

Is the Big Flood Due? Ranking 24 Years of River Peaks

Learning target: I can rank a river's yearly flood peaks, calculate recurrence intervals and yearly chances, graph them, and explain why a '25-year' or '100-year' flood is a probability each year, not a schedule.

Data and graphing About 50 min Individual MS-ESS3-2 · assesses data analysis independent practice

Standards connections

MS-ESS3-2 (assesses, the data-analysis and forecasting half, applied to floods): Students analyze the magnitudes and frequencies of past floods, the kinds of data the clarification statement names, to forecast the yearly chance of floods of different sizes and to recommend how high a new building should sit. The forecast and recommendation they write in q7 and q8 are direct evidence toward the PE. The data are simulated, so the sheet assesses the reasoning, not real-world conclusions.

Purpose: practice · **Time:** about 50 minutes · **Format:** individual

Answers with reasoning

The river record

A streamgauge on Arroyo Verde, a river that runs past the town of Mission Flats, has measured flow since 2001. Flow is the volume of water passing the gage each second, in cubic meters per second (m³/s). When flow rises above 250 m³/s, the river spills over its banks onto the floodplain where part of the town is built. Like most California rivers, Arroyo Verde has its biggest peaks in winters when atmospheric rivers bring days of heavy rain, and small peaks in drought years.

Annual peak flow of Arroyo Verde at Mission Flats, 2001-2024 (teaching dataset)

Simulated teaching dataset: the numbers are invented to be realistic, not measured.

Year	Peak flow (m³/s)	Year	Peak flow (m³/s)
2001	140	2013	50
2002	95	2014	40
2003	210	2015	85
2004	120	2016	190
2005	480	2017	620
2006	360	2018	130
2007	60	2019	410
2008	230	2020	70

Year	Peak flow (m ³ /s)	Year	Peak flow (m ³ /s)
2009	110	2021	45
2010	260	2022	160
2011	300	2023	550
2012	75	2024	280

Simulated data for teaching. The pattern (wet-winter spikes, drought-year lows, big years close together) is realistic, but the river, the town and the values are invented.

How hydrologists turn a record into chances — List all the annual peaks from largest to smallest and give each a rank, m (largest = 1). With n years of record, the recurrence interval of a flood is $T = (n + 1) \div m$ years. The chance of a flood at least that big in any one year is $1 \div T$, written as a percent. Here n = 24, so n + 1 = 25. Example: the 7th largest peak has $T = 25 \div 7 = 3.6$ years and a yearly chance of $7 \div 25 = 28$ percent.

Recurrence intervals and yearly chances (real reference values)

Recurrence interval (years)	Chance in any one year
2	50%
5	20%
10	10%
25	4%
50	2%
100	1%
500	0.2%

These pairs follow from chance = $1 \div T$ and match the table the U.S. Geological Survey uses to explain flood probabilities.

Part A: Rank and calculate

1. Rank all 24 peaks on scrap paper. Then fill in this table for the ranks shown. Show one full calculation of T and of the yearly chance below the table.

Answer: Rank 1: 2017, 620 m³/s, $T = 25 \div 1 = 25.0$ years, chance = $1 \div 25 = 4\%$. Rank 2: 2023, 550, $T = 12.5$ years, 8%. Rank 3: 2005, 480, $T = 8.3$ years, 12%. Rank 4: 2019, 410, $T = 6.3$ years (6.25), 16%. Rank 5: 2006, 360, $T = 5.0$ years, 20%. Rank 8: 2010, 260, $T = 3.1$ years (3.125), 32%. Rank 12: 2022, 160, $T = 2.1$ years (2.08), 48%. Rank 20: 2020, 70, $T = 1.3$ years (1.25), 80%.

Why: Full ranking from largest: 620 (2017), 550 (2023), 480 (2005), 410 (2019), 360 (2006), 300 (2011), 280 (2024), 260 (2010), 230 (2008), 210 (2003), 190 (2016), 160 (2022), 140 (2001), 130 (2018), 120 (2004), 110 (2009), 95 (2002), 85 (2015), 75 (2012), 70 (2020), 60 (2007), 50 (2013), 45 (2021), 40 (2014). The chance equals $m \div 25$.

