

Large Diverter Measurement Files Evaluation (2018-2022)

Telemetry Research Unit, Division of Water Rights, State Water Resources Control Board

May 2025

Executive Summary

The State Water Resources Control Board's Division of Water Rights launched a [Telemetry Pilot Project](#)¹ in 2023 to evaluate data collection and reporting processes and explore telemetered water monitoring practices. For water monitoring data, telemetry is the process of automated collection by sensors and automated transmission using communication systems like cellular towers and satellites. Telemetry enables real-time, remote water data collection that improves efficiency and accuracy by reducing field visits and reporting errors. A typical telemetry system includes a sensor, data logger, antenna, and power source.

As part of the Telemetry Pilot Project, the Division evaluated water diversion data files submitted by a large diverter between 2018 and 2022 in compliance with the Board's [Water Measurement and Reporting Regulation](#)² (also called Senate Bill 88 or SB 88 reporting requirements).³ These files are not telemetered and current regulation does not require telemetry by any reporters. The purpose of the evaluation was to understand trends in equipment selection, how reporters format their data files, and if those files are suitable for automated data analysis.

This reporter was selected as a representative large company that encompasses diverse geographic regions of California and retains multiple water rights. Evaluating the reporter's processes for formatting and reporting data can be used to identify trends and challenges common to large reporters statewide.

¹ <https://www.waterboards.ca.gov/telemetry>

² https://www.waterboards.ca.gov/waterrights/water_issues/programs/diversion_use/water_measurement.html

³ California Code of Regulations 23:931-938

The Large Diverter Measurement Files Evaluation (2018-2022) is one of several reports developed by the Division on this topic. The results of these reports form a baseline understanding of data collection and formatting practices and identify potential barriers to compliance. Insights from the reports will inform future projects aimed at reducing reporting burdens for water users and increasing real-time water availability data for improved water management.

Between 2018 and 2022, the large diverter submitted a total of 523 files for 86 different water rights across 12 different watershed basins in California. The Division's evaluation of these files found that the large diverter submitted different types of water data using three unique parameters (flow, storage, and an unspecified parameter called "mean") and two unique units (cubic-feet per second and acre-feet). Parameter names and unit names were mostly consistent across files. However, 13% of files were missing a parameter name or unit name and 2% were missing both.

Although all data were collected and submitted by a single entity, there was considerable variation among the monitoring equipment. Equipment from 18 different manufacturers was reported, with some watersheds containing equipment from several different manufacturers. Some water rights reported the same measurement device from year to year but also reported different details about the device each year, suggesting reporter confusion or error.

While all the files were in tabular, digital form, variations in data content and format limit data usability, as the data require manual staff processing before any automated analysis. These barriers limit the Division's ability to accurately assess water availability in a timely manner, highlighting that the Water Board may need to establish a standard reporting format and improve the water rights reporting system necessary to enhance data quality and improve the reporting experience.

This technical report details the findings of the Large Diverter Measurement Files Evaluation (2018-2022).

Table of Contents

1. Reporting Characteristics	4
2. Equipment Manufacturer Characteristics	4
2.1 Variation of Equipment Manufacturer	4
2.2 DeviceID and Equipment Manufacturer	5
2.3 Water Right and Equipment Manufacturer	6
2.4 Watershed and Equipment Manufacturer	6
3. Data Format Characteristics	8
3.1 Date Characteristics	8
3.2 Parameter Characteristics	9
3.3 Unit Characteristics	10
3.4 Parameter and Unit Relationship	11
3.5 Parameter and Equipment Manufacturer	12

1. Reporting Characteristics

A total of 523 files were submitted for 86 unique water rights (i.e., ApplIDs) from 2018-2022. An ApplID stands for Application Identifier and is the unique identifier provided when an application to divert water is approved. During the 5-year period from 2018 – 2022, more than 5 files were submitted (i.e., more than one per year) for 17 different ApplIDs (Table 1). This may be the result of the water right holder reporting data for multiple points of diversion, multiple water rights being served by a single point of diversion, or due to switching monitoring equipment during the reporting year. Regulations for SB88 require one data file per point of diversion and per monitoring device.

Table 1. Table depicting the 17 application identifiers (i.e., ApplIDs) that submitted more than 5 files in the 5-year review period (2018-2022).

