

Implementation of Gauss function in determining probability of floods at the gauge station “Dolenci” on the Crna Reka in Republic of Macedonia

Vasileski Dragan^A, Radevski Ivan^A

Received: March 2011 | First Revised: April 2011 | Second Revised: October 2011 | Accepted: November 2011

Abstract

The Gaussian distribution function is symmetric and in hydrological research can be used for the purposes of comparison. After the determination of the parameters and the return probability levels associated with return periods of 2, 5, 10, 25, 50, 100, 1000 and 10.000 years, they are compared with the empirical probabilities, which were previously determined with empirical formulas. To obtain these results a 40 years period was considered with absolute maximum and absolute minimum flows of water recorded at the “Dolenci” water level check point on the Crna Reka. The Gauss distribution function was applied and the results recorded on a probability list were compared with the empirical distribution function.

Key words: maximum discharge [flow], Gaussian distribution function, parameters, probability, standard variable, standard deviation.

Introduction

Crna Reka is located in the western part of Republic of Macedonia. The gauge station Dolenci is located in the upstream section of Crna Reka, in the area known as Demir Hisar with altitude of 739 meters above the sea level with alimentation area of 216,5 km².

The distribution of monthly and annual amounts of precipitation shows that the highest rainfall is observed in the second decade from 1971 to 1980, while the average annual amount of precipitation is 755 mm (period 1961-1990). The largest amount of rainfall was recorded in 1963 (978 mm). The lowest annual amount of rainfall was recorded in 1990 with quantity of 522 mm. Average annual rainfall is 732 mm. Allocation of multi-average amounts of precipitation for the rain gauge “Dolenci” shows that rainfall is higher in autumn than in spring ie the months of November and December are more rainy than April and May. The lowest rainfall amount is observed in August with an average annual quantity of 36 mm, while the month with the highest rainfall is November with an average of 90 mm.

The water discharge at the gauge station is measured by a staff gauge. The maximum water discharge data array for the 40 year period is continuous.

Standard period for flood frequency analysis is 30 years (Srebrenovic, 1986).

Methods and data

As for all other distribution functions it is important to determine the parameters of the Gauss distribution function, which will help the relevant probabilities for a certain period to be obtained. For hydrological research it is particularly important to take a standard period of 30 years in order for the forecast to be accurate. It is also important to note that the Gauss distribution function is symmetric, and this feature does not match with the theoretical distribution of floods in natural conditions.

The empirical distribution is calculated according to Weibull’s formula, where m is the ordinal number in the sequence of the data and N is the total number of members in the thread. $P_m = m/(N+1)$. This formula facilitates the determination of F_e (empiri-

^A Department of geography, Faculty of Natural sciences and mathematics, University “St. Cyril and Methodius” Skopje, Gazi Baba b.b., 1000 Skopje, Republic of Macedonia

cal distribution) of the hydrological series (sequence), which is necessary to calculate the values of the maximum difference D_n between the empirical distribution function estimated using the Weibull formula and the theoretical Gauss distribution function according to the Kolmogorov-Smirnov test. The Gauss distribution function has a dual parameter, and its main parameters are the arithmetic mean (annual average flow) and standard deviation (σ). The mean maximum range of hydrological flows for the period 1960/61 – 1999/00 is $15,38 \text{ m}^3/\text{s}$.

$$Q_{sr} = 15,38^*$$

$$\sigma = 11,25^*$$

*The basic data on maximum annual flow at the gauge station "Dolenci" on the Crna Reka was provided by the Hydro-meteorological Service of the Republic of Macedonia.

The return level of flooding discharge for the appropriate return periods (T) using the Gauss distribution function is equal to the sum of the annual average flow (x_{sr}) for a period of 40 years and the product of the standard variable (z) and the standard deviation (σ) of the sequence.

$$Q_{max} = \bar{x} + z \times \sigma$$

By using the Gauss distribution function the following results were obtained for maximum theoretical water flows of the Crna Reka at the level check point "Dolenci".

