Probabilistic Analytics of Rainfall in Thailand for Lifespan of Hydraulic Structures

Kissada Im-oum[1] Wanwisa Duantrakoonsil[2] Rapin Sunthornwat[3*]
[1]Software Engineering and Information System Program, Faculty of Science and Technology, Pathumwan Institute of Technology
[2]College of Logistics and Supply Chain, Sripatum University
[3]Industrial Technology and Innovation Management Program, Faculty of Science and Technology, Pathumwan Institute of Technology
(* rapin@pit.ac.th)

Abstract

Rainfall is crucial to irrigation systems and agriculture worldwide, especially in the design and analysis of hydraulic structures. Analyzing the feasibility of rainfall in terms of risk and reliability is therefore essential for the analysis of such hydraulic structures. This research aims to compare the risk of rainfall affecting the designed hydraulic structures under binomial distribution and fractional distribution. The study found that under both binomial and fractional distributions, if the return period is constant, the risk increases as the lifespan of the structure increases. On the other hand, if the lifespan of the structure is constant, the risk decreases as the return period increases.

Keywords: binomial distribution, fractional binomial distribution, return period, risk and reliability, precipitation

1. Introduction

A rainfall refers specifically to the amount of water falling to the ground in the form of rain. It is one kind of precipitation which refers to the total amount of water falling, which may include rain, sleet, snow, or hail. Thailand is located in the Southeast Asian region, which has a hot and humid climate, so when reporting weather conditions or meteorological data the term of rainfall is often used in daily life. However, the terms of precipitation are more commonly used in research or academic contexts.

Rainfall plays a significant role in sustaining life, but when its amount exceeds a certain threshold, it can lead to disasters or cause damage to humans. Therefore, managing rainfall and precipitation is crucial. Landslides, for example, are often caused by excessive rainfall, and there are studies that warn about landslides due to heavy rainfall, helping to mitigate the potential risks (see [1]). Additionally, rainfall data is useful for managing water storage systems, allowing for the calculation of maximum rainfall to prepare for or control floodwaters efficiently (see [2]). As we can see, knowing the amount of rainfall enables better planning for disaster preparedness, which aligns with this research.

In this research, the return period is used to understand the occurrence of maximum rainfall, assessing the risk for hydraulic structures to make decisions aimed at enhancing safety. The binomial distribution is applied to determine the probability of the return period, followed by fractional binomial distribution to assess the risk associated with hydraulic structures.

2. Materials and Methods

Mathematical and statistical backgrounds and data collection for analysis of the probability in the precipitation are as follows:

2.1 The binomial distribution

Let X be a random variable which represents the number of successful times from all n trails, occurring with probability p. Thus, $n-x$ trails represents the number of failure times occurring with probability $1-p$ as q (see [3]).

Probability mass function (pmf) of the binomial distribution with the expected value $\mathbf{E}(X)$ and the variance $\mathbf{Var}(X)$ can be defined as:

$$\mathbf{P}(X=x)=\mathbf{P}_{x,n}=\frac{n!}{(n-x)!x!}p^{x}q^{n-x};\ x=0,1,2,...,n$$

$$\mathbf{E}(X)=\mu_X=np$$

$$\mathbf{Var}(X)=\sigma_X^2=npq \text{ ,where } q=1-p$$

2.2 The fractional binomial distribution

Let X be a random variable which extends from binomial distribution in case of the non-integer order of trails, α (see [4]).

Probability mass function (pmf) of the fractional distribution with the expected value $\mathbf{E}(X)$ and the variance $\mathbf{Var}(X)$ can be defined as:

$$\mathbf{P}(X = x) = \frac{\Gamma(\alpha+1)}{\Gamma(x+1)\Gamma(\alpha-x+1)} p^x (1-p)^{\alpha-x}; x = 0,1,2,...,\alpha$$

$$\mathbf{E}(X) = \mu_X = \alpha p$$

$$\mathbf{Var}(X) = \sigma_X^2 = \alpha pq \text{ ,where } q = 1-p$$

where $\Gamma(\alpha)$ is a Gamma function) defined as

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt \text{ for } \alpha > 0$$

The Gamma function corresponds to the following properties.

$\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$ for $\alpha > 0$ and α is a non-integer.

$\Gamma(n+1) = n!$ for $n \geq 0$ and n is an integer.

The main differences between the binomial distribution and the fractional binomial distribution are focused. The binomial distribution is a discrete probability distribution of the number of successes in a sequence of n independent experiments and each with its own Boolean - valued outcome. On the other hand, the fractional binomial distribution is a generalization of the binomial distribution where the number of trials α can be a non-integer (fractional) value.

