Statistical Analysis of Probable Future Flood in Khyber Pakhtunkhwa (KPK)–Pakistan

Ateeq-ur-Rauf[1,3], Izhar Ahmad[1], Kashif Ali[1], S. Bakht Jamal[1], Shah Room[2], Sheeraz Ahmed[3]

[1]University of Engineering and Technology, Peshawar, Pakistan
[2]City University of Sciences & Information Technology, Peshawar, Pakistan
[3]Career Dynamics Research Center, Peshawar, Pakistan

Abstract- **Floods are the nation's greatest natural disaster. The severe floods, resulting from heavy monsoon rains and freak weather systems commenced in July, 2010, in the high altitude, northern parts of the country. According to the U.S. Geological Survey, Floods cause an average of $6 billion of property damage, claim 140 lives, and prompt more Presidential disaster declarations per year than any other hazard. Therefore, planning, design and construction of engineering infrastructure projects often requires consideration of the potential for flood risks. This involves the need of estimates of extreme flood probabilities to assess the infrastructure. This research is focused mainly on the flood frequency analysis, involving the method of probability distribution, working out the exceedance probability from the cumulative frequency, the inverse of which in turn gives the Return Period of the specific magnitude of flood. The project is carried out for different points located on three major hazard prone rivers of KPK. Therefore collecting discharge data on these points, applying frequency analysis methods and thus finding out the probability of return period of the peak discharge (flood) in the next 100, 50, 10 and 5 years.**

KEYWORDS: Discharge, Flood, Return Period, Probability Analysis

I. INTRODUCTION

Flood events are a part of nature. They have existed and will continue to exist. Pakistan has seen many floods, the most worst and destructive is the recent 2010 Pakistan floods [2]. A flood occurs when the overflowing water submerges land and causes deluge. It is a cruel and violent expression of water. Floods have been occurring throughout Earth history, and are expected so long as the water cycle continues to run [8]. Streams receive most of their water input from precipitation, and the amount of precipitation falling in any given drainage basin varies from day to day, year to year, and century to century [6]. In the northern province of Khyber Pakhtunkhwa, formerly known as the North West Frontier Province, large-scale flash floods destroyed thousands of homes and infrastructural elements in a matter of hours. The water receded quickly and flowed down from the mountainous province to the central and southern provinces of Punjab and Sindh [2].

For the design of Hydraulic structures it is not sufficient to take into consideration only the magnitude of the peak discharge, but also the frequency of this peak discharge is necessary. This study aims to evaluate flood occurrence probability in a most probable year using return periods and workout the return periods of maximum peak discharges using percentage of risks.

II. STUDY AREA

Study Area comprises of three main rivers of KPK, which are hazard prone rivers during 2010 flood that are River Indus, River Swat and River Kabul. Figure 1 shows a view of selected study area.

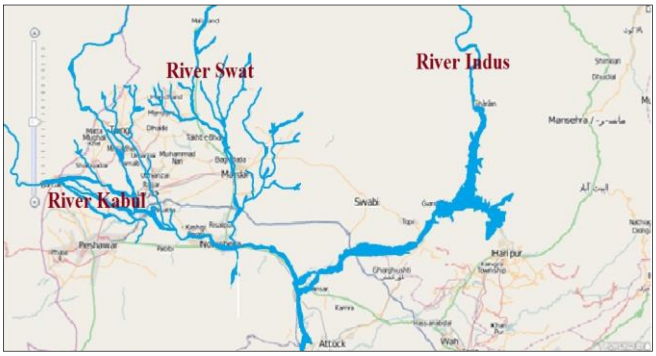

Fig.1: Map of the Selected Rivers of KPK

III. METHODOLOGY

Flood frequency is the concept of the probable frequency of occurrence of a given flood. For the design of engineering works, for example, it is not sufficient that the maximum observed flood was, say, 900m/s. It is also necessary to say what the frequency of occurrence of this flood is [1].

There are a number of probability distributions f (x), which has been suggested by many statisticians [7]. Of these, the more common are:

- Log – normal Distribution
- Pearson Type III Distribution

- Gumbel Distribution
- Frequency Distribution
- Normal Distribution

The first three methods involves parameters more than the last two methods [3]. Therefore the last two methods are easy to use for this research.

IV. ANALYSIS

For analysis two methods were performed on the annual discharge data for each site. But the analysis and results presented in this paper are for the end station (Attock Station) of River Swat and River Kabul. These methods are:

- Frequency Distribution.
- Probability Distribution.

