

AN ASSESSMENT OF VOLUNTEER-COLLECTED WATER DATA IN MONTANA:
HOW MUCH DATA IS THERE AND WHAT QUESTIONS CAN IT ANSWER?

by

Kenna White Amongero

A professional paper submitted in partial fulfillment
of the requirements for the degree

of

Master of Science

in

Land Resources and Environmental Sciences

MONTANA STATE UNIVERSITY
Bozeman, Montana

May 2026

©COPYRIGHT

by

Kenna White Amongero

2026

All Rights Reserved

ACKNOWLEDGEMENTS

I want to thank my advisor, Dr. Adam Sigler, and Gabrielle “G” Jawer for gifting their time, guiding my curiosity, and providing me with a springboard for my professional paper. Learning from them has been an incredible experience and has made reading countless R-script Stack Exchange posts worthwhile. I also want to thank my partner, Anthony Amongero, for all his help correcting my grammar and for patiently listening to my stream of consciousness as I worked through the research process.

TABLE OF CONTENTS

1. INTRODUCTION	1
2. METHODS	6
Study Site Description	6
Data Collection	9
Volunteer Monitoring Group Identification and SAP Retrieval	9
Water Quality Data Acquisition and Processing	12
Qualitative Coding of Goals and Objectives from Sampling Analysis Plans	14
Objectives Evaluation Framework.....	19
3. RESULTS.....	23
Breadth of Volunteer Monitoring and the Parameters of Interest	23
Qualitative Codes Across Sampling Analysis Plan Goals and Objectives	29
Goal and Objective Evaluation Framework.....	33
Groups with a Baseline Goal Type	33
Groups with a Source Assessment Goal Type	36
Groups with a Trend Goal Type	38
Groups with an Education Goal Type	39
Groups with an Effectiveness Monitoring Goal Type.....	41
Cross Goal Comparisons.....	42
4. DISCUSSION	44
Volunteer Water Quality Monitoring Data Across Montana	45
Insights from the Qualitative Coding of Group Sampling Analysis Plans	48
The Goal and Objectives Evaluation Framework	52
Study Limitations.....	56
Future Research Directions	58
5. CONCLUSION.....	61
REFERENCES CITED.....	63

LIST OF TABLES

Table	Page
1. Table 1. Parameter Type Groups.....	13
2. Table 2. Monitoring Goal Type Coding Scheme.	15
3. Table 3. Monitoring Objectives Coding Scheme.	17
4. Table 4. Objective Context Coding Scheme.	18
5. Table 5. Goal Type and Objective Co-occurrence Matrix	30

LIST OF FIGURES

Figure	Page
1. Figure 1. Montana Regional Watershed Map.	7
2. Figure 2. Volunteer Group Filtering Flowchart.	10
3. Figure 3. Goals and Objectives Hierarchy Flowchart.	19
4. Figure 4. Qualitative Coding and Quantitative Water Quality Data Integration Framework.	20
5. Figure 5. Map of Volunteer Monitoring Sites and Density.	24
6. Figure 6. Monitoring Group Data Collection Timeline.	26
7. Figure 7. Monitoring Group Count by Year.	27
8. Figure 8. Data Collection Quantified by Parameter Type.	28
9. Figure 9. Top Parameter Collection Tree Map.	29
10. Figure 10. Goal and Objective Type Counts.	31
11. Figure 11. Goal and Objective Co-occurrence Network Diagram.	32
12. Figure 12. Baseline Goal Projects Data Collection Boxplots.	35
13. Figure 13. Source Assessment Goal Projects Data Collection Boxplots	37
14. Figure 14. Trend Analysis Goal Projects Data Collection Boxplots.	39
15. Figure 15. Education Goal Projects Data Collection Boxplots.	40
16. Figure 16. Effectiveness Monitoring Goal Projects Data Collection Boxplots	42

