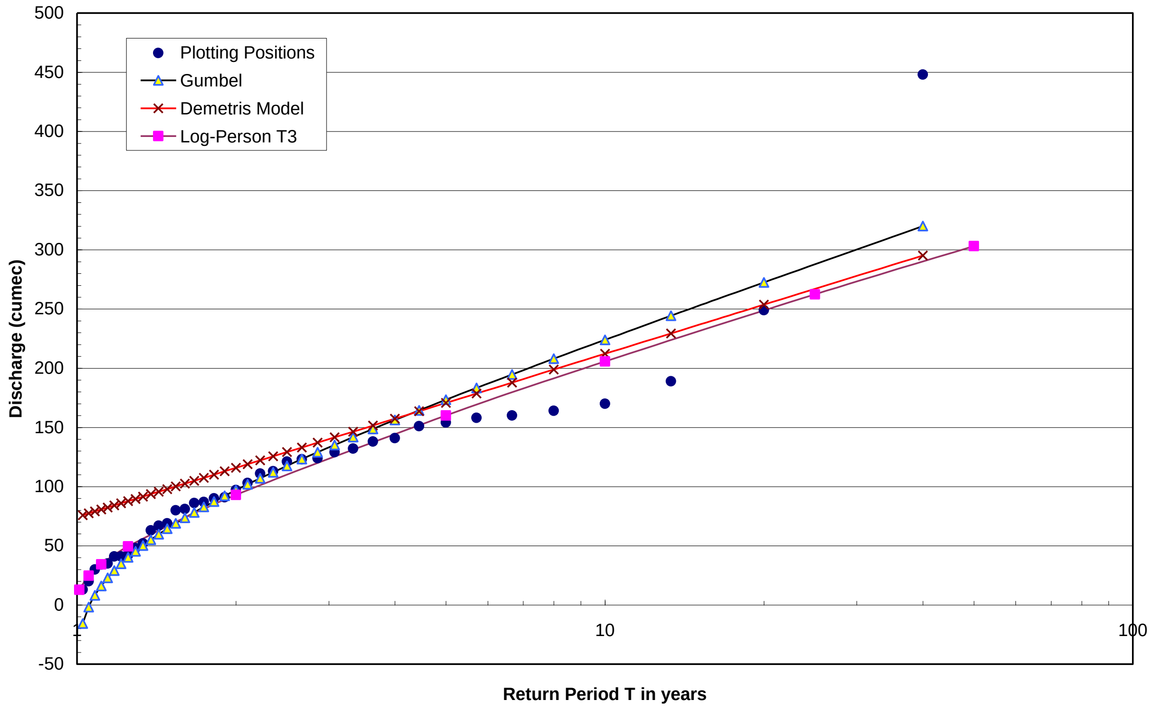


Fig.(4): Annual Maximum Yearly Inflow Series of Tigris River at Adhaim Stream, Fitted By Several Empirical and Theoretical Distributions.