Count of ApplIDs	Count of Files Submitted
1	7
9	10
2	12
3	15
2	16

2. Equipment Manufacturer Characteristics

2.1 Variation of Equipment Manufacturer

The 523 submitted files contain water data recorded by devices from 18 different manufacturers (i.e., Make). There are two categories of Make where the submitter failed to identify the manufacturer (i.e., empty_make) or the submission did not contain enough information to identify the manufacturer (i.e., Unspecified). The most commonly identified manufacturers in the reviewed files were YSI, Hydrological Services, ParoScientific, and Sutron (Figure 1). However, 42% of the submitted files were missing manufacturer information (i.e., empty_make). The files represented in Figure 1 account for approximately 81% of the total file submissions. Overall, the data set shows significant variety in equipment considering that all data is from a single reporter and demonstrates need for improvements to data validations in the reporting system to enhance data quality.

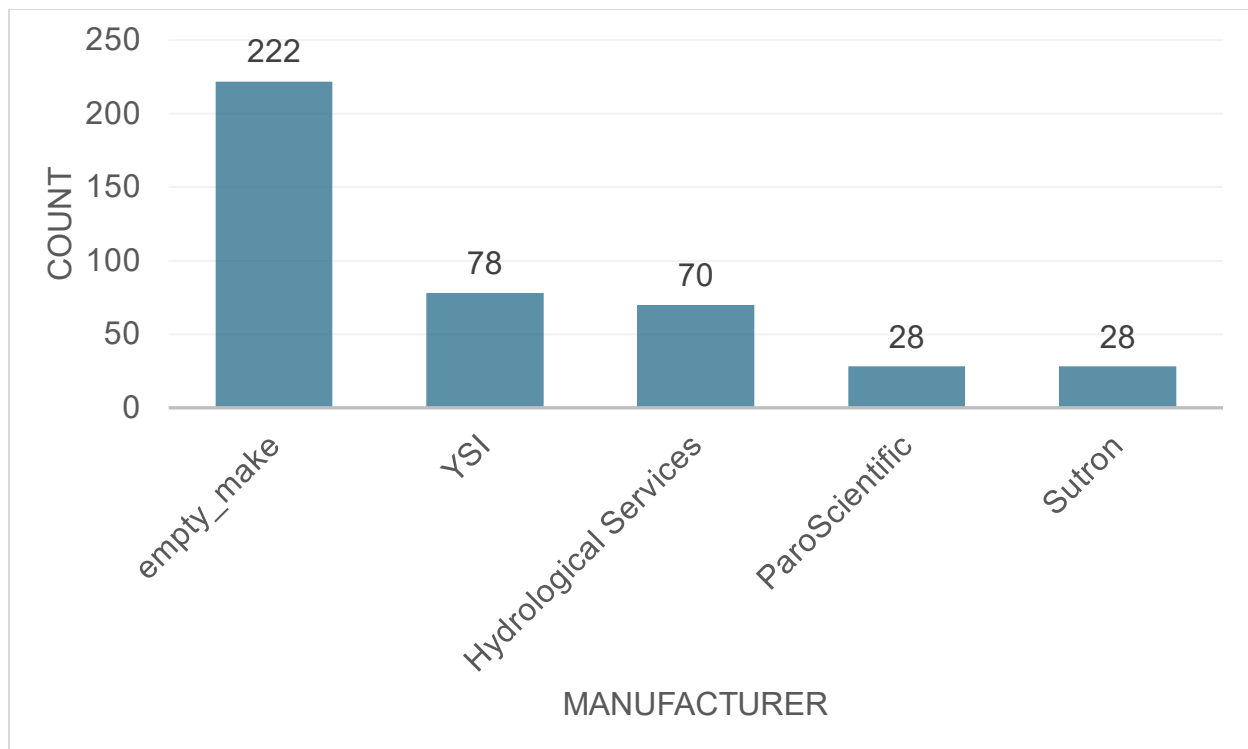


Figure 1. Bar graph displaying selected manufacturers (i.e., Make) on the x-axis, the number of files submitted of that manufacturer on the y-axis, and exact number of files with that manufacturer appears above each bar for the period 2018-2022.

2.2 DeviceID and Equipment Manufacturer

When a reporter registers a monitoring device with the Division of Water Rights, it is assigned a DeviceID. A total of 79 DeviceIDs were reported across all 86 AppIDs and 523 submitted files. In some cases, the same DeviceID was used for multiple AppIDs. SB88 regulations allow for a single device to serve multiple water rights.