In this research, X is a random variable representing the number of times it rains or the number of events. Let p is the probability of rain occurring at least a given number of times.

$$p = \mathbf{P}(X > x) = \frac{m-a}{N+b}$$

where m is the order number of event

N is the total number of even in data

and a,b are the Weibull parameters setting with $a = 0, b = 1$, see [5]. In the application of hydrology, the project lifespan or duration of hydraulic structures is represented by n for integer duration or α for non-integer duration.

Let T be return period which refers to the average time interval between events that have a magnitude or severity equal to or greater than a specified value. That is, $p = \frac{1}{T}$.

2.3 Data collection

Dataset for this research is gathered from the website of National Hydroinformatics Data Center [6]. The dataset is the annual rainfall of Thailand over the past 40 years, from 1985 to 2024.

2.4 Algorithm for computation of the probability of rain occurring at least a given number of times

For this research, the data analysis can be computed as the following steps.

Step 1. Given a dataset of N data points, sort the data from highest to lowest.

Step 2. Assign the highest value the rank $m = 1$, the next highest $m = 2$, and so on up to $m = N$.

Step 3. Calculate the return period $T = \frac{1}{p} = \frac{N+1}{m}$.

Step 4. Calculate the probability of the event occurring $p = \frac{1}{T}$.

2.5 Risk and Reliability

For this research, risk is defined as

Risk $= \mathbf{P}(X \geq 1) = 1 - \mathbf{P}(X = 0)$

Therefore,

Risk $= 1-(1-p)^n$ for the binomial distribution with integer n.

Risk $= 1-(1-p)^\alpha$ for the fractional binomial distribution with non-integer α.

In contrast, reliability is defined as

Reliability = 1 − Risk.

3. Results and discussion

This section consists of the results from data analysis of precipitation in Thailand from year 1985 to 2024. Also, the discussion is provided as the followings.

3.1 The maximum annual rainfall in Thailand

From Figure 1, it shows the maximum rainfall amount for each year, with the independent variable being the year of occurrence and the dependent variable being the annual rainfall amount. It can be observed that from 1985 to 2024, the maximum rainfall each year did not exceed 1600 millimeters (mm). Only in the years 2011, 2017, and 2021 did the annual rainfall surpass the amounts of other years.

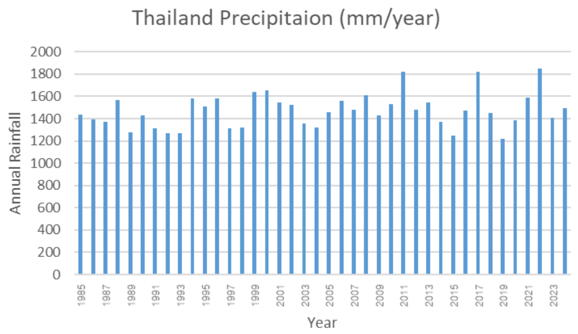

Figure 1: The maximum annual rainfall in Thailand

3.2 Return period and probability of the maximum rainfall in Thailand

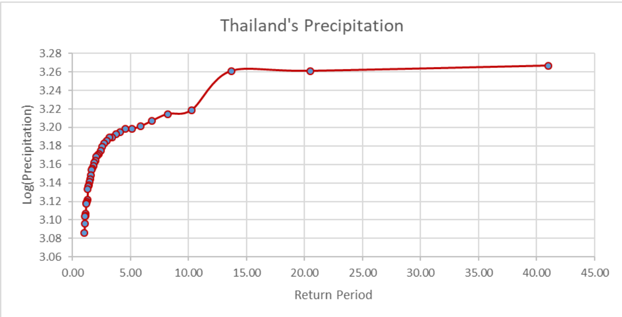

Figure 2: The maximum rainfall in Thailand

Figure 2 explains the maximum rainfall amount for the return period in Thailand, with the independent variable being the return period and the dependent variable being the maximum rainfall amount. As the return period increases, the maximum rainfall amount also increases. However, by the 14th return period, the maximum rainfall amount stabilizes at approximately 3.26 mm in log scales.

Table 1: Return period of the maximum rainfall in Thailand

Return period (years)	The maximum rainfall (mm.)
2	1465.98
6	1593.28
10	1680.25
14	1819.70
40	1862.09

From Table 1, it can be observed that if the return period increases, the maximum rainfall in Thailand will also increase.

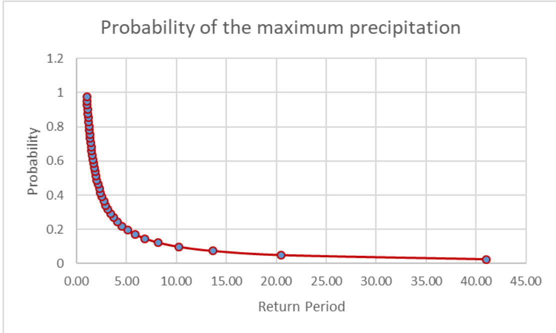

Figure 3: Probability of the maximum rainfall in Thailand

From Figure 3, it can be observed that the probability of maximum rainfall decreases as the return period increases.