Kolmogorov-Smirnov test

To determine the coincidence of the empirical distribution with the theoretical Gauss distribution it is necessary to perform a test. In hydrology, the

most popular tests are the Kolmogorov-Smirnov tests and the χ^2 tests. In this paper the Kolmogorov-Smirnov tests are primarily used because of the difference between the empirical and the theoretical distribution (D_n), which is obtained for each member in the sequence. Using the Kolmogorov-Smirnov test as a non parameter is particularly advantageous in hydrological research, particularly because the information is stored for all recorded data and each result of the theoretical function is compared to the respective empirical value. The resulting difference in probabilities between the empirical and the theoretical distribution is denoted by D_n , which should be less than the critical value $-D_0$ (0,215) in order to accept the zero hypothesis that the theoretical function coincides with the empirical. The biggest disadvantage of such testing is the determination of the difference between the empirical and the theoretical distribution function, which is only possible for the maximum water flow recorded in that period ($47 \text{ m}^3/\text{s}$).

$$D_n = \max|F_o(x_i) - F_t(x_i)|$$

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, -\infty < x \leq +\infty$$

After determining the Gauss distribution function (F_t) the largest difference between the empirical and the Gauss distribution function is calculated. The largest difference D_n is $-0,16$. This value is less than the critical value D_0 and therefore the Gauss distribution function provides a good agreement to the empirical distribution.

Table 1. Determining the probability of flooding of the Crna Reka for the period 1960/61 – 1999/00 at the level check point "Dolenci"

Return period [T]	Probability [P%]	z	z · σ	Qmax [m ³ /s]
10000	0.01	3.715	41.7938	57.17
1000	0.1	3.09	34.7625	50.14
200	0.5	2.576	28.9800	44.36
100	1	2.326	26.1675	41.55
50	2	2.054	23.1075	38.49
25	4	1.752	19.7100	35.09
20	5	1.64	18.4500	33.83
10	10	1.28	14.4000	29.78
5	20	0.842	9.4725	24.85
2	50	0	0.0000	15.38

Table 2. Calculating the Kolmogorov-Smirnov parameter D_n for the Gauss distribution

m	Q [m ³ /s]	Fe%	Fe	z	Ft	1-Ft	Dn
1	2.20	98.27	0.98	-1.17156	0.120688	0.88	0.10
2	3.30	95.79	0.96	-1.07378	0.141461	0.86	0.10
3	3.50	93.32	0.93	-1.056	0.145484	0.85	0.08
4	3.84	90.84	0.91	-1.02578	0.152498	0.85	0.06
5	4.82	88.37	0.88	-0.93867	0.173951	0.83	0.06
6	4.99	85.89	0.86	-0.92356	0.177859	0.82	0.04
7	5.46	83.42	0.83	-0.88178	0.188948	0.81	0.02
8	6.30	78.47	0.78	-0.80711	0.209801	0.79	-0.01
9	6.30	80.94	0.81	-0.80711	0.209801	0.79	0.02
10	6.90	73.51	0.74	-0.75378	0.225491	0.77	-0.04
11	6.90	75.99	0.76	-0.75378	0.225491	0.77	-0.01
12	7.88	71.04	0.71	-0.66667	0.252493	0.75	-0.04
13	7.92	68.56	0.69	-0.66311	0.25363	0.75	-0.06
14	8.06	66.09	0.66	-0.65067	0.257631	0.74	-0.08
15	9.60	63.61	0.64	-0.51378	0.303704	0.70	-0.06
16	9.64	61.14	0.61	-0.51022	0.304948	0.70	-0.08
17	9.96	58.66	0.59	-0.48178	0.314982	0.69	-0.10
18	10.20	56.19	0.56	-0.46044	0.322599	0.68	-0.12
19	10.40	53.71	0.54	-0.44267	0.329003	0.67	-0.13
20	10.50	51.24	0.51	-0.43378	0.332225	0.67	-0.16
21	12.00	48.76	0.49	-0.30044	0.381919	0.62	-0.13
22	12.60	46.29	0.46	-0.24711	0.402411	0.60	-0.13
23	14.10	43.81	0.44	-0.11378	0.454707	0.55	-0.11
24	15.00	41.34	0.41	-0.03378	0.486527	0.51	-0.10
25	15.80	38.86	0.39	0.037333	0.51489	0.49	-0.10
26	16.90	36.39	0.36	0.135111	0.553738	0.45	-0.08
27	18.70	33.91	0.34	0.295111	0.616045	0.38	-0.04
28	18.80	31.44	0.31	0.304	0.619436	0.38	-0.07
29	20.00	28.96	0.29	0.410667	0.659342	0.34	-0.05
30	20.40	26.49	0.26	0.446222	0.672282	0.33	-0.06
31	20.80	24.10	0.24	0.481778	0.685018	0.31	-0.07
32	21.80	21.53	0.22	0.570667	0.715887	0.28	-0.07
33	23.30	19.06	0.19	0.704	0.759284	0.24	-0.05
34	23.80	16.58	0.17	0.748444	0.772904	0.23	-0.06
35	25.90	14.11	0.14	0.935111	0.825134	0.17	-0.03
36	35.00	11.63	0.12	1.744	0.95942	0.04	0.08
37	35.80	9.16	0.09	1.815111	0.965247	0.03	0.06
38	38.00	6.68	0.07	2.010667	0.97782	0.02	0.04
39	41.00	4.21	0.04	2.277333	0.988617	0.01	0.03
40	47.00	1.73	0.02	2.810667	0.997528	0.00	0.01