When comparing DeviceID and equipment manufacturer, there were 53 AppIDs that reported 43 DeviceIDs with an identified manufacturer in one year and an unidentified manufacturer in a different year during the 5-year review period. In most cases, the unidentified manufacturer appeared for only one year, while the same identified manufacturer appeared for the remaining 4 years. The reason for the same DeviceID to be reported with an identified and unidentified manufacturer is unknown, but it can be the result of human error, a change in personnel submitting or creating the file, or a problem with the user interface when reporters are submitting their annual reports. The State Water Board is building a new geospatial data reporting system called CalWATRS to replace the outdated EWRIMS (Electronic Water Rights Information Management System) reporting system. The new CalWATRS system will be more user-friendly and flag errors such as when a reporter does not select the device manufacturer, which will reduce data errors.

2.3 Water Right and Equipment Manufacturer

When comparing ApplID to equipment manufacturer, only 6 ApplIDs (of 86 total) submitted files with more than one identified manufacturer. However, when accounting for unidentified manufacturers, approximately 55% of the ApplIDs submitted files with at least one equipment manufacturer that was identified and one that was unidentified. These findings suggest that data quality could be improved with an improved reporting system that enforced required fields and data validations.

2.4 Watershed and Equipment Manufacturer

This data set represents 12 California watershed basins, defined using their Hydrologic Unit Code 8 (HUC8). Comparing HUC8 with manufacturer (i.e., Make) showed a variety of manufacturers were used, even in a single watershed with monitoring performed in-house by a single reporter. Of the 12 HUC8s, 8 contained equipment from only 1 or 2 identified manufacturers, while the remaining 4 contained equipment from 3 or more identified manufacturers. Figure 2 displays the HUC8s with the most variation in manufacturer. One of the 12 HUC8s contained equipment from 7 different identified manufacturers.

About 40% of submitted files did not contain details about the manufacturers; these empty_make files occurred in all HUC8s. However, two HUC8s (18020121 and 18040012) accounted for about 56% of the instances.

These findings suggest that even for a particular water reporter in a particular area, a variety of equipment may be in use. A standardized data format for reporting could support data consistency as a reporter operates a variety of equipment or switches equipment and manufacturers. The current lack of standardization decreases data usability, increases staff labor, and delays water management actions.