ABSTRACT

Citizen science plays an increasingly important role in environmental data collection, and one of the most prominent forms of citizen science in the United States is volunteer water quality monitoring. However, the outcomes of volunteer water quality monitoring projects are rarely compiled comprehensively. In Montana, no synthesis of the extent of volunteer-collected water quality data exists, and little is known about how well volunteer data is poised to address the goals and objectives of the groups collecting it. Therefore, this study develops and applies a mixed-methods framework to evaluate whether volunteer water quality monitoring programs in Montana are collecting data sufficient to achieve their monitoring goals and objectives articulated in monitoring groups' Sampling Analysis Plans. Leveraging connections with volunteer water quality practitioners, I identified 43 independent volunteer monitoring projects across Montana. Project goals and objectives were qualitatively coded using deductive, inductive, and thematic approaches, then merged with project-level National Water Quality Portal datasets by parameter type. Project-goal combinations were evaluated using metrics describing monitoring duration, spatial extent, analyte diversity, and sampling frequency to assess how well dataset structures aligned with their stated monitoring intentions. Results show that volunteer monitoring has expanded statewide since 2001, with sampling concentrated in western Montana. Baseline and trend analysis goals are the most common across projects, while source assessment, effectiveness monitoring, and education goals are less frequent. Most groups sample primarily during summer baseflows, resulting in limited coverage during hydrologically important runoff periods. Projects with long-term datasets or broad spatial and temporal coverage are comparatively well positioned to address their stated goals, while projects seeking to evaluate pollutant sources or restoration effectiveness often lack the temporal resolution required. This study demonstrates that assessing volunteer-collected data against monitoring intent provides a clearer understanding of dataset adequacy and highlights opportunities to strengthen volunteer monitoring designs through expanded seasonal sampling, more intentional articulation of education goals, and integration of additional monitoring metrics.

INTRODUCTION

Citizen science can be broadly defined as the intentional involvement of volunteers in the scientific process, often including individuals with no specific scientific training (Haklay et al. 2021). Citizen science is generating an increasing amount of environmental data (Silvertown 2009), ranging from documenting biodiversity shifts to collecting evidence of climate change (Kress et al. 2018, Jansen et al. 2024). Within this environmental science domain, volunteer water quality monitoring has emerged as a prominent form of citizen science, and participation in these efforts has gained momentum over the past two decades (Stepenuck and Genskow 2018). Citizen-led water quality monitoring efforts enhance community awareness and add value to many organizations by expanding their sampling capabilities (Peters et al. 2015, McCauley 2017). These programs additionally function as both data-generating systems, expanding spatial and temporal monitoring coverage beyond what governments alone can support (Devictor et al. 2010), and as decision-support tools, informing watershed planning, restoration prioritization, and community-based management (Peters et al. 2015).

Volunteer water-monitoring programs have generated large datasets over the past 20 years, measuring parameters across diverse aquatic systems (Stepenuck and Genskow 2018, Shao and Bishop 2025). Throughout the United States, these programs vary widely in scale, from local watershed groups conducting seasonal sampling to long-standing volunteer networks contributing information to centralized databases (Green et al. 2013, Albus et al. 2019, Stepenuck 2023). Up to 1,700 volunteer monitoring programs are estimated to operate in the U.S. (Stepenuck 2023). Many of these local volunteer water monitoring groups fall under a “parent” program, such as Stream Teams, Adopt-A-Lake, Watershed Keepers, and other

volunteer programs that operate across multiple states (Stepenuck 2023). Many volunteer water quality programs are state, local agency, tribal, university, or non-profit affiliated (Nerbonne and Nelson 2004, Green et al. 2013, Stepenuck and Genskow 2018), through which data can be made publicly accessible by uploading it to the Water Quality Exchange (WQX). The WQX is the Environmental Protection Agency's (EPA) data warehouse, which has been integrated into the national Water Quality Portal (WQP, 2021) alongside the United States Geological Survey National Water Information System (NWIS) (National Water Quality Monitoring Council 2026a). Through the online WQP data portal, select volunteer data from across the US has become easily accessible.

While volunteer monitoring data is becoming increasingly integrated into a national database, details on program goals and monitoring objectives remain scattered. The following studies demonstrate the breadth of these goals and objectives. In a 2013 survey of 293 volunteer water quality monitoring programs, Stepenuck and Genskow (2018) reported that 84% of program objectives focused on engaging and educating the community and on creating long-term datasets. A synthesis of 72 peer-reviewed articles on citizen science in surface water quality monitoring revealed that measurements of water clarity, pH, and dissolved oxygen (DO) were the most frequent (Ramírez et al. 2023). Rivers and lakes were the most prevalent water bodies monitored by groups with aims ranging from pollution monitoring to studying restoration outcomes (Stepenuck and Genskow 2018, Ramírez et al. 2023, Shao and Bishop 2025). Rather than investigating the types of questions volunteer programs are most interested in, the scientific literature has largely focused on evaluating the accuracy, credibility, and reliability of volunteer water quality data (Burgess et al. 2017, Albus et al. 2019, 2020). Professional and volunteer data