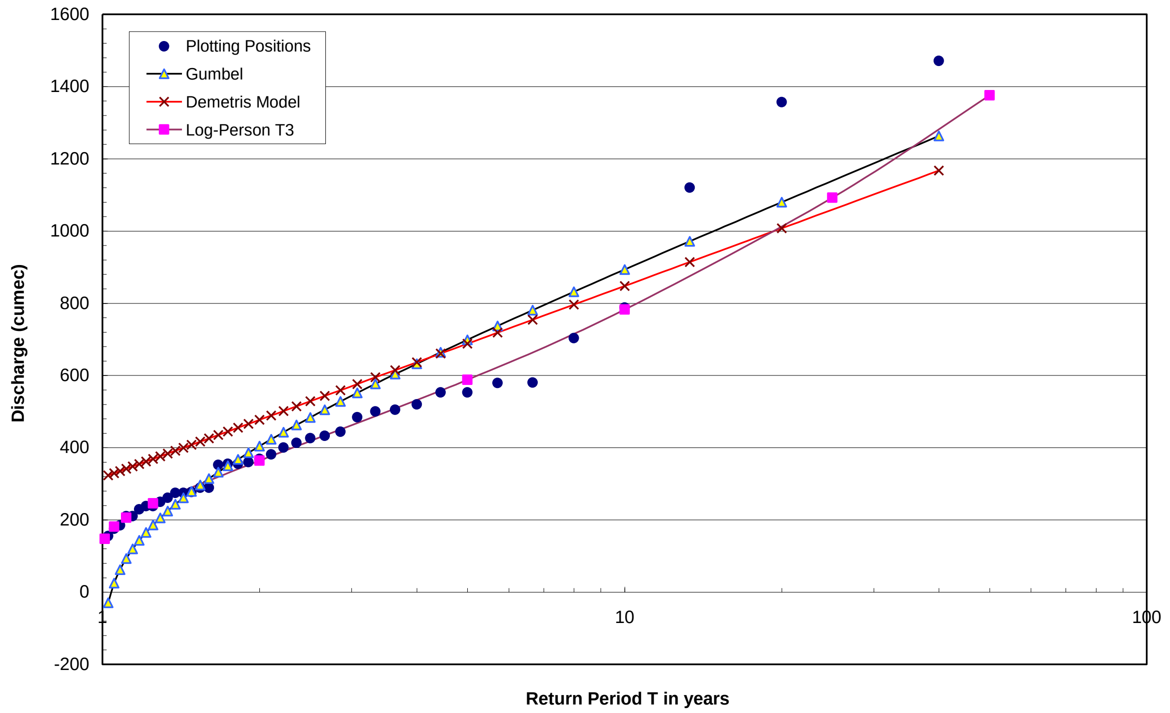


Fig.(5): Annual Maximum Yearly Inflow Series of Tigris River at Derbendi-khan Stream, Fitted By Several Empirical and Theoretical Distributions.