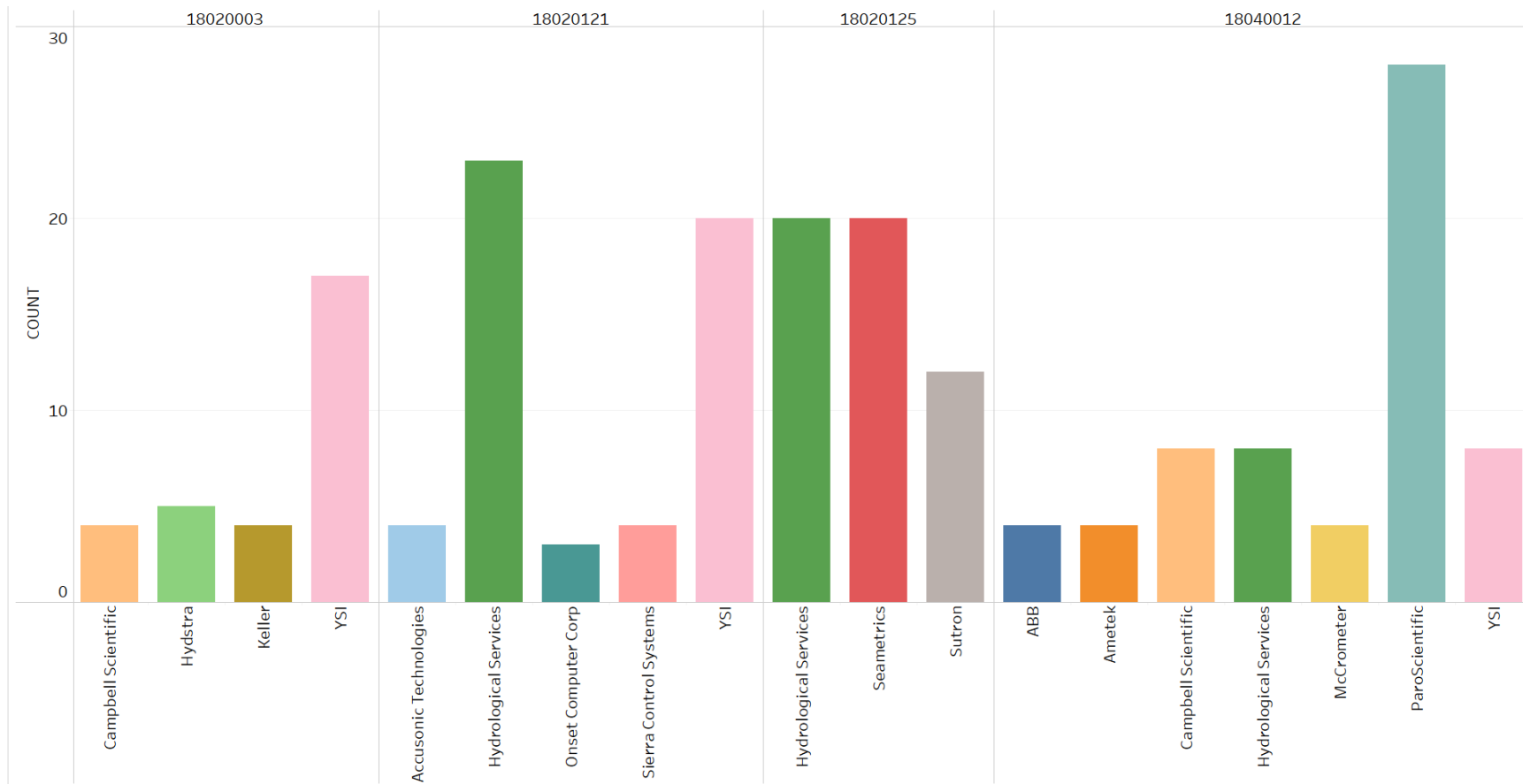


Figure 2. Bar chart displaying watershed (HUC8) on the upper x-axis compared to equipment manufacturer (i.e., Make) on the lower x-axis and total number of files from each manufacturer within the HUC8 on the y-axis for period 2018-2022. Figure only displays these 4 HUC8 areas with the most variation in equipment manufacturer. HUC18020003, Lower Pit. HUC18020121, North Fork River. HUC18020125, Upper Yuba. HUC18040012, Upper Mokelumne.

3. Data Format Characteristics

All submissions were *xlsx* file type and monitoring data was entered directly into the cells, showing that the data has potential to be machine-readable. Machine-readable files can be more easily collated and summarized for business applications. In these ways, this data set is more machine-readable than the data sets examined in “Clear Lake Watershed Measurement Files Evaluation (2018-2022)” and “Reporting Year 2021 Measurement Files Evaluation”, since those data sets contained embedded images of data or files with no discernible telemetry data.

Overall, the 523 files followed a similar (not identical) format regardless of manufacturer. Most files (*n*=513) contained a header row that contained named columns clearly identifying the data. Only 10 files contained metadata, such as station name, in one or multiple rows above the header row and data. The data appeared in two columns for all files, except in 5 files which contained three columns. In all files the first column contained date data (Section 3.1) and the second column contained a combination of parameter data (Section 3.2) and unit data (Section 3.3).

There was no clear trend when comparing format with manufacturer. For instance, a manufacturer that named the first column header “Date” also generated files with other names for date, suggesting that variation is caused either during user configuration or by specific equipment. The reporter’s representative shared in private correspondence that data files for reporting are created by downloading data from their centralized data system after the data has been reviewed by trained staff, confirming that differences in data format are likely caused by user configurations of the data export.

The overall consistency observed among the submitted files is remarkable among files submitted for SB88. As detailed in the companion reports, “Clear Lake Watershed Measurement Files Evaluation (2018-2022)” and “Reporting Year 2021 Measurement Files Evaluation”, the regulations for SB88 do not describe a data standard and submitted files vary widely in content and format. The submitted data was not entirely consistent since 13% of files were missing a parameter or unit name and 2% were missing both parameter and unit names.

3.1 Date Characteristics

The column name for date was recorded in many formats (e.g., [Date], Date, and date), but all files recorded date in *MM/DD/YYYY* format (format excluded the leading 0 for single-digit months and single-digit days). The column name “Date” was by far the most common among the files, appearing 483 times. The name “date” appeared 14 times and “[Date]” appeared 7 times. The remaining 19 files did not provide a label for the date column and were blank. A timestamp recording the time of data collection was not included in any of the files.