3.3 Risk and reliability based on binomial distribution

For Table 2 based on binomial distribution, if return period is fixed, then risk increase as duration increase. If duration is fixed, then risk decrease as return period increase.

For Table 3 based on binomial distribution, if return period is fixed, then reliability decrease as duration increase. If duration is fixed, then reliability increase as return period increase.

For Table 4 based on fractional binomial distribution, if return period is fixed, then risk increase as duration increase. If duration is fixed, then risk decrease as return period increases.

For Table 5 based on fractional binomial distribution, if return period is fixed, then reliability decrease as duration increase. If duration is fixed, then reliability increase as return period increases.

3.4 Discussion

The research used binomial distribution (with n as an integer) and fractional binomial distribution (with α as a non-integer) to show that when the values of n and α are constant, as the return period (T) increases, the risk values diverge more. On the other hand, when the return period (T) decreases, the risk values for both distributions converge in the same direction. This is evident when comparing $n = 7$ and $\alpha = 6.5$, where the risk values for both distributions are similar. Additionally, when the return period (T) remains constant, as the values of n and α increase, the risk values for both distributions tend to become more similar. For application in real problem, hydraulic structure designers can choose risk and reliability to help decide the level of risk or reliability when determining the lifespan of hydraulic structures. For example, from Tables 2 and 4, when $n = 5, T = 2, p = 0.5$, the risk of the hydraulic structure is 0.969, or when $\alpha = 3.5, T = 2, p = 0.5$, the risk of the hydraulic structure is 0.912. Therefore, it can be concluded that the probabilities or outcomes of both equations are quite close to each other.

4. Conclusion

The conclusion for this research can be summarized as follows.

The maximum rainfall in Thailand over the past 40 years reached its highest at 1848 millimeters in 2022 and its lowest at 1218 millimeters in 2019. Overall, the average maximum rainfall in Thailand over the 40-year period is approximately 1472.7 millimeters. The maximum rainfall of 1465.98 millimeters in Thailand occurs every 2 years.

For both binomial distribution and fractional binomial distribution, if the return period is constant, the risk will decrease as the project lifespan increases. If the project lifespan is constant, the risk will decrease as the return period increases.

Therefore, the information resulting from the data analysis of this research enhances decision-making for the design and analysis of hydraulic structures such as dams, spillways, and sluice gates.

5. References

[1] Soralump S, Nuimak T. Estimation of Statistical Critical Rainfall Envelope for Landslide Warning. In: Proceedings of the 18th National Convention on Civil Engineering; 2013 May; Chiang Mai, Thailand. p. 8-10.

[2] Aueteerasrun P, Sukosi K, Thepprasit C, Marjang N. Probable mximum precipitation and probable maximum flood. In: Proceedings of the 11th THAICID National Symposium; 2018 Jun 20; Nonthaburi, Thailand.

[3] DeSanto , Moscatelli R, et al. Introduction to Statistics. 10th ed. New York (NY): Publisher; 2015.

[4] Hino M, Namba R. Fractional binomial distributions induced by the generalized binomial theorem and their applications. arXiv. 2024 Aug 21;2408.12011v1.

[5] Weibull W. A statistical theory of strength of materials. Ing Vetensk Akad Handl. 1939; 151: 1-45.

[6] National Hydroinformatics Data Center. Hydro Situation [document on the Internet]. Bankok: Hydro - Informatics Institute (HII); 2024 [updated 2024 January 24; cited 2024 Feb 15]. Available from: https://www.thaiwater.net.

Table 2: Risk of the maximum rainfall based on binomial distribution

T	p	**Integer duration (n)**													
		2	**5**	**7**	**10**	**15**	**20**	**25**	**30**	**35**	**40**	**45**	**50**	**55**	**60**
2	0.5	0.750	0.969	0.992	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5	0.2	0.360	0.672	0.790	0.893	0.965	0.988	0.996	0.999	1.000	1.000	1.000	1.000	1.000	1.000
10	0.1	0.190	0.410	0.522	0.651	0.794	0.878	0.928	0.958	0.975	0.985	0.991	0.995	0.997	0.998
25	0.04	0.078	0.185	0.249	0.335	0.458	0.558	0.640	0.706	0.760	0.805	0.841	0.870	0.894	0.914
50	0.02	0.040	0.096	0.132	0.183	0.261	0.332	0.397	0.455	0.507	0.554	0.597	0.636	0.671	0.702
100	0.01	0.020	0.049	0.068	0.096	0.140	0.182	0.222	0.260	0.297	0.331	0.364	0.395	0.425	0.453
500	0.002	0.004	0.010	0.014	0.020	0.030	0.039	0.049	0.058	0.068	0.077	0.086	0.095	0.104	0.113
1000	0.001	0.002	0.005	0.007	0.010	0.015	0.020	0.025	0.030	0.034	0.039	0.044	0.049	0.054	0.058