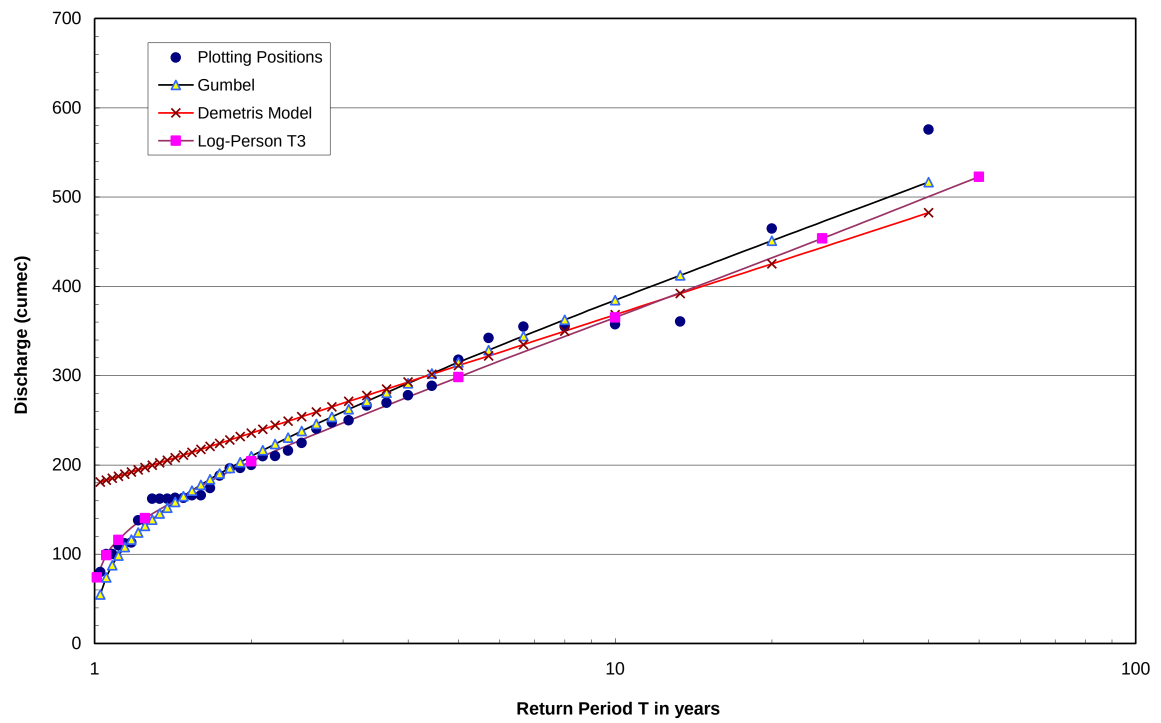


Fig.(6): Annual Maximum Yearly Inflow Series of Tigris River at Hemreen Stream, Fitted By Several Empirical and Theoretical Distributions.

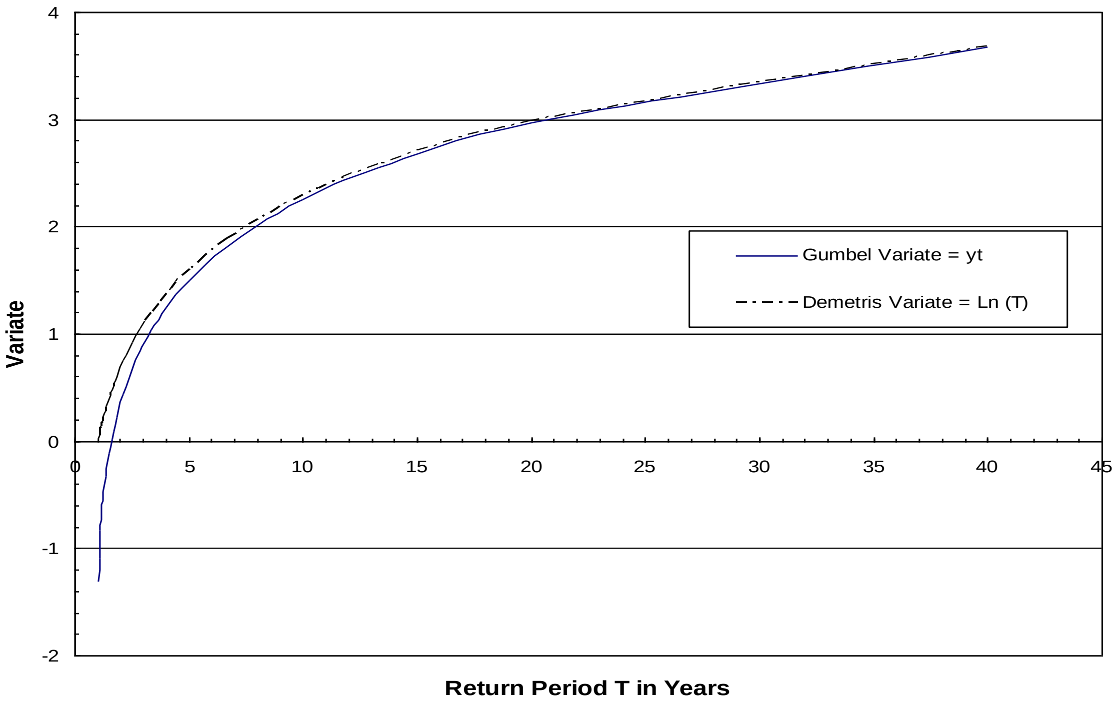


Fig.(7): Comparing Results of Gumbel and Demetris Variates.

