# Surface Water Information Modeling System (SWIMS)

# Gap Tolerance Analysis

**Author: Cheng-Wei Huang**

## Executive Summary

**Purpose**
The Oregon Water Resources Department (OWRD) relies on flow duration curves (FDCs) to characterize natural streamflow for water-availability evaluations. This study establishes defensible gap-tolerance criteria—how much missing data can be present while keeping FDCs reliable.

**Data and scope**
We analyzed 51 minimally disturbed index gages with continuous daily records over WY 1991–2020. Gages were selected from low-disturbance HUC-12 watersheds to reflect near-natural hydrologic conditions.

**Method overview**
For each gage, the baseline 30-year FDC was computed from complete data. We then simulated missing data by randomly selecting 15 of the 30 years and removing 10–75% of daily values within those years while retaining all data in the other 15 years (overall retention range: 62.5–95%). Each gap level was iterated 100 times. We compared the mean FDC (±1 SD) to baseline at 20%, 50%, and 80% exceedance. Shannon entropy of monthly flows was used to interpret sensitivity.

**Key findings**

- For the majority of index gages, FDCs remain within ±15% at key percentiles when ≥80% of the 30-year period is retained (≈ 60% completeness within the gapped 15 years, with the other 15 intact).
- Lower-entropy (more skewed) regimes are more sensitive and typically require ≥82.5–95% overall retention to satisfy the same ±15% criterion.

**Program implications for "index gages"**
OWRD previously assumed complete records for every year in the 30-year base period. Results indicate that the candidate pool can be expanded to include relatively minimally disturbed gages with less than 30 years of complete record, provided overall retention ≥80% (and higher thresholds for sensitive, low-entropy sites).

**Recommendations**

1. Baseline acceptance: require ≥80% overall retention.
2. Sensitive sites (low entropy): require ≥85–90% (up to 95%) overall retention.
3. Document entropy (or comparable distribution diagnostics) to flag gages needing stricter thresholds.

These criteria enable broader, defensible use of available records while maintaining FDC integrity for water-availability and related hydrologic analyses.

# Introduction

At the Oregon Water Resources Department (OWRD, referred to as "the Department"), flow duration curves (FDCs) are a fundamental analytical tool for characterizing natural streamflow to support water availability evaluations. FDCs provide critical insight into streamflow variability, which informs resource management decisions, including allocation, conservation, and policy development. Despite their importance, the effect of incomplete streamflow records on FDC accuracy is not well understood, and standardized criteria for allowable data gaps have not been formally established.

This analysis seeks to address these knowledge gaps by determining the maximum percentage of missing data that can be tolerated while preserving the integrity of FDCs. Specifically, FDCs generated from datasets with varying levels of artificially introduced missing data (ranging from 5% to 37.5%) are compared against FDCs derived from complete datasets to identify an acceptable threshold. The analysis identifies a gap tolerance level as the maximum proportion of missing data that results in deviations of no more than 15% from FDCs based on complete data. The findings are based on streamflow data from 51 selected index gages, defined as sites with continuous 30-year mean-daily records (WY 1991–2020) located in minimally impacted watersheds—classified as low or low-to-moderate disturbance in Andrews & Huang (2024).

# Methods

## Data

The analysis utilized mean daily streamflow data from 51 index gages selected for their ability to represent low-disturbance hydrologic conditions. Each gage provides a continuous 30-year record spanning water years 1991 through 2020, ensuring sufficient temporal coverage for robust flow duration curve (FDC) analysis. These gages were strategically chosen to minimize anthropogenic influences and ensure that the streamflow data reflect natural hydrologic conditions, which are essential for accurate and reliable FDC computations.