A) FREQUENCY DISTRIBUTION:

The analysis of the data collected at end station (Attock Station) of the River Swat and River Kabul using frequency distribution method are as follows:

Table 1: Frequency Distribution

Year	Peak Discharge (cusec)	Cumulative Frequency Distribution	Cumulative Relative Frequency	Exceedance Probability	Return Period
1970	278390	278390	0.024	0.976	1.0245902
1971	300000	300000	0.048	0.952	1.0504202
1972	306000	306000	0.072	0.928	1.0775862
1973	316800	316800	0.096	0.904	1.1061947
1974	331400	331400	0.12	0.88	1.1363636
1975	356800	356800	0.144	0.856	1.1682243
1976	373800	373800	0.168	0.832	1.2019231
1977	380000	380000	0.192	0.808	1.2376238
1978	381000	381000	0.216	0.784	1.2755102
1979	383320	383320	0.24	0.76	1.3157895
1980	386460	386460	0.264	0.736	1.3586957
1981	392000	392000	0.288	0.712	1.4044944
1982	398000	398000	0.312	0.688	1.4534884
1983	404000	404000	0.336	0.664	1.5060241
1984	408400	408400	0.36	0.64	1.5625
1985	409500	409500	0.384	0.616	1.6233766
1986	421000	421000	0.408	0.592	1.6891892
1987	428100	428100	0.432	0.568	1.7605634
1988	431400	431400	0.456	0.544	1.8382353
1989	441700	441700	0.48	0.52	1.9230769
1990	447600	447600	0.504	0.496	2.016129
1991	462333	462333	0.528	0.472	2.1186441
1992	469200	469200	0.552	0.448	2.2321429
1993	474000	474000	0.576	0.424	2.3584906
1994	475200	475200	0.6	0.4	2.5
1995	486000	486000	0.624	0.376	2.6595745
1996	503000	503000	0.673	0.327	3.058104
1997	503000	504200	0.697	0.303	3.30033
1998	504200	532000	0.721	0.279	3.5842294
1999	532000	546200	0.745	0.255	3.9215686
2000	546200	548400	0.769	0.231	4.3290043
2001	548400	551500	0.793	0.207	4.8309179
2002	551500	552060	0.817	0.183	5.4644809
2003	552060	561000	0.841	0.159	6.2893082
2004	561000	562430	0.865	0.135	7.4074074
2005	562430	567000	0.889	0.111	9.009009
2006	567000	572000	0.913	0.087	11.494253
2007	572000	575000	0.937	0.063	15.873016
2008	575000	578200	0.961	0.039	25.641026
2009	578200	994600	0.985	0.015	66.666667