3.2 Parameter Characteristics

Across all 523 files, three unique parameters were identified: flow, storage, and mean (Figure 3). Overall, 462 (88%) reported one identified parameter and only 5 files reported two identified parameters, while 56 files had unidentified parameters. The parameter flow was the most reported parameter, appearing 274 times in 269 files.

Files used the same format to name the parameters storage and mean: storage was named “Storage (AF)” and mean was named “Mean”. However, the parameter flow was named “Flow” or “flow” and the unit cubic-feet per second was included in the name as “(cfs)” (e.g., “Flow (cfs)”).

Some files contained data that would not be useful without further context. A total of 79 files did not identify the parameter (n=56) or listed “Mean” (n=23). For the parameter “Mean”, there was not enough information to determine what “Mean” represented. For instance, all files that reported “Mean” lacked units. Of these 79 files, 62 were for reporting year 2018. The remaining 17 were for 2 measurement devices (M001690 and M001724) each associated with two ApplIDs (A030258/A030415 and A002996/A026739, respectively).

These results suggest the large diverter may have increased monitoring in 2018 to meet compliance requirements during initial implementation of SB 88, and then later the large diverter improved its reporting standards for most of its reporting devices. In private correspondence, the reporter’s representative shared that common practice is to provide volumetric (storage) per acre-foot unit in midnight daily storage and flow rate data (direct diversion) in average daily flow per unit of cubic foot per second. The unspecified mean is likely one of these two parameters. The representative affirmed that each year they improve their reporting habits by streamlining and automating their reporting processes, which also reduces their labor costs for reporting compliance.

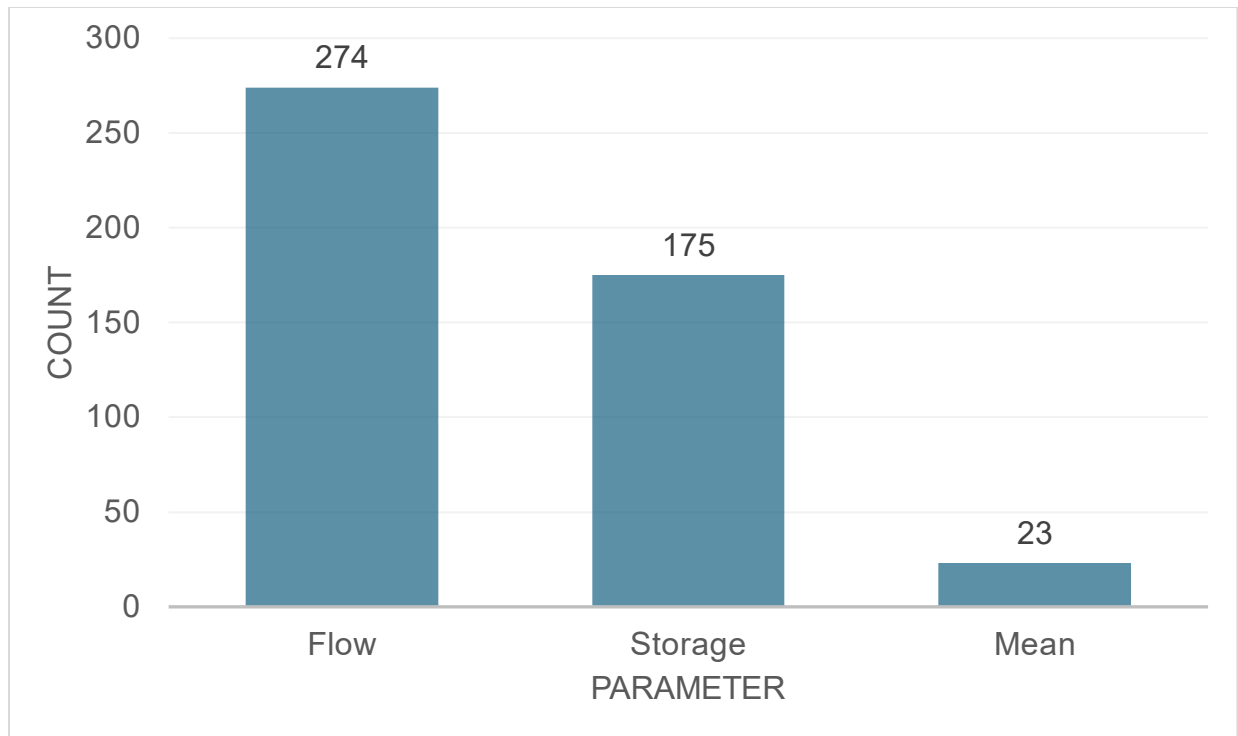


Figure 3. Bar graph showing the number of times parameters appeared in submitted files for the period 2018-2022. Parameter names are on the x-axis, number of files submitted with that parameter are on the y-axis, and exact count of files with that parameter appears above each bar.