To achieve this, all selected index gages are situated within minimally impacted HUC 12 sub-watersheds. The degree of anthropogenic disturbance in each sub-watershed was assessed in Andrews and Huang (2024) using a methodology adapted from Falcone et al. (2010, 2011), which incorporates multiple metrics to evaluate human impacts on hydrologic systems. While various disturbance variables were considered, priority was given to avoiding impacts from surface water withdrawals and dam storage. Sub-watersheds were deemed relevant (i.e., low disturbance) if they were classified as low or low-to-moderate disturbance, or if they met the following thresholds: an irrigation withdrawal ratio (ratioed by natural flow estimates using the National Hydrography Dataset [NHDPlus V2.1]) of less than or equal to 0.10 and a storage ratio (also ratioed by natural flow estimates) of less than or equal to 0.05. This rigorous selection process ensured that the resulting dataset excluded significant hydrologic disruptions, which could otherwise compromise the validity of the FDCs. By focusing on minimally impacted sub-watersheds, the selected data provides a reliable basis for evaluating the sensitivity of FDCs to missing streamflow records and establishing defensible gap tolerance levels.

Table 1. List of index gages with low disturbance.

| Station number | Station name | Latitude | Longitude | Basin |
|---|---|---|---|---|
| **13216500** | N Fk Malheur R ab Beulah Res nr Beulah, OR | 43.948797 | -118.172772 | Malheur |
| **13331500** | Minam R nr Minam, OR | 45.61958739 | -117.7243687 | Grande Ronde |
| **14010000** | S Fk Walla Walla R nr Milton, OR | 45.829925 | -118.170003 | Umatilla |
| **14010800** | N Fk Walla Walla R nr Milton Freewater, OR | 45.885003 | -118.201097 | Umatilla |
| **14032000** | Butter Cr nr Pine City, OR | 45.54650278 | -119.3049056 | Umatilla |
| **14020000** | Umatilla R ab Meacham Cr nr Gibbon, OR | 45.71874559 | -118.3235713 | Umatilla |
| **14044000** | M Fk John Day R at Ritter, OR | 44.88848965 | -119.1410902 | John Day |
| **14091500** | Metolius R nr Grandview, OR | 44.62622717 | -121.4831067 | Deschutes |
| **14074900** | Snow Cr nr Sisters, OR | 44.116472 | -121.662323 | Deschutes |
| **14057500** | Fall R nr La Pine | 43.796484 | -121.572753 | Deschutes |
| **14055600** | Odell Cr nr La Pine | 43.575578 | -121.881118 | Deschutes |
| **14054500** | Browns Cr nr La Pine, OR | 43.723513 | -121.796139 | Deschutes |
| **14050000** | Deschutes R bl Snow Cr nr La Pine, OR | 43.814069 | -121.777198 | Deschutes |
| **14052500** | Quinn R nr La Pine, OR | 43.783542 | -121.836581 | Deschutes |
| **14050500** | Cultus R ab Cultus Cr nr La Pine, OR | 43.81806 | -121.795863 | Deschutes |
| **14051000** | Cultus Cr ab Crane Prairie Res nr La Pine, OR | 43.82094 | -121.823488 | Deschutes |
| **14052000** | Deer Cr ab Crane Prairie Res nr La Pine, OR | 43.804606 | -121.839252 | Deschutes |
| **14097100** | Warm Springs R nr Kahneeta Hot Springs, OR | 44.85623014 | -121.1494931 | Deschutes |
| **14096850** | Beaver Cr bl Quartz Cr nr Simnasho, OR | 44.95789576 | -121.3942242 | Deschutes |
| **14092750** | Shitike Cr at Peters Pasture nr Warm Springs, OR | 44.75122815 | -121.6331176 | Deschutes |
| **14137000** | Sandy R nr Marmot, OR | 45.39095523 | -122.1300863 | Sandy |
| **14139800** | S Fk Bull Run R nr Bull Run, OR | 45.44456683 | -122.1095256 | Sandy |
| **14138800** | Blazed Alder Cr nr Rhododendron, OR | 45.45261842 | -121.891747 | Sandy |
| **14153800** | Layng Cr ab Prather Cr nr Disston, OR | 43.7092 | -122.72757 | Willamette |
| **14154500** | Row R ab Pitcher Cr nr Dorena, OR | 43.73456751 | -122.8706253 | Willamette |
| **14158500** | McKenzie R at Outlet of Clear Lake, OR | 44.36095405 | -121.9939477 | Willamette |
| **14161500** | Lookout Cr nr Blue R, OR | 44.20984768 | -122.2572887 | Willamette |
| **14178000** | N Santiam R bl Boulder Cr nr Detroit, OR | 44.70762303 | -122.102573 | Willamette |
| **14182500** | Little N Santiam R nr Mehama, OR | 44.79095516 | -122.5795271 | Willamette |
| **14187000** | Wiley Cr nr Foster, OR | 44.37262371 | -122.6248055 | Willamette |
| **14185000** | S Santiam R bl Cascadia, OR | 44.391518 | -122.5098066 | Willamette |
| **14185900** | Quartzille Cr nr Cascadia, OR | 44.53984671 | -122.4359162 | Willamette |
| **14210000** | Clackamas R at Estacada, OR | 45.30039964 | -122.3536978 | Willamette |
| **14209500** | Clackamas R ab Three Lynx Cr, OR | 45.12484321 | -122.0734113 | Willamette |
| **14301000** | Nehalem R nr Foss, OR | 45.70399837 | -123.755404 | North Coast |
| **14303600** | Nestucca R nr Beaver, OR | 45.26621573 | -123.8470605 | North Coast |
| **14301500** | Wilson R nr Tillamook, OR | 45.48399439 | -123.6901199 | North Coast |
| **14301300** | Miami R nr Garibaldi, OR | 45.575115 | -123.874433 | North Coast |
| **14306030** | Yaquina R nr Chitwood, OR | 44.657397 | -123.83875 | Mid Coast |