Also acceptable: Accept T rounded to whole numbers (8, 6, 3, 2, 1) if the rounding is consistent, and chances computed from the rounded T if the method is correct.

- Correct year and flow for each rank
- $T = 25 \div m$ for each row
- Chance = $1 \div T$ as a percent for each row
- One full calculation shown

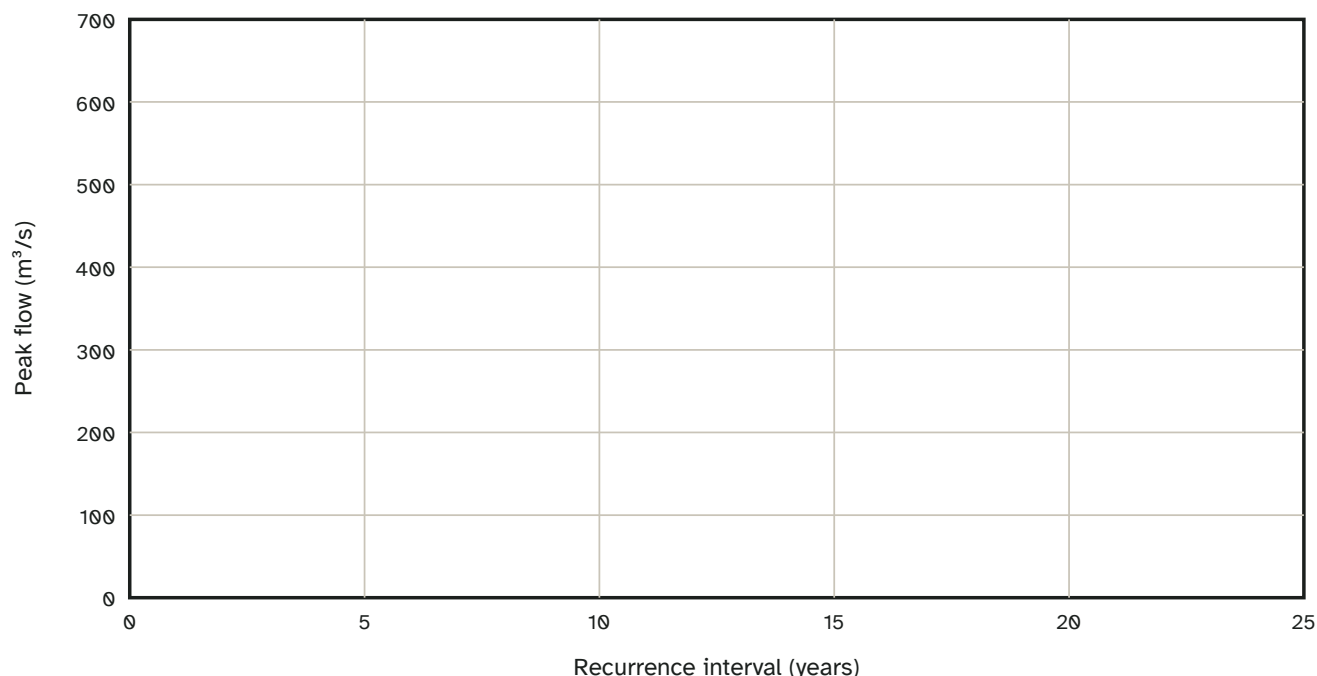
2. In how many of the 24 years did the river spill onto the floodplain (flow above 250 m³/s)? Use the rank of the smallest of those floods to find the recurrence interval and yearly chance of an 'overflow year.'

Answer: 8 years were above 250 m³/s: 2017 (620), 2023 (550), 2005 (480), 2019 (410), 2006 (360), 2011 (300), 2024 (280) and 2010 (260). The smallest of these, 260, is rank 8, so $T = 25 \div 8 = 3.1$ years and the yearly chance is $8 \div 25 = 32$ percent, roughly one year in three.