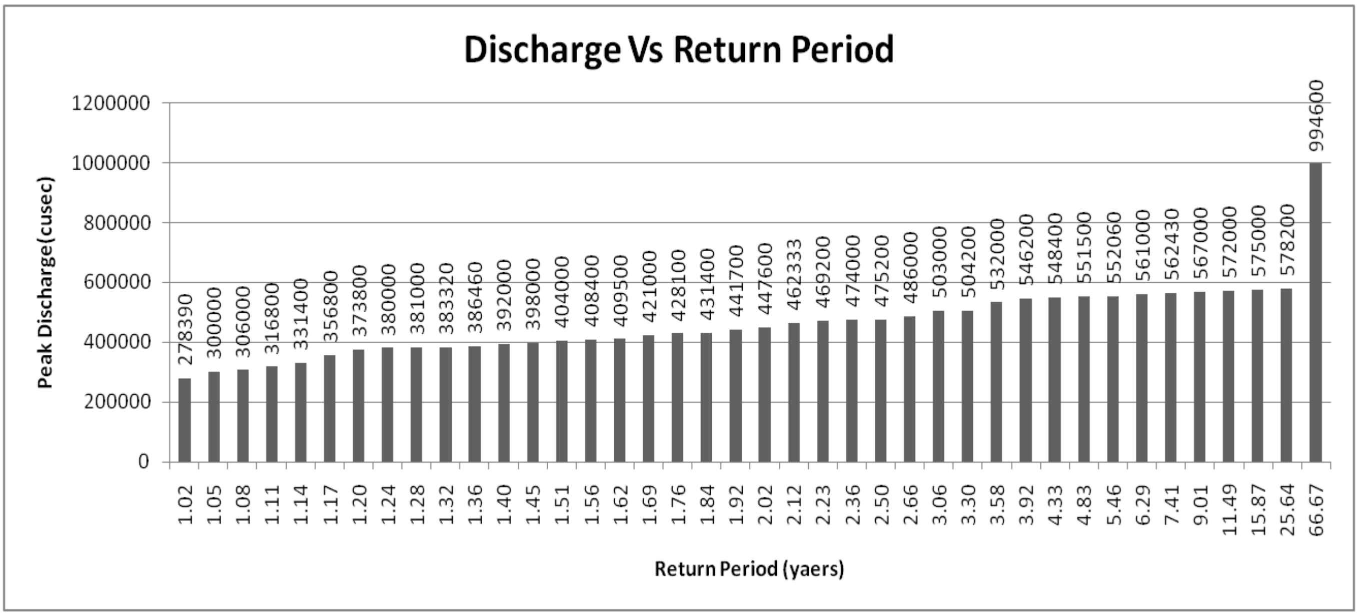

Fig.2: Showing Discharge Vs Return Period of Attock Station

B) PROBABILITY DISTRIBUTION

Probability distribution analysis are shown in Table2; while its graphical distribution is shown in Fig.3.

Table 2: Probability Distribution

Year	Peak Discharge	Z	F(x)	Exceedance probability	Return period
2004	278390	-1.530669133	0.063	0.937	1.067235859
2001	300000	-1.351728953	0.0885	0.9115	1.097092704
2007	306000	-1.302046349	0.0968	0.9032	1.107174491
2008	316800	-1.212617661	0.1131	0.8869	1.127522832
2000	331400	-1.091723324	0.1379	0.8621	1.159958242
1987	356800	-0.8814003	0.1894	0.8106	1.233654083
1981	373800	-0.740632921	0.2297	0.7703	1.298195508
2002	380000	-0.68929423	0.2483	0.7517	1.330317946
1970	381000	-0.681013796	0.2483	0.7517	1.330317946
1971	383320	-0.661803189	0.2546	0.7454	1.341561578
2003	386460	-0.635802626	0.2643	0.7357	1.359249694
1980	392000	-0.589929021	0.2776	0.7224	1.38427464
1985	398000	-0.540246417	0.2946	0.7054	1.417635384
1979	404000	-0.490563813	0.3121	0.6879	1.453699666
1972	408400	-0.454129903	0.3264	0.6736	1.48456057
1974	409500	-0.445021426	0.33	0.67	1.492537313
1999	421000	-0.349796434	0.3632	0.6368	1.570351759
1993	428100	-0.291005352	0.3859	0.6141	1.628399284
1984	431400	-0.26367992	0.3974	0.6026	1.659475606
2009	441700	-0.178391449	0.4325	0.5675	1.762114537
1983	447600	-0.129536889	0.4483	0.5517	1.8125793
1975	462333	-0.007541254	0.5	0.5	2
1996	469200	0.049320487	0.516	0.484	2.066115702
1982	474000	0.08906657	0.5319	0.4681	2.136295663
1997	475200	0.099003091	0.5359	0.4641	2.154708037
1986	486000	0.188431779	0.5714	0.4286	2.333177788
1991	503000	0.329199157	0.6255	0.3745	2.670226969
1998	503000	0.329199157	0.6255	0.3745	2.670226969
1990	504200	0.339135678	0.6293	0.3707	2.697599137
1989	532000	0.569331745	0.7123	0.2877	3.475842892
1973	546200	0.686913908	0.7517	0.2483	4.027386226
1994	548400	0.705130863	0.758	0.242	4.132231405
2006	551500	0.730800209	0.7673	0.2327	4.297378599
2005	552060	0.735437252	0.7673	0.2327	4.297378599
1976	561000	0.809464332	0.7881	0.2119	4.719207173
1977	562430	0.821305353	0.7939	0.2061	4.852013586
1978	567000	0.859146936	0.8051	0.1949	5.130836326
1992	572000	0.900549106	0.8159	0.1841	5.431830527
1988	575000	0.925390409	0.8212	0.1788	5.592841163
1995	578200	0.951887797	0.8289	0.1711	5.844535359
2010	994600	4.399860533	0.998	0.002	500