| | | | | |
|---|---|---|---|---|
| **14305500** | Siletz R at Siletz, OR | 44.71456114 | -123.8862182 | Mid Coast |
| **14306500** | Alsea R nr Tidewater, OR | 44.38817887 | -123.8340015 | Mid Coast |
| **14306340** | E Fk Lobster Cr nr Alsea, OR | 44.25012201 | -123.6373243 | Mid Coast |
| **14316700** | Steamboat Cr nr Glide, OR | 43.34900889 | -122.7289393 | Umpqua |
| **14309500** | W Fk Cow Cr nr Glendale, OR | 42.80455873 | -123.6103463 | Umpqua |
| **14308000** | S Umpqua R at Tiller, OR | 42.93040021 | -122.9475532 | Umpqua |
| **14325000** | S Fk Coquille R at Powers, OR | 42.89122414 | -124.0698188 | South Coast |
| **14332000** | S Fk Rogue R nr Prospect, OR | 42.70790846 | -122.3939192 | Rogue |
| **14330000** | Rogue R bl Prospect, OR | 42.72929607 | -122.5158717 | Rogue |
| **14400000** | Chetco R nr Brookings, OR | 42.12455317 | -124.1870364 | South Coast |
| **10396000** | Donner Und Blitzen R nr Frenchglen, OR | 42.79099436 | -118.8674185 | Malheur Lake |
| **14118500** | W Fk Hood R nr Dee, OR | 45.599375 | -121.6363611 | Hood |

## Procedure

The analysis employed a systematic approach to assess the impacts of missing data on flow duration curves (FDCs). The methodology involved the following steps:

### Baseline FDC Computation

FDCs were first generated using the complete dataset for each of the 51 selected index gages, spanning 30 years (water years 1991 through 2020). These uninterrupted daily streamflow records served as the baseline for comparison, providing an accurate representation of natural flow conditions under complete data availability.

### Simulated Missing Data

To evaluate the effects of missing data, subsets of the baseline dataset were created with controlled data gaps. Specifically:

- From the 30-year daily streamflow record, 15 years were randomly selected, and random data gaps of 10–75% were introduced within those selected years.
- All data from the remaining 15 years were retained in full, yielding an overall retention of 62.5–95% of the 30-year record (i.e., 5–37.5% missing).

This approach replicated real-world conditions where missing data occur due to equipment malfunctions, data transmission failures, or other operational challenges. The assumption is that gaps occur randomly throughout the record; however, in practice, gaps may be autocorrelated (e.g., a failing sensor may cause consecutive missing days) or more likely during conditions that stress equipment (e.g., freezing temperatures). These characteristics are not expected to meaningfully influence the results, as both the resampling method and the overall extent of data reduction help minimize the influence of any specific gap pattern.