Why: An overflow year is any year whose peak is at least 260 here, so the eighth-ranked peak sets the threshold.

Also acceptable: Accept $8 \div 24 = 33$ percent with the explanation that 8 of the 24 years overflowed; both describe about one year in three.

Arroyo Verde: peak flow versus recurrence interval



Part B: Interpret and forecast

3. Describe the shape of your curve. What happens to the size of the flood as the recurrence interval gets longer, and is the increase steady or does it change?

Answer: The curve rises from about $70 \text{ m}^3/\text{s}$ at 1.3 years to $620 \text{ m}^3/\text{s}$ at 25 years. It rises steeply at short recurrence intervals and then flattens: going from a 1-year to a 5-year flood adds almost $300 \text{ m}^3/\text{s}$, but going from a 12.5-year to a 25-year flood adds only about $70 \text{ m}^3/\text{s}$. Rarer floods are bigger, but each extra year of recurrence interval adds less flow.

Why: Flood-frequency curves are usually steep at short intervals and flatter at long intervals on a linear time axis.

Also acceptable: Accept 'it curves and levels off' with numbers.

- Bigger floods have longer recurrence intervals
- The increase is not steady: steep at first, then flatter
- Uses at least one pair of values

4. Use your curve to estimate the peak flow of a 10-year flood on Arroyo Verde. Explain how you read it from the graph.

Answer: About $500 \text{ m}^3/\text{s}$. Find 10 years on the horizontal axis, go up to the curve, and read across to the flow axis. The value falls between the rank 3 point (8.3 years, $480 \text{ m}^3/\text{s}$) and the rank 2 point (12.5 years, $550 \text{ m}^3/\text{s}$); straight-line interpolation gives $480 + (10 - 8.3) \div (12.5 - 8.3) \times 70 \approx 508 \text{ m}^3/\text{s}$.

Why: Estimates depend on the hand-drawn curve.

Also acceptable: Accept 480 to 530 m^3/s with a clear description of reading the graph.

5. A 25-year flood has a 4 percent chance each year, so the chance of **no** 25-year flood in a year is 96 percent. (a) What is the chance of no 25-year flood in either of the next two years? (b) What is the chance of at least one in the next two years? (c) What is the chance of 25-year floods in **both** of the next two years? Show your multiplication.

Answer: (a) $0.96 \times 0.96 = 0.9216$, so about 92 percent. (b) $1 - 0.9216 = 0.0784$, about 7.8 percent. (c) $0.04 \times 0.04 = 0.0016$, about 0.16 percent, small but not zero.

Why: Each year is treated as independent, so chances for separate years multiply. Part (c) shows that back-to-back rare floods are unlikely but possible, as 2017 and 2019 (both in the top four) and 2023 suggest.

Also acceptable: Accept 92%, 8% and 0.2% when rounded.

6. The town wants to know the size of Arroyo Verde's 100-year flood. Why can this 24-year record not answer that question well? Name one kind of extra data that would help.

Answer: The record covers only 24 years, so the 100-year flood is far outside it. The largest flood in the record has $T = 25$ years; anything rarer has to be estimated by stretching the curve beyond the data, which is very uncertain. Helpful extra data: longer records from nearby gages, historical accounts and high-water marks from older floods, flood deposits in the floodplain's soil layers, or computer models of very large storms.

Why: Estimating a 1 percent flood from 24 years is extrapolation; hydrologists combine longer and regional records.

Also acceptable: Accept 'we would have to guess past the end of the graph' with one data source.

- States that the record is too short compared with 100 years
- Names one useful source of extra data

7. A resident says, 'The 2017 flood was the biggest in 24 years, a 25-year flood. So 2023 must have been a fluke, and the next flood that big won't come until about 2042.' Use the dataset and your calculations to explain what is wrong with this reasoning, and write a correct forecast statement for floods of that size.