This is because both (Demetris and Gumbel variates), tend to *Ln(T)* for large return periods *T*, and this could be demonstrated as follows:

Applying the Maclaurin series expansion to *Ln(1+ x)* yields, [15]:

$$0 + x - \frac{x^2}{2} + ... \qquad \text{-------------------------- (14)}$$

For absolute (x) << 1, this gives a first order approximation to

$$Ln(1 + x) \approx x \qquad \text{-------------------------- (15)}$$

So, for large *T* the Gumbel reduced variate will be:

$$Y_T = -Ln\left(-Ln\left(1 - \frac{1}{T}\right)\right) \rightarrow -Ln\left(\frac{1}{T}\right) = LnT \qquad \text{-------------------------- (16)}$$

For this, Demetris model tends to be overestimate with return periods of less than 10 years comparing to Gumbel distribution over all the time.

However the figures show that the Iraqi stream flow could be represented by the three models in a very good way, which is also supported by the chi square test [16]. The Chi square test is carried out to find the best fit method by comparing the computed values with the corresponding observed values for some significant levels **(α)**.

It was found that in some streams and for Gumbel distribution, the results are not accepted at 0.05 significant level contrary to other two methods, especially for Log Pearson T-3 which is found to be the best fit one. However, all these results are shown in Table (2)

Table (2): The Results of CHI Square Test Analysis for the Used Models

	Model											
Stream's Name	**Gumbel**				**Log pearson T-3**				**Demetris**			
	χ	D.F	α	$_{\alpha}\chi$	χ	D.F	α	$_{\alpha}\chi$	χ	D.F	α	$_{\alpha}\chi$
Musul	0.52	2	0.05	5.99	0.94	2	0.5	5.99	0	3	0.05	7.81
Upper Zab	0.94	2	0.05	5.99	4.61	2	0.05	5.99	0.29	3	0.05	7.81
Lower Zab	7.38	1	0.005	7.88	1.15	1	0.05	3.84	0.52	2	0.05	5.99
Adhaim	1.09	1	0.05	3.84	0.16	1	0.05	3.84	1.22	2	0.05	5.99
Derbendi-khan	4.73	1	0.01	6.63	2.31	2	0.05	5.99	0.94	2	0.05	5.99
Hemreen	4.96	2	0.05	5.99	3.53	2	0.05	5.99	0	3	0.05	7.81

Table (3) shows some of important return periods estimated by each one of the above three methods. It can be seen that the differences start to be particularly unnoticeable at high return periods especially between Gumbel and Demetris models according to the same reasons above.

Table (3): Predicted Floods Frequency Depending on the Applied Models, (M^3/s)

Stream's Name	Return Period In Years	Demetris	Gumbel	Log Pearson T-3
Musul	1000	6314	6905	5670
	200	5173	5619	4941
	100	4682	5064	4573
	50	4191	4507	4211
	25	3700	3946	3784
	10	3050	3190	3159
	5	2559	2591	2604
	2	1910	1687	1686
Upper Zab	1000	2575	2803	3158
	200	2135	2307	2423
	100	1946	2093	2142
	50	1756	1878	1877
	25	1567	1662	1626
	10	1316	1370	1313
	5	1127	1139	1083
	2	876	790	766
Lower Zab	1000	2201	2416	2860
	200	1788	1950	2077
	100	1610	1749	1787
	50	1432	1547	1521
	25	1254	1343	1276
	10	1019	1069	979
	5	841	852	769
	2	605	525	496

	1000	488	538	464
	200	392	429	380
	100	350	382	343
Adhaim	50	309	335	303
	25	267	288	262
	10	212	224	206
	5	171	173	160
	2	116	97	93
	1000	1909	2101	3340
	200	1538	1683	2110
	100	1378	1503	1712
	50	1219	1321	1376
	25	1059	1139	1092
	10	848	893	783
	5	688	698	588
	2	477	404	364
Derbendi-khan	Return Period In Years	Demetris	Gumbel	Log Pearson T-3
	1000	747	816	854
	200	615	667	668
	100	558	602	594
	50	501	537	523
	25	444	472	454
	10	368	384	365
	5	311	315	298
	2	236	210	204

4- Conclusion

The standard probability flood distributions of Gumbel and log pearson type-3 are still considered a very good models to estimate the flood frequency of the annual flow rate for Iraqi streams comparing to the recent models, in spite of that, the recent models may be more easy to parameters estimation and their application. Furthermore, log Pearson type-3 is considered to be the best one comparing with Gumbel and Demetris models.