3.3 Unit Characteristics

There were two unique units identified in the files: cubic feet per second (CFS) and acre-feet (AF) (Figure 4). There were 312 occurrences (60%) of CFS and 181 occurrences (35%) of AF. The unit CFS was named “(cfs)” or “cfs”. The unit AF was named “AF,” “(AF)”, and “ac-ft”.

When CFS was named “cfs”, it occurred without a defined parameter, but when named as “(cfs)” always appeared in the header name with the associated parameter “Flow” or “flow”. The majority of instances named AF as “(AF)”, with only 8 instances for “AF” and 4 instances for “ac-ft.” The unit “(AF)” appeared in the header name with the associated parameter “Storage,” while the other names of AF appeared without a defined parameter.

As with parameter, the unit was sometimes unspecified although data was provided. Of the 523 files, 23 (4%) contained an identifiable parameter but did not have an identifiable unit and 12 (2%) had no identifiable parameters or units.

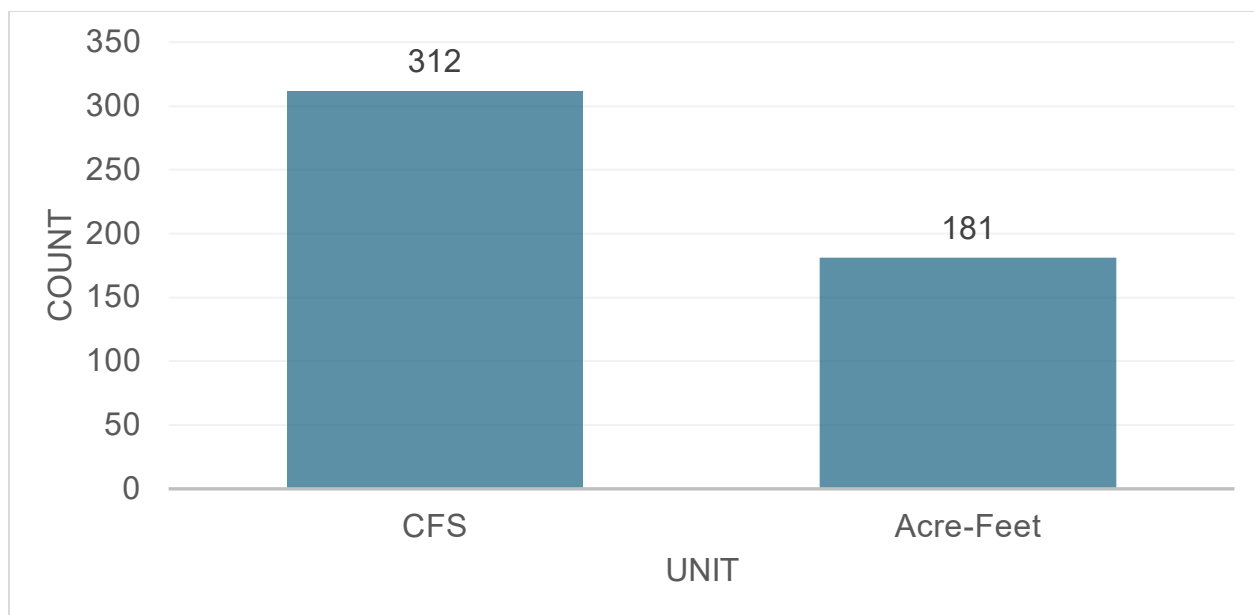


Figure 4. Bar graph showing the number of times units appeared in submitted files for the period 2018-2022. Unit names are on the x-axis, number of files submitted with that unit are on the y-axis, and exact count of files with that unit appears above each bar. CFS - cubic feet per second

3.4 Parameter and Unit Relationship

Examination of unique parameter/unit pairs showed that the parameter “flow” was always assigned the unit of CFS and “storage” was always assigned the unit of AF (Figure 5). About 86% of files recorded an identifiable value for one parameter/unit pair, and about 1% of files recorded an identifiable value for two parameters/unit pairs. All files that recorded two parameter/unit pairs were measurements of flow in units of CFS.