### Iterative Sampling

To ensure robust results and account for variability, the random sampling and gap-introduction process was repeated 100 times for each level of missing data (10%, 20%, 30%, etc.). For each iteration, an FDC was computed, yielding a distribution of 100 FDCs for each gap level. This iterative process reduced bias and allowed for a comprehensive evaluation of uncertainty.

### Statistical Analysis

For each level of missing data:

- The mean and standard deviation of the 100 FDCs for each month were calculated to summarize the central tendency and variability.
- The mean FDC ± 1 standard deviations were compared against the baseline FDC for each month. Note that ± 1 standard deviations covers 68.3% of the data when assuming normal distribution.
- Deviations were quantified as percentage differences at key exceedance probabilities: 20%, 50%, and 80% of flow duration. These exceedance values were chosen to represent high, median, and low flows, respectively, which are critical for water availability assessments.

### Gap Tolerance Determination

The gap tolerance level was defined as the maximum percentage of missing data at which deviations between the mean FDC (with gaps) ± 1 standard deviations and the baseline FDC remained within 15% across the evaluated exceedance flows. Generally, a 15% percent bias is considered acceptable for hydrological modelling and 15% change presumptively protective for environmental flows (Moriasi et al, 2007, 2015; Richter et al., 2012). Thus, this threshold reflects an acceptable level of accuracy loss while maintaining the integrity of FDC-based analyses.

### Entropy Analysis

To assess the degree of uniformity in streamflow distributions, Shannon entropy was calculated for each gage and month based on the full (baseline) dataset. Entropy provides a measure of distribution spread—higher entropy indicates a more uniform distribution, while lower values reflect more skewed or peaked distributions. For this analysis, daily streamflow values for each month were binned into 200 equal-width intervals, and the resulting frequency distributions were normalized into probabilities before computing entropy. These entropy values were then used to help interpret the sensitivity of FDCs to missing data, as gages with more uniform flow distributions tended to exhibit higher tolerance to data gaps.

By systematically introducing and evaluating the effects of missing data, this method enabled the identification of robust and defensible gap tolerance levels. The findings provide a clear framework for managing incomplete streamflow records while ensuring reliable FDC computations for water availability and streamflow analysis.

## Results

Table 2 summarizes the minimum data completeness required within the 15 randomly selected years to maintain deviations of ≤15% at the 20%, 50%, and 80% exceedance flow levels.  On average, maintaining at least 60% data completeness within these 15 years (equivalent to 20% missing data across the full 30-year period) was sufficient for most index gages to minimize bias in the resulting FDCs relative to those computed from complete datasets. This finding is likely due to the relatively uniform distribution of streamflow observed at many index gages (Figure 1), which makes the FDC less sensitive to random patterns of missing data and reduces the likelihood of significant deviations from the baseline.

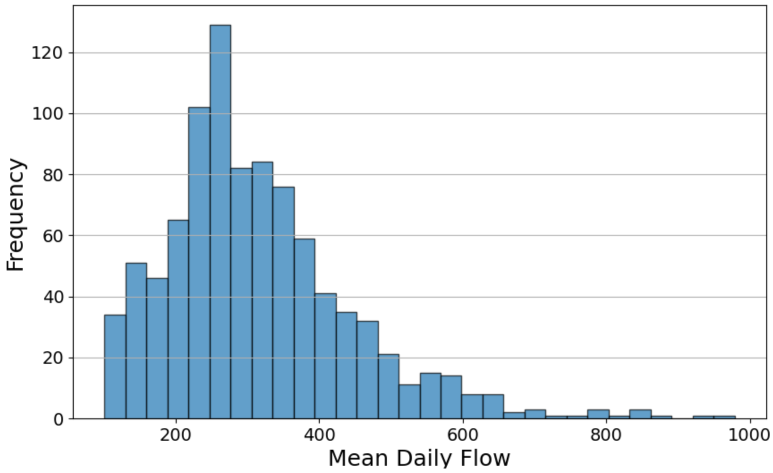


Figure 1. Distribution of mean daily flow data for May from 1991 to 2020 at the gaging station 14010000.