can be comparable across a range of conditions (Kosmala et al. 2016, Albus et al. 2019); however, concerns about high variability in citizen science have left much data underutilized (Burgess et al. 2017, Albus et al. 2020). While volunteer water quality monitoring falls under a broad purview and invites analytical criticisms, interest in environmental stewardship, declining agency resources, advances in low-cost monitoring technologies, and increased volunteer engagement in formal decision-making processes continue to drive program growth (Silvertown 2009, Conrad and Hilchey 2011, McKinley et al. 2017, Grudens-Schuck and Sirajuddin 2019).

Despite the growing volume of volunteer water quality data, synthesis efforts have largely occurred at high levels or on coarse scales. Volunteer data synthesis has focused on program spatial and demographic coverage (e.g. Stepenuck 2023), participation structures (e.g., Stepenuck and Genskow 2018), monitoring methods (e.g., Ramírez et al. 2023, Shao and Bishop 2025), and data usability and accuracy (e.g., Albus et al. 2019, 2020), rather than the goals and objectives articulated by monitoring groups themselves. Studies evaluating volunteer data have been essential for establishing the credibility of citizen science programs (McCauley 2017, Albus et al. 2020). However, the outcomes of citizen science and its volunteer programs have been rarely compiled and comprehensively evaluated (Stepenuck and Genskow 2018, Shao and Bishop 2025). I have found no literature to date that examines how volunteer water quality monitoring intentions translate into and are leveraged to meet specific data requirements. Establishing data validity alone does not address whether the structure and scope of volunteer datasets are appropriate for answering the questions monitoring groups seek to answer. As a result, the connection between what volunteer groups intend to learn and what their datasets can elucidate remains poorly understood. Significant effort is put into citizen science and volunteer

monitoring; yet, if practitioners lack sufficient insight into successful volunteer monitoring programs, resources may not be allocated efficiently to foster future effective programs (Measham and Barnett 2008). Therefore, it is valuable for citizen-led monitoring groups and supporting entities to evaluate the alignment of group datasets with project goals and objectives.

In Montana, there has been no synthesis of volunteer water quality monitoring projects. Yet Montana provides a favorable setting for analyzing volunteer-collected water quality data, given its strong institutional support for volunteer monitoring programs. Institutional backing and collaboration are consistently identified as key factors in the consistency of volunteer monitoring programs and the usability of their data (San Llorente Capdevila et al. 2020). Throughout Montana, volunteer monitoring efforts are supported by state agencies such as the Montana Department of Environmental Quality (MDEQ, n.d.), university programs such as Montana State University Extension Water Quality (MSUEWQ) (Montana State University [MSU] n.d.), and the Flathead Lake Biological Station's Monitoring Montana Waters (MMW) program (Flathead Lake Bio Station 2025). Monitoring efforts are also supported by statewide coordination networks, including the Montana Watershed Coordination Council (MWCC, 2025), and the Montana Water Center (2017). Both MDEQ and MMW offer funding and extensive guidance on volunteer monitoring to help groups draft comprehensive Sampling and Analysis Plans (SAPs) and implement water quality monitoring programs (Flathead Lake Bio Station 2025, Montana DEQ n.d.). MSUEWQ also helps develop monitoring programs across Montana, provides monitoring education resources and guidance on data analysis, and hosts a water quality data hub specifically for data collected by volunteer monitoring groups (MSU n.d.). Montana volunteer water quality institutions have integrated their educational resources and developed

cohesive planning documents, ensuring that monitoring groups across the state receive consistent guidance. Coordinated institutional networks are important to reducing redundancy and supporting program persistence across changing funding and volunteer dynamics (San Llorente Capdevila et al. 2020).