Table 3: Reliability of the maximum rainfall based on binomial distribution

T	p	**Integer duration (n)**													
		2	**5**	**7**	**10**	**15**	**20**	**25**	**30**	**35**	**40**	**45**	**50**	**55**	**60**
2	0.5	0.250	0.031	0.008	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
5	0.2	0.640	0.328	0.210	0.107	0.035	0.012	0.004	0.001	0.000	0.000	0.000	0.000	0.000	0.000
10	0.1	0.810	0.590	0.478	0.349	0.206	0.122	0.072	0.042	0.025	0.015	0.009	0.005	0.003	0.002
25	0.04	0.922	0.815	0.751	0.665	0.542	0.442	0.360	0.294	0.240	0.195	0.159	0.130	0.106	0.086
50	0.02	0.960	0.904	0.868	0.817	0.739	0.668	0.603	0.545	0.493	0.446	0.403	0.364	0.329	0.298
100	0.01	0.980	0.951	0.932	0.904	0.860	0.818	0.778	0.740	0.703	0.669	0.636	0.605	0.575	0.547
500	0.002	0.996	0.990	0.986	0.980	0.970	0.961	0.951	0.942	0.932	0.923	0.914	0.905	0.896	0.887
1000	0.001	0.998	0.995	0.993	0.990	0.985	0.980	0.975	0.970	0.966	0.961	0.956	0.951	0.946	0.942

Table 4: Risk of the maximum rainfall based on fractional binomial distribution

T	p	Non-integer duration (α)													
		0.5	**3.5**	**6.5**	**8.5**	**13.5**	**18.5**	**23.5**	**28.5**	**33.5**	**38.5**	**43.5**	**48.5**	**53.5**	**58.5**
2	0.5	0.293	0.912	0.989	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5	0.2	0.106	0.542	0.766	0.850	0.951	0.984	0.995	0.998	0.999	1.000	1.000	1.000	1.000	1.000
10	0.1	0.051	0.308	0.496	0.592	0.759	0.858	0.916	0.950	0.971	0.983	0.990	0.994	0.996	0.998
25	0.04	0.020	0.133	0.233	0.293	0.424	0.530	0.617	0.688	0.745	0.792	0.831	0.862	0.887	0.908
50	0.02	0.010	0.068	0.123	0.158	0.239	0.312	0.378	0.438	0.492	0.541	0.585	0.625	0.661	0.693
100	0.01	0.005	0.035	0.063	0.082	0.127	0.170	0.210	0.249	0.286	0.321	0.354	0.386	0.416	0.445
500	0.002	0.001	0.007	0.013	0.017	0.027	0.036	0.046	0.055	0.065	0.074	0.083	0.093	0.102	0.111
1000	0.001	0.001	0.003	0.006	0.008	0.013	0.018	0.023	0.028	0.033	0.038	0.043	0.047	0.052	0.057

Table 5: Reliability of the maximum rainfall based on fractional binomial distribution

T	p	Non-integer duration (α)													
		0.5	**3.5**	**6.5**	**8.5**	**13.5**	**18.5**	**23.5**	**28.5**	**33.5**	**38.5**	**43.5**	**48.5**	**53.5**	**58.5**
2	0.5	0.707	0.088	0.011	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
5	0.2	0.894	0.458	0.234	0.150	0.049	0.016	0.005	0.002	0.001	0.000	0.000	0.000	0.000	0.000
10	0.1	0.949	0.692	0.504	0.408	0.241	0.142	0.084	0.050	0.029	0.017	0.010	0.006	0.004	0.002
25	0.04	0.980	0.867	0.767	0.707	0.576	0.470	0.383	0.312	0.255	0.208	0.169	0.138	0.113	0.092
50	0.02	0.990	0.932	0.877	0.842	0.761	0.688	0.622	0.562	0.508	0.459	0.415	0.375	0.339	0.307
100	0.01	0.995	0.965	0.937	0.918	0.873	0.830	0.790	0.751	0.714	0.679	0.646	0.614	0.584	0.555
500	0.002	0.999	0.993	0.987	0.983	0.973	0.964	0.954	0.945	0.935	0.926	0.917	0.907	0.898	0.889
1000	0.001	0.999	0.997	0.994	0.992	0.987	0.982	0.977	0.972	0.967	0.962	0.957	0.953	0.948	0.943