Table 2. Minimum percentage of data needed (within 15 randomly selected years) to represent 20%, 50%, and 80% exceedance flows within a 15% deviation of the full 30-year dataset. The table also includes the corresponding monthly entropy values, which reflect the uniformity of streamflow distributions—with higher entropy indicating more uniform flow patterns across months.

| Station | % data | Jan | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **13216500** | 50 | 3.37 | 3.39 | 3.87 | 4.32 | 4.23 | 4.27 | 3.97 | 3.63 | 3.44 | 3.66 | 3.76 | 3.24 |
| **13331500** | 35 | 3.29 | 3.75 | 4.07 | 4.22 | 4.66 | 4.64 | 4.27 | 4.19 | 3.86 | 3.75 | 3.62 | 4.03 |
| **14010000** | <25 | 3.77 | 2.69 | 4.08 | 4.33 | 4.51 | 4.29 | 4.11 | 3.63 | 3.51 | 3.49 | 3.84 | 3.85 |
| **14010800** | 55 | 3.26 | 2.30 | 4.16 | 3.92 | 4.27 | 3.57 | 3.91 | 3.99 | 3.87 | 2.72 | 3.53 | 3.36 |
| **14032000** | 65 | 2.96 | 3.42 | 3.81 | 3.86 | 3.83 | 2.97 | 3.96 | 3.56 | 4.13 | 4.45 | 4.20 | 2.64 |
| **14020000** | 40 | 3.60 | 2.65 | 3.98 | 4.04 | 4.30 | 3.53 | 4.27 | 3.99 | 3.63 | 2.74 | 3.53 | 3.70 |
| **14044000** | 45 | 3.08 | 4.04 | 4.28 | 4.26 | 4.08 | 3.99 | 3.58 | 4.37 | 3.82 | 4.28 | 4.24 | 3.00 |
| **14091500** | <25 | 4.08 | 4.06 | 4.22 | 4.11 | 4.23 | 4.17 | 3.87 | 3.57 | 3.36 | 3.50 | 3.72 | 4.00 |
| **14074900** | 30 | 3.44 | 3.60 | 3.54 | 3.83 | 4.02 | 3.68 | 3.89 | 4.19 | 4.10 | 3.84 | 3.60 | 3.75 |
| **14057500** | <25 | 3.69 | 3.71 | 3.72 | 3.54 | 3.44 | 3.35 | 3.38 | 3.48 | 3.48 | 3.66 | 3.71 | 3.71 |
| **14055600** | <25 | 4.13 | 4.31 | 4.59 | 4.50 | 4.72 | 4.80 | 4.56 | 4.27 | 4.17 | 4.32 | 4.38 | 4.21 |
| **14054500** | <25 | 3.21 | 3.18 | 3.15 | 3.17 | 3.07 | 3.25 | 3.36 | 3.38 | 3.30 | 3.31 | 3.39 | 3.27 |
| **14050000** | 55 | 4.46 | 4.10 | 4.05 | 4.23 | 4.37 | 4.46 | 4.72 | 4.65 | 4.84 | 4.84 | 4.69 | 4.48 |
| **14052500** | 50 | 3.35 | 3.31 | 3.34 | 3.36 | 3.60 | 3.54 | 3.62 | 3.70 | 3.56 | 3.39 | 3.26 | 3.33 |
| **14050500** | <25 | 3.43 | 3.37 | 3.40 | 3.67 | 4.01 | 4.02 | 3.83 | 3.72 | 3.94 | 3.90 | 3.79 | 3.69 |
| **14051000** | 90 | 3.67 | 3.82 | 3.82 | 4.10 | 4.58 | 4.72 | 4.12 | 4.07 | 3.04 | 2.55 | 2.76 | 2.67 |
| **14052000** | 90 | 3.73 | 3.82 | 3.97 | 4.03 | 4.10 | 4.09 | 3.78 | 3.49 | 3.24 | 3.17 | 3.06 | 2.88 |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **14097100** | 40 | 3.48 | 3.89 | 4.08 | 3.59 | 4.47 | 4.48 | 4.74 | 4.37 | 3.79 | 3.80 | 3.79 | 2.90 |
| **14096850** | 50 | 3.31 | 3.83 | 3.96 | 3.28 | 3.86 | 4.26 | 4.16 | 3.70 | 3.47 | 3.59 | 4.01 | 2.16 |
| **14092750** | <25 | 3.46 | 3.74 | 3.48 | 3.51 | 4.23 | 4.25 | 4.53 | 4.67 | 2.51 | 2.71 | 3.24 | 3.24 |
| **14137000** | 35 | 3.45 | 4.05 | 4.29 | 3.83 | 4.40 | 4.21 | 4.19 | 3.96 | 2.96 | 2.78 | 3.61 | 3.86 |
| **14139800** | 45 | 3.38 | 3.88 | 4.09 | 3.97 | 4.35 | 3.99 | 4.28 | 2.75 | 2.40 | 3.10 | 3.42 | 3.97 |
| **14138800** | 45 | 3.40 | 3.70 | 3.99 | 3.82 | 4.16 | 3.78 | 3.15 | 1.95 | 1.67 | 2.92 | 3.48 | 3.66 |