The institutional support and coordination structures across Montana allow diverse local monitoring groups to operate autonomously while still contributing to a more cohesive statewide monitoring landscape. Montana's water quality volunteers are encouraged to identify and articulate the goals and objectives of a proposed monitoring program before drafting their more detailed and technical SAP. The MDEQ, MMW, and MSUEWQ direct volunteers to the Monitoring Method Selection Guide (MMSG) (Makarowski and Sigler 2020) to help groups state clear goals and write objectives that directly support specific goal types. The MMSG identifies five goal types and provides examples of objectives for each (Makarowski and Sigler 2020). Once goals and objectives are defined, groups are directed to complete a SAP outlining field procedures, analytical requirements, quality assurance/quality control, data management, and a project budget using either the MDEQ or MMW template (MDEQ 2022, MMW 2025). In Montana, the MDEQ data portal (which feeds into the larger WQX) and the MSUEWQ Data Hub are recommended data storage options for all monitoring groups (MSUEWQ 2021). Given the importance of groups defining goals and objectives, the dissemination of uniform MMSG recommendations for volunteer quality monitoring design, and the availability of statewide volunteer data storage resources, Montana provides a unique and structured context in which to evaluate whether volunteer monitoring efforts are aligned with their stated purposes. A systematic evaluation of the alignment between volunteer monitoring data and their goals and

objectives could provide valuable insight to improve monitoring designs and ensure that volunteer efforts are positioned to meaningfully inform water resource management.

Addressing the questions surrounding the extent of volunteer-collected water quality data in Montana and how well that data is poised to address the goals and objectives of the groups collecting it requires a framework that can integrate qualitative and quantitative information. Therefore, the aim of this study is to develop and apply a mixed-methods framework to evaluate whether volunteer water quality monitoring programs in Montana are collecting data sufficient to achieve their monitoring goals and objectives. To meet this aim, my objectives are to characterize the available citizen science and water quality monitoring group data across the state of Montana; to compile and qualitatively analyze the goals and objectives from the volunteer monitoring groups' Sample Analysis Plans; and to establish a framework to compare characteristics of water quality datasets against standards for data requirements to address stated goals and objectives.

METHODS

Study Site Description

Montana is the fourth-largest state by land area in the US and encompasses a range of climatic diversity. Montana has seven National Oceanic and Atmospheric Administration (NOAA) climate divisions (NOAA n.d.) and seven level III ecoregions (Sando et al. 2025). Water resources in Montana depend on winter snowpack and climate greatly influences streamflow patterns because the snowpack varies across regions, depending on elevation, precipitation distribution, and side of the continental divide (Whitlock et al. 2017 chap. 3).

Montana is designated as a “headwaters state,” with 80% of its water coming from within state boundaries (Whitlock et al. 2017 chap. 3). The waters in Montana are categorized into four sub-regions, 16 basins, and 115 sub-basins under hydrologic unit code (HUC) system (MDEQ 2025 p. 5, Figure 1). Roughly 49 billion m³ (40 million acre feet) of water are exported from Montana each year (Whitlock et al. 2017 chap. 3), traveling through 93,127 km (57,867 mi) of perennial rivers and streams (MDEQ 2025). There are 1,417 named ponds, lakes, and reservoirs greater than 20,000 m² (roughly five acres) that dot Montana’s landscape (Montana Lake Book Steering Committee [MLBS] 2017 p. 10). The most noteworthy of which is Flathead Lake, with a capacity of 23.27 km³ and 302 km (187.6 mi) of shoreline (Flathead Lake Bio Station 2026a).