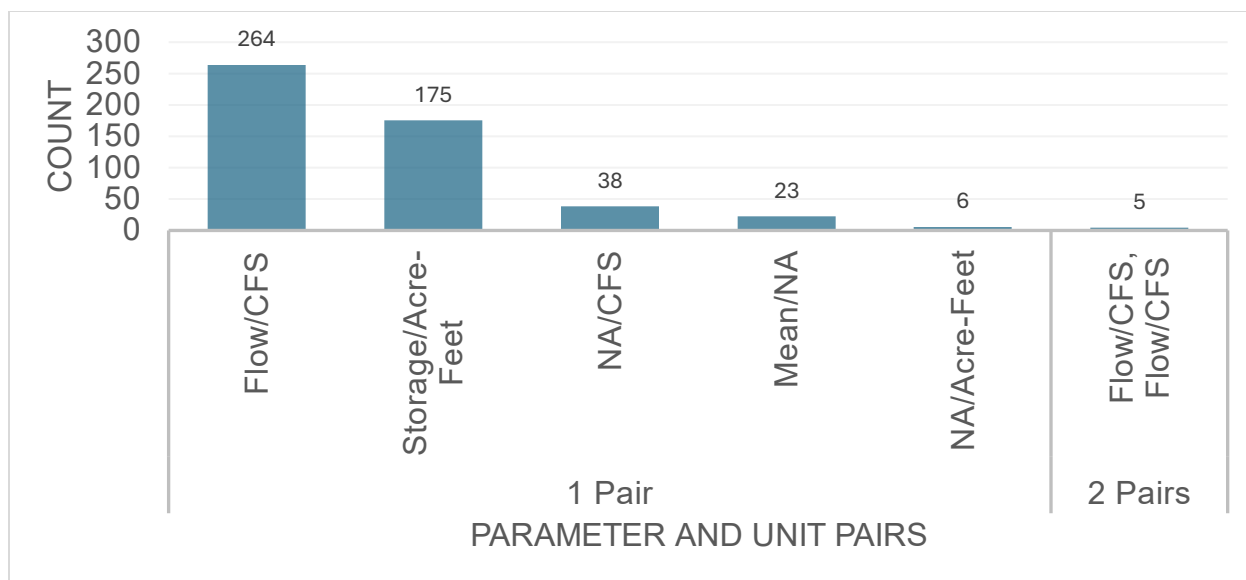


Figure 5. Bar graph depicting the parameter and corresponding unit (parameter/unit pairs) in submitted files during the review period 2018-2022. Appearance of parameter and unit pairs in the file is on the x-axis, the count of files submitted with those pairs is on the y-axis, and exact count of files with that parameter and unit pair appears above each bar.

The remaining 2% of files contained data that was disorganized or unlabeled, with no discernible parameters or units. When averaging across all the files, 90% of files had identifiable parameters and 94% had identifiable units, showing parameter names were missing slightly more often than unit names.

The parameter “mean” appeared in 23 files and always appeared without an identifiable unit. These 23 files contained ten unique AppIDs and all files came from the same HUC8, North Fork Feather River (18020121). The parameter “mean” is ambiguous; it can represent mean flow, mean storage, or a different parameter altogether. The use of ambiguous parameter names is another reason a standardized data format can improve data quality because it will eliminate the use of vague labels.

3.5 Parameter and Equipment Manufacturer

Comparing the 16 identified manufacturers (i.e., Make) and the parameter names in the files demonstrated that only equipment from a few manufacturers was used to report data from more than one unique parameter across all files (Figure 6) and never reported data from more than one unique parameter in the same file (Figure 5). Across the files, only the manufacturer YSI was used to record values for all three parameters (flow, storage, and mean). Three other manufacturers were used to record values related to both flow and storage (namely ABB, Ametek, Seametrics, and YSI), while only Accusonic Technologies recorded values for both flow and mean. The remaining manufacturers only recorded data for either flow or for storage, but not both.