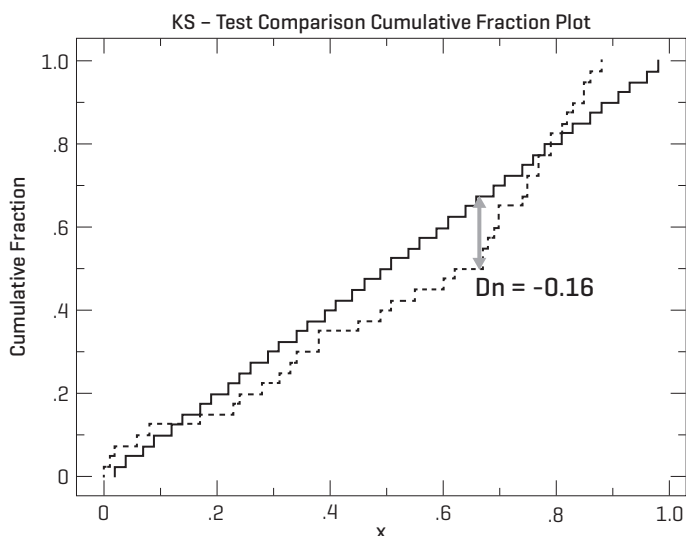


Figure 1. Position of the maximum difference D_n between the empirical and the Gauss distribution function

It is evident from the above graph that the Gauss distribution function presents smaller values than the empirical distribution function. The largest difference $D_n = -0.16$ is expressed on the empirical probability of 0.51.

The above table presents the results for maximum theoretic water discharge for different return periods (2-10.000 years), calculated with five

probability distributions. Results of the Gauss distribution function have the lowest values compared with other distributions.

From the above graph the difference between the Gaussian normal distribution and the asymmetric distributions, like Gumbel, LogNormal, Pearson III and Log Pearson III distributions, is evident. There is a better agreement between the asymmetric distributions, especially logarithmic distributions and empirical distribution functions.

From the above table it is evident that the maximum difference D_n for the Gauss distribution function has the highest D_n value compared with the other four distributions functions. The asymmetrical distributions used here provide a better match with the empirical distribution function. The distribution providing the best fit is the LogPearson type III.

An empirical distribution function of maximum annual flows for the period 1960/61 – 1999/00 is presented with blue dots on the above chart. The Gauss distribution is marked with a green line and has a symmetrical shape. Notable is the fact that the five largest recorded flows for the analyzed period of 40 years not only deviate from the Gauss distribution function, but also do not belong within the 95% confidence intervals obtained by using "Monte Carlo" simulation.

Table 3. Comparing the results of Gauss distribution function and four other asymmetric distributions

Distribution	Gauss	Pearson III	LogPearson III	Gumbel	Lognormal
T / Parametres	x_{sr}, σ	Q_{sr}, C_v, C_s	C_v, C_s, y_{sr}	α, β	y_{sr}, σ
10000	57,17	86,11	168,66	91,14	196,11
1000	50,14	68,60	109,57	70,96	122,32
200	44,36	55,91	77,09	56,81	82,90
100	41,55	50,41	64,86	50,69	68,69
50	38,49	44,57	53,59	44,63	55,94
25	35,09	38,62	43,26	38,40	44,53
20	33,83	36,71	40,12	36,38	40,92
10	29,78	30,42	30,90	30,24	31,17
5	24,85	23,69	22,44	23,48	22,40
2	15,38	16,05	12,01	13,54	11,86