| **14153800** | 55 | 3.78 | 3.79 | 3.94 | 2.82 | 4.15 | 3.18 | 3.55 | 2.78 | 1.38 | 2.63 | 3.16 | 3.40 |
| **14154500** | 55 | 3.78 | 3.88 | 4.02 | 3.29 | 4.13 | 2.98 | 3.29 | 3.37 | 1.30 | 2.57 | 3.22 | 3.71 |
| **14158500** | <25 | 4.24 | 4.57 | 4.31 | 4.40 | 4.62 | 4.60 | 4.57 | 4.94 | 4.87 | 4.37 | 4.27 | 4.40 |
| **14161500** | 65 | 3.42 | 3.47 | 3.84 | 3.42 | 4.19 | 3.47 | 4.25 | 4.13 | 1.87 | 2.40 | 3.84 | 3.70 |
| **14178000** | <25 | 3.85 | 3.94 | 4.17 | 3.52 | 4.26 | 4.17 | 4.28 | 4.75 | 2.82 | 3.07 | 3.98 | 3.85 |
| **14182500** | 55 | 3.66 | 3.65 | 3.96 | 3.72 | 4.31 | 3.68 | 3.95 | 2.00 | 1.65 | 2.61 | 3.44 | 3.86 |
| **14187000** | 60 | 3.86 | 4.08 | 4.22 | 3.52 | 4.00 | 3.16 | 4.24 | 3.82 | 1.24 | 2.50 | 3.69 | 4.20 |
| **14185000** | 65 | 3.51 | 3.86 | 4.24 | 3.59 | 4.25 | 3.52 | 4.08 | 3.30 | 1.50 | 2.65 | 3.79 | 3.79 |
| **14185900** | 65 | 3.91 | 3.56 | 4.07 | 3.69 | 4.13 | 3.55 | 4.07 | 2.43 | 1.15 | 2.47 | 3.87 | 4.08 |
| **14210000** | 30 | 3.91 | 4.12 | 4.25 | 3.72 | 4.39 | 4.12 | 4.03 | 4.21 | 2.64 | 2.80 | 3.65 | 4.00 |
| **14209500** | <25 | 3.91 | 4.12 | 4.14 | 3.72 | 4.33 | 4.11 | 4.02 | 4.35 | 2.55 | 2.80 | 3.64 | 4.06 |
| **14301000** | 45 | 4.36 | 4.07 | 4.28 | 4.18 | 4.04 | 4.14 | 4.56 | 3.25 | 1.47 | 2.75 | 3.90 | 4.13 |
| **14303600** | 60 | 4.06 | 4.15 | 4.25 | 4.39 | 4.47 | 4.12 | 4.71 | 2.98 | 1.86 | 2.80 | 3.77 | 4.01 |
| **14301500** | 45 | 4.02 | 3.79 | 4.07 | 3.96 | 4.17 | 3.84 | 4.67 | 2.65 | 1.08 | 2.64 | 3.54 | 3.75 |
| **14301300** | 75 | 3.57 | 3.89 | 3.92 | 4.27 | 4.15 | 3.02 | 4.31 | 2.70 | 2.07 | 3.28 | 3.33 | 3.96 |
| **14306030** | 60 | 3.55 | 3.86 | 4.21 | 4.00 | 3.91 | 3.50 | 4.20 | 3.98 | 2.07 | 2.71 | 3.41 | 3.82 |
| **14305500** | 60 | 3.71 | 4.09 | 4.15 | 4.32 | 4.20 | 3.81 | 4.61 | 2.95 | 1.44 | 2.63 | 3.57 | 4.10 |
| **14306500** | 65 | 3.82 | 4.11 | 4.11 | 3.81 | 4.22 | 3.76 | 4.54 | 4.03 | 1.20 | 2.69 | 3.59 | 4.27 |
| **14306340** | 60 | 3.40 | 3.45 | 3.61 | 3.42 | 3.89 | 3.33 | 4.51 | 4.04 | 0.58 | 2.35 | 3.47 | 3.93 |
| **14316700** | 60 | 3.94 | 3.98 | 4.27 | 3.32 | 4.30 | 2.99 | 3.77 | 3.91 | 1.25 | 2.15 | 3.50 | 3.65 |
| **14309500** | 70 | 3.87 | 3.18 | 3.65 | 3.15 | 3.63 | 2.72 | 4.62 | 4.62 | 1.36 | 1.33 | 2.72 | 3.60 |
| **14308000** | 65 | 3.66 | 3.76 | 4.22 | 3.30 | 3.90 | 3.24 | 4.12 | 4.45 | 1.58 | 2.58 | 3.43 | 3.32 |
| **14325000** | 70 | 3.75 | 3.55 | 3.70 | 3.40 | 3.55 | 2.57 | 4.51 | 4.84 | 0.84 | 1.75 | 3.19 | 3.97 |
| **14332000** | 70 | 2.76 | 2.81 | 3.83 | 3.97 | 4.38 | 3.81 | 3.46 | 3.78 | 2.63 | 3.59 | 2.82 | 2.21 |
| **14330000** | <25 | 3.97 | 4.04 | 4.40 | 3.95 | 4.51 | 4.63 | 4.55 | 4.59 | 4.11 | 3.91 | 4.25 | 3.54 |
| **14400000** | 90 | 3.93 | 3.81 | 3.96 | 3.21 | 3.59 | 2.39 | 4.61 | 4.51 | 0.79 | 2.01 | 3.30 | 4.11 |
| **10396000** | 35 | 2.52 | 3.53 | 3.98 | 3.97 | 3.85 | 4.00 | 3.65 | 3.93 | 4.35 | 4.18 | 3.82 | 2.38 |
| **14118500** | 30 | 3.78 | 4.24 | 3.99 | 3.97 | 4.39 | 4.35 | 4.04 | 3.95 | 2.89 | 2.84 | 3.56 | 3.90 |