While interesting, it is unclear what implications this finding has for files that may be submitted in the future by this reporter or others. Most of these manufacturers produce equipment that can measure all these parameters and more, so there are several possible reasons for this finding, including 1) water users reporting in compliance with SB88 may be monitoring or reporting the minimum data set that they think will produce compliance, 2) reporters may have complicated operational or historical reasons for operating equipment by different manufacturers or reporting different parameters.

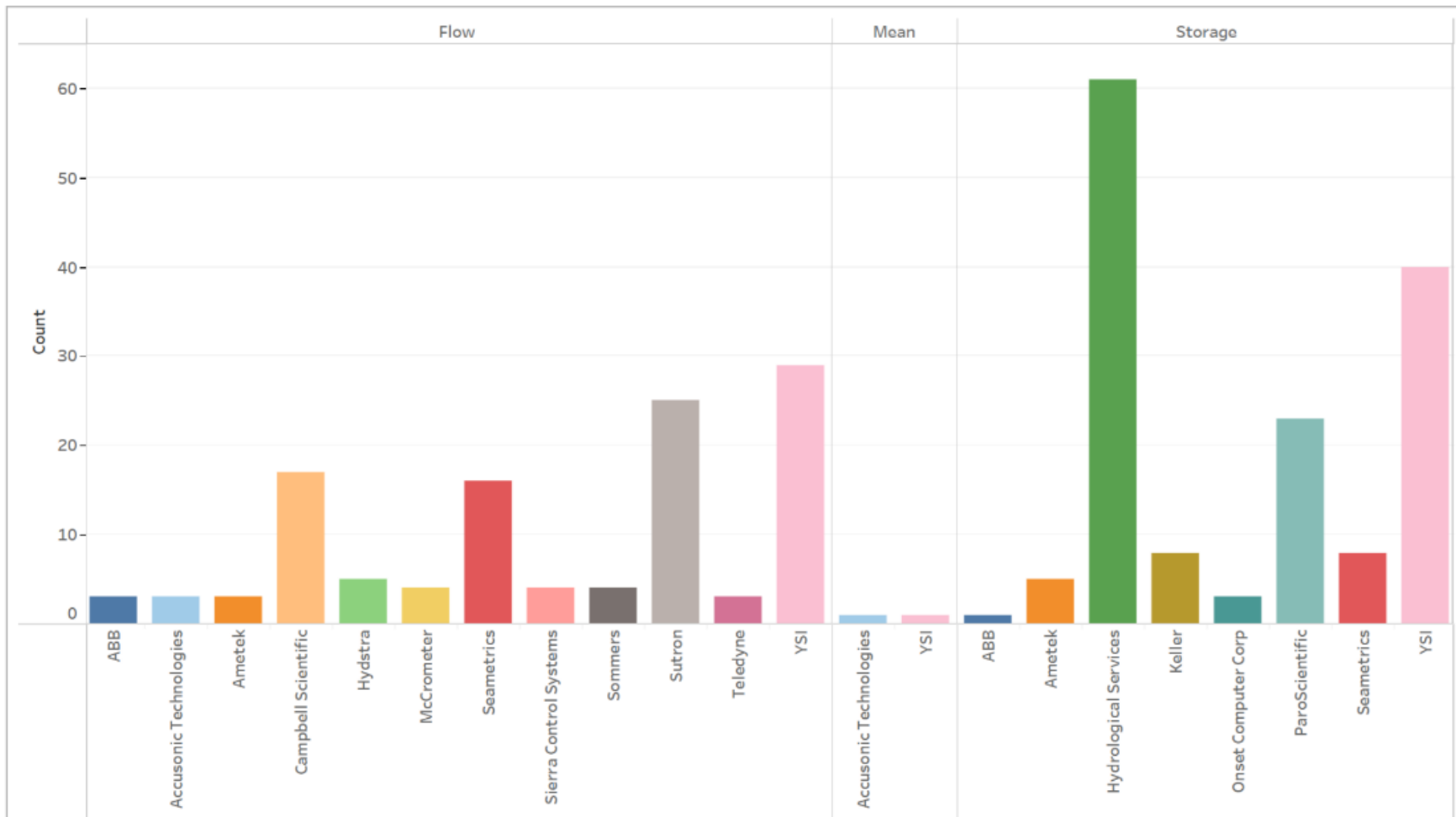


Figure 6. Bar graph showing parameter on the upper x-axis, manufacturer (I.e., Make) on the lower x-axis, and total count of submitted files on the y-axis for period 2018-2022.