Table 4. Kolmogorov-Smirnov test for five probability distributions with maximum difference D_n

K-S test	$\alpha=1\%$	$\alpha=5\%$	$\alpha=10\%$	Attained α	D_n
Normal	ACCEPT	ACCEPT	ACCEPT	26,47%	0,15574
LogNormal	ACCEPT	ACCEPT	ACCEPT	88,60%	0,08901
Pearson III	ACCEPT	ACCEPT	ACCEPT	83,42%	0,09515
Log Pearson III	ACCEPT	ACCEPT	ACCEPT	98,59%	0,06871
Gumbel	ACCEPT	ACCEPT	ACCEPT	65,96%	0,11238

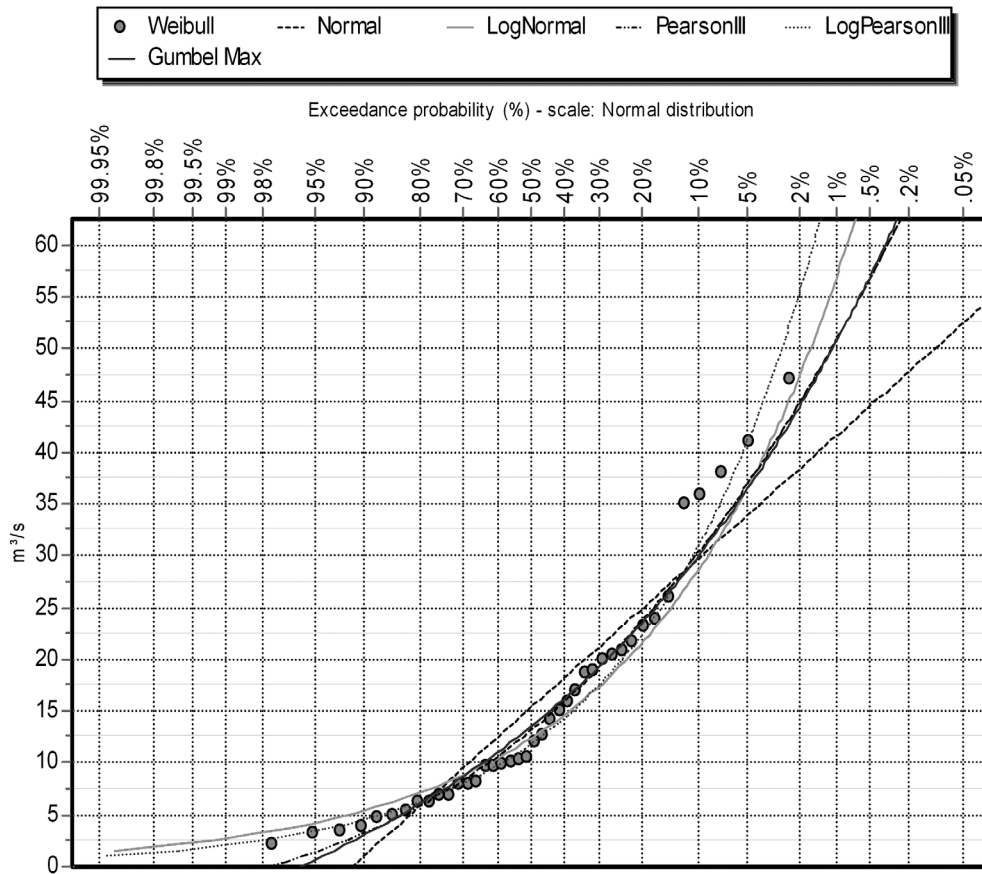


Figure 2. Probability chart with five distributions: Normal [Gauss], Gumbel, LogNormal, Pearson III and LogPearson III