However, certain index gages exhibited lower gap tolerance levels, requiring a minimum of 65-90% data completeness to achieve deviations within the acceptable 15% threshold. The reduced tolerance for these gages is primarily due to their highly skewed streamflow distributions (Figure 2). In such cases, missing data disproportionately affect the FDCs, amplifying deviations and leading to significant biases, particularly at exceedance probabilities dominated by extreme low or high flows.

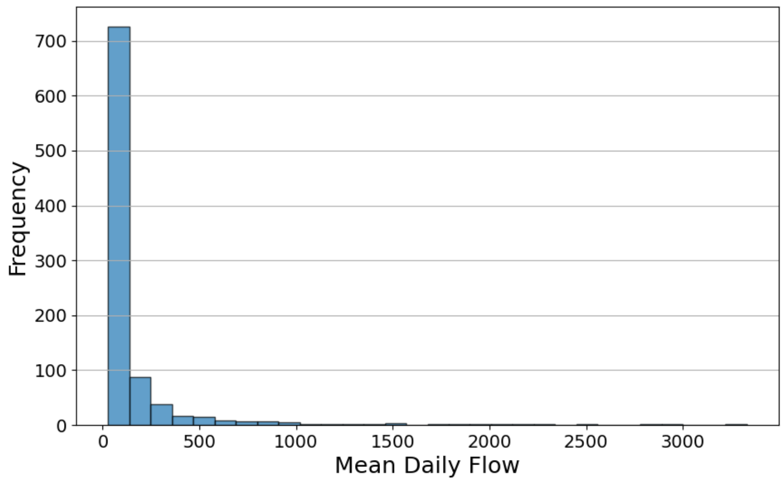


Figure 2. Distribution of mean daily flow data for October from 1991 to 2020 at the gaging station 14325000.