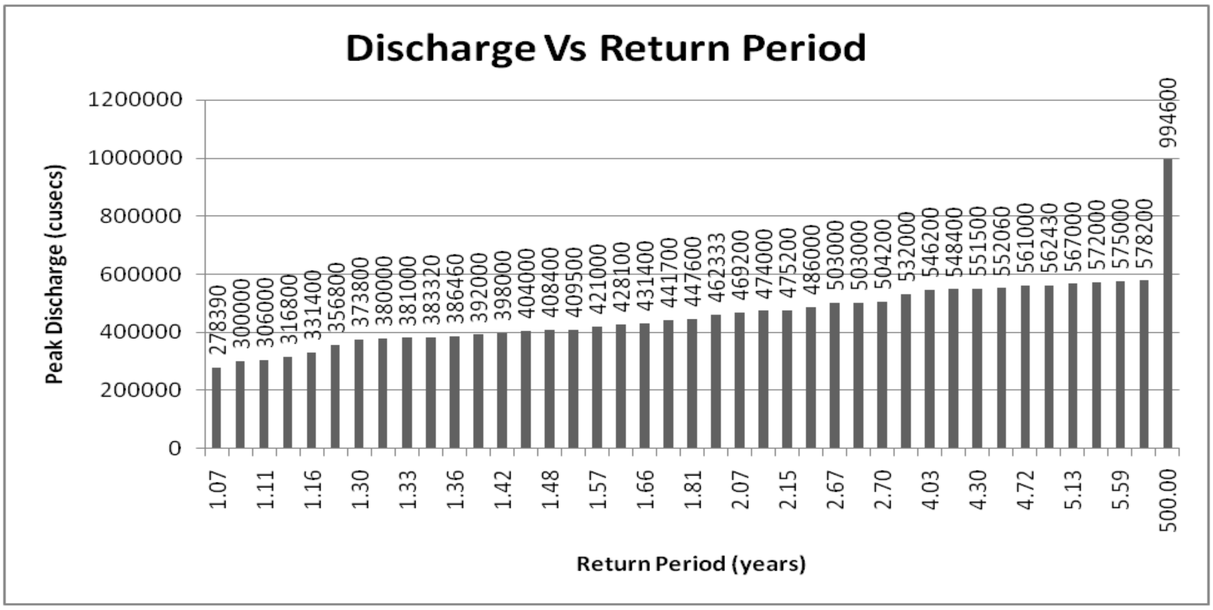

Fig.3: Showing Discharge Vs Return Period of Attock Station

V. RESULTS

Two approaches are used:

- Results based on the probability of maximum annual flood for each station in coming 5, 10, 20, 50, and 100 years.
- Results on the Risk base for each station using both methods of analysis [4].

A) Probability in Next Years:

The probabilities in Next years for the analysis through Frequency Distribution and Normal Distribution for Attock Station are given in Table 3 and Table 4 respectively.

Table 3: Results at Attock Station by Frequency Distribution

Year	Peak Discharge (Cusecs)	Return Period (Years)	P In Next 5 Years	P In Next 10 Years	P In Next 20 Years	P In Next 50 Years	P In Next 100 Years
1978	567000	9	0.445071043	0.692053852	0.90516917	0.997230675	0.999992331
1,992	572000	12	0.352772151	0.581096112	0.824519533	0.987100527	0.999833604
1998	575000	16	0.275803566	0.475539525	0.72494121	0.960320717	0.998425554
1995	578200	26	0.178072893	0.324435831	0.543613054	0.859287385	0.98019996
2010	994600	67	0.072432199	0.139617975	0.259742772	0.528527201	0.777713399

Table 4: Results at Attock Station by Normal Distribution

Year	Peak Discharge (Cusecs)	Return Period (Years)	P In Next 5 Years	P In Next 10 Years	P In Next 20 Years	P In Next 50 Years	P In Next 100 Years
1978	567000	6	0.598122428	0.838494417	0.973915947	0.999890115	0.999999988
1,992	572000	6	0.598122428	0.838494417	0.973915947	0.999890115	0.999999988
1998	575000	6	0.598122428	0.838494417	0.973915947	0.999890115	0.999999988
1995	578200	6	0.598122428	0.838494417	0.973915947	0.999890115	0.999999988
2010	994600	500	0.00996008	0.019820957	0.039249043	0.095253182	0.181433195

B) Risk Based Calculation

Results are obtained on the Basis of assuming Percentage of Risk and analysis of data through Frequency Distribution and Normal Distribution for Attock Station are given in Table 5 and Table 6 respectively.