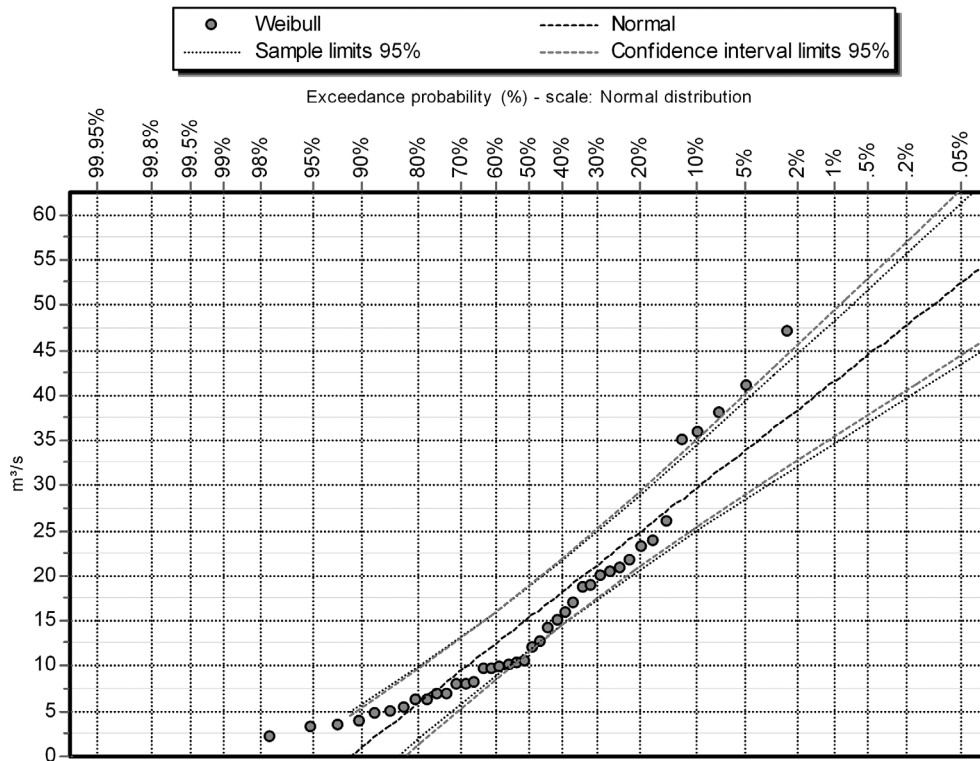


Figure 3. Probability chart of Gauss function with 95% confidence intervals using "Monte Carlo" simulation

Discussion and Conclusions

The Gauss distribution function in flood analysis could be used for reference purposes, but the results, especially for the smaller probabilities that correspond to return periods of 100, 1000, or 10.000 years, cannot be considered as relevant, i.e. high floods are underestimated.

This is confirmed by the probability paper, where the five highest values of the empirical distribution function do not match with the theoretical Gaussian distribution function, i.e. the Gauss straight line for small probabilities shows lower calculated floods. The general importance of this paper was significance of using a Gaussian distribution in determine a probability of flood features in nature conditions (upstream section of Crna Reka). The results showed the fact that Gaussian function can be used like comparing distribution in parallel with other distributions. The regime of flood features on Crna Reka is approximately equal to other neighboring streams. Beside this large deviation, the theoretical floods do not coincide with the 95% confidence intervals calculated with "Monte Carlo" simulation (simulation used for testing confidence between empirical and theoretical distribution, using confidence limits for 95% comparing).

References

- Apollov, B.A. 1963. A study of rivers. Moscow University, Moscow, 422 pp.
- Chow, V.T. 1964. Handbook of Applied Hydrology. McGraw-Hill Book Company, 367 pp.
- Gumbel, E.J. 1958. Statistics of Extremes. Columbia University Press, New York, 375 pp.
- Hrelja, H., Isailović, D. 1984. Hydrology – Collection of calculated tasks. Faculty of civil engineering, Sarajevo, 142 pp.
- Hrelja, H. 2000. Probability and Statistics in Hydrology. Faculty of Civil engineering, Sarajevo, 248 pp.
- Jevđević, V. 1956. Hydrology. Institute for the Development of Water Resources "Jaroslav Cerni", Belgrade, 401 pp.
- Jovanović, S. 1987. Application of mathematical statistics in hydrology. Faculty of Civil Engineering, Belgrade 1-8-65 pp.
- Popović, B., Blagojević, B. 1997. Mathematical Statistics applied in Hydrology. University of Nish, Nish, 260 pp.
- Srebrenović, D. 1986. Applied Hydrology. Scientific Book, Zagreb, 509 pp.
- Viessman, W., Lewis G. 1995. Introduction to Hydrology. Addison-Wesley Educational Publishers, New York. 612 pp.
- Vukmirović, V. 1990. Analysis of the probabilities of Hydrologic events. Science book, Belgrade. 174 pp.