Answer: The resident treats the recurrence interval as a schedule. A 25-year flood has a 4 percent chance every year, no matter when the last one happened, so it can happen two years in a row or not for decades. The dataset shows this: 2023 ($550 \text{ m}^3/\text{s}$) was nearly as big as 2017 ($620 \text{ m}^3/\text{s}$) only six years later, and 2019 ($410 \text{ m}^3/\text{s}$) was the fourth-biggest peak, two years after 2017. 'About 2042' has no special meaning. A correct forecast: 'Based on 24 years of records, a flood at least as large as $620 \text{ m}^3/\text{s}$ has about a 4 percent chance of happening in any given year, including next year; over the next two years the chance is about 8 percent.'

Why: The PE asks students to forecast from frequency data; a forecast is a probability over a time window, not a date.

Also acceptable: Accept the point that a 24-year record gives only a rough estimate of the 25-year flood, as long as the schedule error is also addressed.

- Identifies the error: treating the interval as a schedule
- States the 4 percent yearly chance applies every year
- Uses specific years and flows from the dataset
- Writes a forecast as a probability over a time window

8. Mission Flats will build a new fire station that should last 50 years. Over 50 years, the chance of at least one 25-year flood is about 87 percent, and the chance of at least one 100-year flood is about 39 percent. Recommend where or how high the station should be built, and use these numbers to justify your choice.

Answer: Example: Build the fire station outside the floodplain on higher ground, or raise its floor well above the height of a 100-year flood. Designing only for a 25-year flood is not enough, because there is about an 87 percent chance such a flood will happen during the station's life. Even the 100-year flood has about a 39 percent chance over 50 years, and a fire station must work during floods, so it should sit above that level or out of the floodplain entirely.

Why: The chances come from $1 - 0.96^{50} \approx 0.87$ and $1 - 0.99^{50} \approx 0.39$. The recommendation uses forecast probabilities to inform mitigation, which is the second half of the PE.

Also acceptable: Accept a recommendation to build on stilts or a raised pad, or to protect it with a levee designed above the 100-year flood, if justified with the numbers; note that levees can fail or be overtopped.

- Clear recommendation (location or design height)
- Uses the 87 percent figure to reject designing for the 25-year flood
- Uses the 39 percent figure or the station's emergency role to justify a higher standard

Teacher notes

The dataset is simulated. The recurrence-interval formula $T = (n + 1) \div m$ is the simple ranking method the U.S. Geological Survey uses to teach flood frequency; agencies use more complex statistics for official estimates. The reference table is accurate: the USGS pairs 2, 5, 10, 25, 50, 100 and 500-year intervals with 50, 20, 10, 4, 2, 1 and 0.2 percent yearly chances, and it is shifting to the term annual exceedance probability. The 50-year chances in q8 are $1 - 0.96^{50} = 0.870$ and $1 - 0.99^{50} = 0.395$. Common errors: dividing m by 25 to get T (that gives the chance), ranking smallest as 1, and connecting the graph points dot to dot. Plan to discuss q7 aloud; the idea that a rare flood is 'due' or 'used up' is very persistent and returns in W7.

Scaffold offered to students: Ranking without getting lost

Copy the 24 flows onto 24 small slips of paper with their years, then lay the slips in a line from biggest to smallest and number them 1 to 24. For each rank in q1, divide 25 by the rank to get T, then divide the rank by 25 and multiply by 100 to get the yearly chance. Check: T times the chance (as a decimal) should always equal 1. For example, rank 5 gives $T = 5.0$ years and a chance of 20 percent, and $5.0 \times 0.20 = 1$.

Extension: A changing river

Suppose Mission Flats paved large areas upstream between 2010 and 2015. Explain how more runoff could make recent peaks larger than older ones even with the same storms, and why that would make the whole 24-year record less trustworthy for forecasting. Then propose how a hydrologist could check whether the river's behavior has changed, using only the dataset on this page.