References

1. C. Bayliss & D. W. Reed," The use of historical data in flood frequency estimation", Report of Natural Environment Research Council, 2001, www.ceh.ac.uk.

2. M. C. Acreman, & R. J. Horrocks, "Flood frequency analysis for the 1988 Truro Floods", J. Inst. Water & Environ. Manage., 4, 62-69,1990

3. D. R. Archer, "Improvement in flood estimates using historical flood information on the River Wear at Durham", Proc. BHS First National Hydrology Symp., Hull, 14-16 September, 1987.

4. R. M. Hirsch, "Probability plotting position formulas for flood records with historical information", J. Hydrol., 96, 185-199, 1987.

5. J. V. Sutcliffe, "The use of historical records in flood frequency analysis" , J. Hydrol., 96,159-171, 1987.

6. J. R. M. Hosking, & J. R. Wallis, "The value of historical data in flood frequency analysis", Wat. Resour. Res., 22, 1606-1612, 1986.

7. V. R. Baker, " Paleo flood hydrology and extraordinary flood events", J. Hydrol., 96, 79-99, 1987.

8. J. R. Stedinger, R. M. Vogel, & E. Foufoula-Georgiou, " Handbook of Hydrology - Frequency analysis of extreme events", (ed: Maidment, D.R.). McGra Hill, 18.1-18.66, 1992.

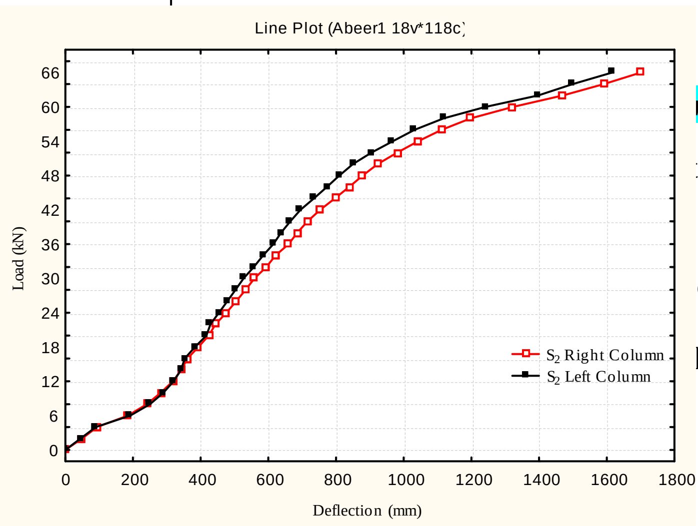
9. P. Prodanovic and P. Simonovic," Impacts of Changing Climatic Conditions in the Upper Thames River Basin", Canadian Water Resources Journal, Vol. 32(4): 265–284, 2007.

10. D. W. Reed, D. Jakob & A. J. Robson, " Flood Estimation Handbook ",Vol. 3, Institute of Hydrology, Wallingford, 28-39, 1999.

11. V. T. Chow, "Handbook of Applied Hydrology", McGraw-Hill, New york, N. Y., 1964.

12. K Subramanya, "Engineering Hydrology", Tata McGraw-Hill, New Delhi, 2008.

13. Demetris Koutsoyiannis, ”A critical review of probability of extreme rainfall: principle and models”, Advances in Urban Flood Management, edited by ,S. Grawin, E.pasche, A.Vassilopouls ,and C.Zevenbergen, 139-166, Taylor and Francis ,London ,2007.



ter Resources Engineering", 7th , Khanna Publishers,

ering Mathematics", 9th, McGraw-Hill, 2005.

ls of Probability and Statistics for Engineers", T.T.