Table 5: Risk Base Result at Attock Station by Frequency Distribution

Year	Peak Discharge (Cusecs)	Return Period (Years)	5% Risk	10% Risk	25% Risk	50% Risk	75% Risk
1978	567000	9	0.435489662	0.894530482	2.442474596	5.884949192	11.76989838
1,992	572000	12	0.589501007	1.210882063	3.306258129	7.966167236	15.93233447
1998	575000	16	0.794770216	1.632521381	4.457525016	10.74005367	21.48010733
1995	578200	26	1.307811365	2.686348799	7.334952614	17.67298769	35.34597537
2010	994600	67	3.410939798	7.006342259	19.13049731	46.09341889	92.18683778

Table 6: Risk Base Result at Attock Station by Normal Distribution

Year	Peak Discharge (Cusecs)	Return Period (Years)	5% Risk	10% Risk	25% Risk	50% Risk	75% Risk
1978	567000	6	0.281334228	0.577882931	1.577882931	3.801784017	7.603568034
1,992	572000	6	0.281334228	0.577882931	1.577882931	3.801784017	7.603568034
1998	575000	6	0.281334228	0.577882931	1.577882931	3.801784017	7.603568034
1995	578200	6	0.281334228	0.577882931	1.577882931	3.801784017	7.603568034
2010	994600	500	25.62099199	52.62755999	143.6971472	346.226901	692.4538021

VI. CONCLUSION

From this research, it is concluded that the highest flood [2] recorded in the history of Pakistan should occur once with a probability of 55% to 77 % in the next 100 years. The Results of frequency distribution are adopted for future design purposes because the results of the frequency distribution method are more reasonable and realistic as compared to probability (Normal) distribution method. These probabilities of the highest Flood to occur in future are given in Table 7.

Table 7: Showing Probability of Highest Flood in Next Coming Years

Coming Years	Percentage Probability
Next 5 Years Probability	(3-7) %
Next 10 Years Probability	(7-13) %
Next 20 Years Probability	(14-25) %
Next 50 Years Probability	(33-52) %
Next 100 Years Probability	(55-77) %

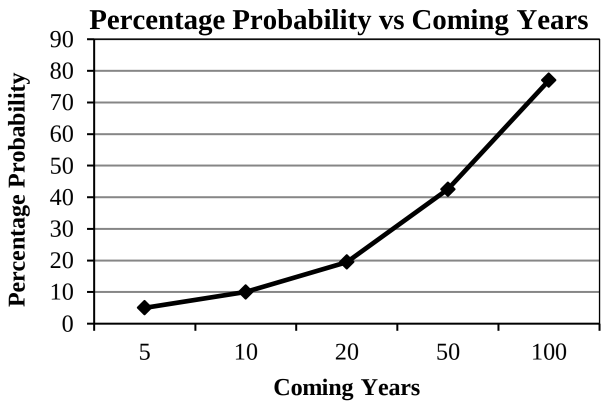

Figure 4: Showing Graphical View of Table 7

VII. RECOMMENDATIONS

Recommendations based on the findings of the study are:

- Research should be extended to other rivers of Pakistan.
- Research should be carried out using other methods as well.
- New structures should be designed as per Designed Flood.
- Existing structures that overtopped by the flood should be re-built.
- Annual Peak data shows different results at the same place using different methods therefore need of daily as well as hourly data should be made possible.
- During collection of Data, some of the gauge station data was absent therefore measurement of Data as well as availability of data at each gauge station should be made sure.
- Sophisticated equipment should be used for measurement of Discharge and for rainfall.
- During Data Collection from different departments, their data shows variability in their values so there should be coordination among different (hydrological, meteorological etc.) departments.

REFERENCES

[1] Glos, E., and R. Krause. "Estimating the accuracy of statistical flood values by means of long-term discharge records and historical data." (1969).

[2] Shaukat, Bilal, and Abdul Majid Khalil. "Vulnerability of Tourism Industry to Disasters:(A Case Study of July 2010 Flood in SWAT, Pakistan).

[3] Meyer, Michael D. Design standards for US transportation infrastructure: the implications of climate change. Washington, DC: Transportation Research Board, 2008.

[4] Melentijevich, M: Estimation Of Flood Flows Using Mathematical Statistics. Int. assoc. sci. hydrol. Publ. 84, pp164-172, 1967.

[5] Paulhus, J. L. H., and J. F. Miller: Flood Frequencies Derived From Rainfall Data, j. Hydraulic Div. ASCE, vol. 83, pap. 1451, Dec 1957.

[6] Stephen A. Nelson, River Systems & Causes of Flooding, Natural Disasters, Tulane University, EENS 3050.

[7] Chaudhary, Mohammad A., and Sally C. Stearns. "Estimating Confidence Intervals For Cost-effectiveness Ratios: An Example From A Randomized Trial." Statistics in medicine 15, no. 13 (1996): 1447-1458.

[8] Zhang, Qiang, Xihui Gu, Vijay P. Singh, Peng Sun, Xiaohong Chen, and Dongdong Kong. "Magnitude, frequency and timing of floods in the Tarim River basin, China: Changes, causes and implications." Global and Planetary Change 139 (2016): 44-55.