

# Flood Return Interval

Marc Los Huertos

2025-04-05

## Basic Hydrology

“/home/mwl04747/RTricks/11\_Flood\_Return\_Interval/Regional\_Climate\_Trends/USGS\_Discharge\_20250405.png”

### Base and Flood Flow

Baseflow is the portion of streamflow that is sustained between precipitation events, fed to streams by delayed pathways such as groundwater flow.

Flood flow is the portion of streamflow that is generated by precipitation events, and is typically associated with surface runoff.

### Return Interval

In hydrology, the **flood interval** (or **return period**) represents the average time between floods of a certain magnitude. The recurrence interval, or return period, of a flood is the average time, in years, between floods of a certain magnitude. For example, a “100-year flood” means that, statistically, a flood of that magnitude has a 1% chance of occurring in any given year.’

It is calculated by taking the reciprocal of the probability of the event occurring in any given year.

It’s important to remember that recurrence intervals are statistical averages, and a flood of a certain magnitude could occur more or less frequently than the average.

Recurrence intervals can be influenced by factors like climate change and changes in land use, such as urbanization, deforestation, etc.

The return interval (or recurrence interval) is the average time between events of a certain size or larger. For example, a 100-year flood is an event that has a 1% chance of occurring in any given year.

The return interval can be calculated using the following formula:

$$RI = \frac{N + 1}{M}$$

Where: -  $RI$  = return interval -  $N$  = number of years of record -  $M$  = rank of the event (1 for the largest event, 2 for the second largest, etc.) -  $P$  = probability of exceedance

## Methods for Estimating Flood Intervals

There are several methods used in hydrology to estimate flood intervals or return periods. These methods vary in complexity and applicability depending on the available data and purpose. Below are some of the most common approaches:

### 1. Empirical Method (Plotting Position Formula)

This is the simplest method and involves ranking annual peak discharges and applying a plotting position, where the return period  $T$  is defined as:

$$T = \frac{n + 1}{m}$$

Where:

- $n$  is the number of years of data
- $m$  is the rank ( $1 = \text{largest}$ )

This method does not assume any underlying distribution.

**2. Gumbel Distribution (Extreme Value Type I)** A commonly used theoretical distribution for modeling annual maxima is the Gumbel distribution. The return period is related to the probability of exceedance  $P$  as:

$$P = \frac{1}{T}$$

The Gumbel cumulative distribution function (CDF) is:

$$F(x) = \exp \left[ - \exp \left( - \frac{x - \mu}{\beta} \right) \right]$$

Where  $\mu$  is the location parameter and  $\beta$  is the scale parameter. These can be estimated using method of moments or maximum likelihood.

**3. Log-Pearson Type III Distribution** This method is widely recommended in the United States for flood frequency analysis. It applies a Pearson Type III distribution to the log-transformed discharge data. The procedure involves:

- Taking the logarithm of the discharge data
- Calculating the mean, standard deviation, and skewness
- Using these to estimate quantiles for different return periods

```
# Load necessary library
if (!require(e1071)) install.packages("e1071", dependencies = TRUE)

## Loading required package: e1071
library(e1071) # for skewness function

# Step 1: Input annual peak discharge data (in m^3/s)
discharge <- c(400, 520, 610, 700, 650, 720, 800, 750, 680, 670)
```

```

# Step 2: Log-transform the discharge data
logQ <- log10(discharge)

# Step 3: Compute statistics of log-values
mean_logQ <- mean(logQ)
sd_logQ <- sd(logQ)
skew_logQ <- skewness(logQ)

# Step 4: Define return periods and calculate exceedance probabilities
T <- c(2, 5, 10, 25, 50, 100)      # Return periods
P <- 1 - 1 / T                      # Exceedance probability
z <- qnorm(P)                       # Standard normal variate

# Step 5: Apply Log-Pearson Type III quantile formula
K <- z + ( (z^2 - 1) * skew_logQ ) / 6 # Pearson III frequency factor

logQ_est <- mean_logQ + K * sd_logQ  # Estimated logQ
Q_est <- 10^logQ_est                # Convert back to original scale

# Combine results in a table
lp3_results <- data.frame(
  Return_Period = T,
  Exceedance_Prob = round(P, 4),
  Discharge_Estimate = round(Q_est, 2)
)

```

```

library(xtable)
print(xtable(lp3_results))

```

```

## % latex table generated in R 4.2.2 by xtable 1.8-4 package
## % Sat Apr 5 18:39:28 2025
## \begin{table}[ht]
## \centering
## \begin{tabular}{rrrr}
## \hline
## & Return\_Period & Exceedance\_Prob & Discharge\_Estimate \\
## \hline
## 1 & 2.00 & 0.50 & 663.81 \\
## 2 & 5.00 & 0.80 & 766.28 \\
## 3 & 10.00 & 0.90 & 808.48 \\
## 4 & 25.00 & 0.96 & 842.30 \\
## 5 & 50.00 & 0.98 & 857.23 \\
## 6 & 100.00 & 0.99 & 865.72 \\
## \hline
## \end{tabular}
## \end{table}

```

**4. Bayesian Methods** Bayesian approaches allow incorporating prior information and uncertainty in parameter estimates. They are especially useful for short or regional datasets and are implemented via MCMC simulations.

**5. Regional Frequency Analysis (RFA)** Instead of analyzing one site in isolation, RFA uses data from multiple similar sites in a region to improve flood frequency estimates. The Index Flood method is a popular RFA technique.

**6. Generalized Extreme Value (GEV) Distribution** GEV unifies the Gumbel, Fréchet, and Weibull families under one formulation:

$$F(x) = \exp \left\{ - \left[ 1 + \xi \left( \frac{x - \mu}{\sigma} \right) \right]^{-1/\xi} \right\}$$

Where:

- $\mu$  is the location
- $\sigma$  is the scale
- $\xi$  is the shape parameter

These methods are often evaluated based on goodness-of-fit tests, diagnostic plots, and expert judgment.