## Discussion

These findings highlight the need to account for streamflow distribution characteristics when determining appropriate gap tolerance levels. Gages with irregular or highly skewed flow distributions—reflected by lower entropy values in Table 2—are inherently more sensitive to missing data, requiring higher levels of data completeness to maintain FDC deviations within acceptable thresholds. In such cases, missing values disproportionately affect key portions of the FDC, especially at exceedance probabilities dominated by extreme low or high flows, amplifying bias. Conversely, gages with higher entropy values, indicating more uniform streamflow distributions, tend to exhibit greater tolerance to data gaps and produce more stable FDCs under conditions of incomplete data. Additionally, the analysis was only applied to minimally disturbed gages for the 1991-2020 base period, and results may not be applicable to disturbed gages or other time periods.

Nevertheless, the analysis confirms that for the majority of index gages, FDCs are relatively robust to missing data, provided that at least 80% of the data is retained across the full 30-year period. In other words, generally, 30-year FDCs at minimally disturbed stations are generally within 15% at key percentiles when at least 80% of daily values are retained — for example, 15 years intact and ≤40% gaps in the other 15 years. Beyond this threshold, deviations from the baseline FDCs become substantial, potentially compromising the reliability of water availability assessments and other hydrologic analyses.

These results show that stations with less-than-perfect data completeness for a period of interest can be useful for characterizing longer-term FDCs, underscoring the importance of setting defensible gap tolerance levels to balance data availability constraints with the need for accurate and reliable FDC computations.

## Conclusion

This study shows that flow duration curves (FDCs) at minimally disturbed Oregon gages are generally robust to incomplete daily records so long as ≥80% of the full 30-year period (WY 1991–2020) is retained. On average, that standard (equivalent to ~60% completeness within the 15 randomly sampled years and all data retained in the other 15) keeps deviations at key percentiles within the 15% threshold. Sites with more skewed, low-entropy flow regimes require stricter completeness—typically ≥82.5–95% overall retention—to meet the same criterion.

We initially assumed each index gage must have complete records for every year of the 30-year base period. These results indicate that the Department can expand the candidate pool to include relatively minimally disturbed gages with less than 30 years of complete record, provided their overall retained data meet or exceed 80% (and higher where flow distributions are highly skewed).

Practically, we recommend a two-tier standard:

1. Baseline acceptance: ≥80% overall retention with no systematic seasonal gaps.
2. Sensitive/low-entropy sites: ≥85–90% (up to 95%) overall retention.

These findings enable more flexible, defensible use of available records in water-availability evaluations while maintaining FDC integrity. They also highlight where caution is warranted—gages with highly skewed flows.

# References

Andrews, R.M., and Huang, C. 2024. WARS Update: Streamgage Network Expansion Documentation. Oregon Water Resources Department.

Falcone, J.A., Carlisle, D.M., and Weber, L.C. 2010. Quantifying human disturbance in watersheds: variable selection and performance of a GIS-based disturbance index for predicting the biological condition of perennial streams. Ecological Indicators 10, no. 2: 264-273.

Falcone, J.A. 2011. GAGES-II: Geospatial attributes of gages for evaluating streamflow. US Geological Survey.

Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D., and Veith, T.L. 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Transactions of the ASABE 50.3: 885-900.

Moriasi, D.N., Gitau, M.W., Pai, N., and Daggupati, P. 2015. Hydrologic and water quality models: Performance measures and evaluation criteria. Transactions of the ASABE 58.6: 1763-1785.

Richter, B.D., Davis, M.M., Apse, C., and Konrad, C. 2012. A presumptive standard for environmental flow protection. River Research and Applications 28.8: 1312-1321.