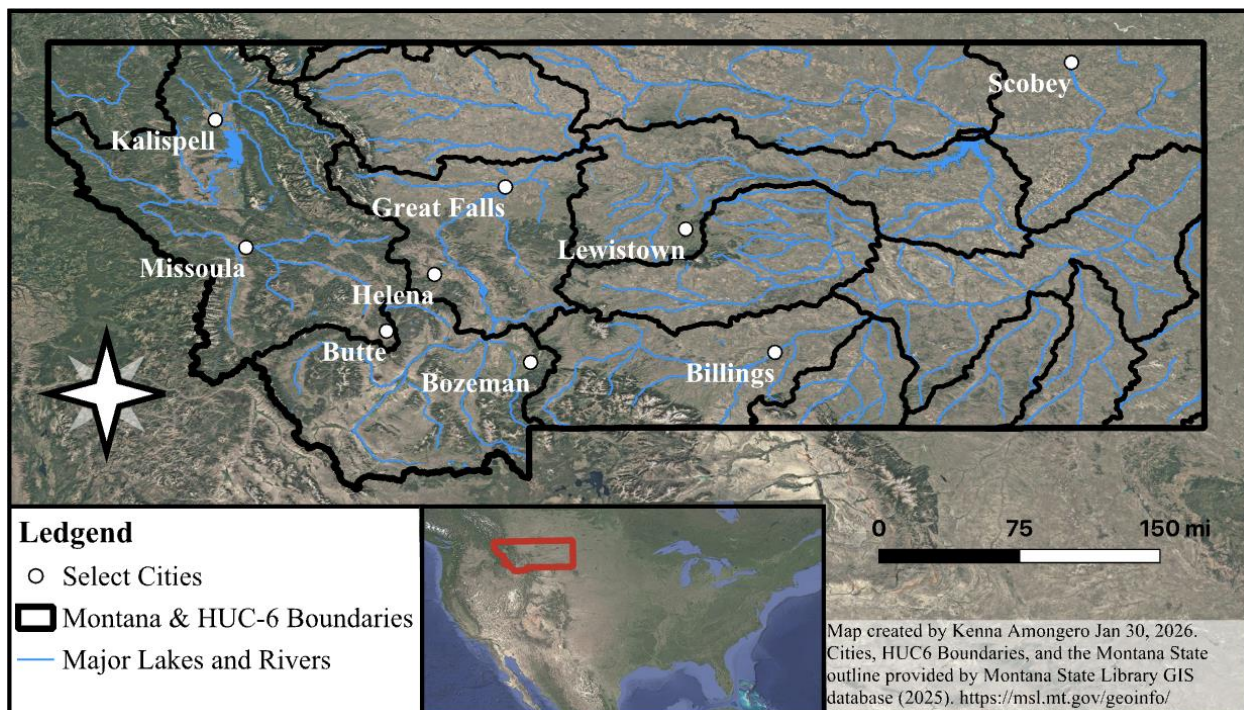


Figure 1. Montana Regional Watershed Map. A map of Montana’s position within the United States with HUC-6 watershed districts projected on the NAD83 coordinate system. Major rivers, lakes, and selected Montana Cities across the state are included.

Most watersheds in Montana are snowmelt-dominated, with stream discharge peaking in the spring and transitioning to low flows during late summer. Spring snow melt is a primary driver of annual runoff, particularly west of the Continental Divide and in higher-elevation watersheds (Pederson et al. 2011). However, the Montana Climate Assessment reports that warming temperatures have reduced snowpack in many regions, leading to earlier peak run-off and longer summer base-flow periods (Whitlock et al. 2017). Rivers that receive water from both low- and high-elevation snowpacks can exhibit a dual-peak hydrograph with peaks in both early spring and summer (Whitlock et al. 2017 chap. 3). Low-elevation rivers in eastern Montana can exhibit irregular runoff patterns, and they often have lower baseflows and greater sensitivity to seasonal precipitation and drought (Pederson et al. 2011, Whitlock et al. 2017 chap. 3). Studies have observed a decline in late-summer stream flows consistent with warm-season drought trends observed in the Columbia and Upper Missouri River basins (Das et al. 2011, Leppi et al. 2012, Norton et al. 2014). Therefore, water quality sampling timing and frequency must be tailored to distinct flow regimes to capture representative conditions, given Montana's hydrologic heterogeneity and drought trends.

Varied land and water uses in Montana influence water management and contribute to spatial variability in water quality. The state comprises forests, rangelands, agricultural valleys, growing urban areas, and regions historically and currently influenced by mining and energy development. Agricultural land use, chiefly irrigated crop production and livestock grazing, has been linked to elevated nutrient, sediment, and pathogen loads in watersheds (Sharpley et al. 1994, Lyons et al. 2000, Berka et al. 2001, MDEQ 2025 p. 33). Urbanization, hydrologic modification, forestry, and mining activities in Montana have resulted in nonpoint-source water

contamination that continues to influence monitoring priorities (MDEQ 2025). As of 2025, MDEQ has listed 44% of river and stream kilometers and 72% of lake square kilometers within state jurisdiction as impaired under 303(d) Category 5 of the Clean Water Act (MDEQ 2025 pp. 32–33). Category 5 listing includes waters with designated uses that are impaired or threatened to be impaired by a pollutant and therefore require a Total Maximum Daily Load (TMDL) plan (MDEQ 2025 p. 72). Excess algal growth in waters is the number one pollutant between 2022 and 2024, causing new 303(d) impairment listings in Montana (MDEQ 2025 p. 37).

Data Collection

Volunteer Monitoring Group Identification and SAP Retrieval

Water quality data compiled by citizen science and monitoring group projects across Montana are available through the WQX (<https://www.waterqualitydata.us/>) and the MSUEWQ data portal (datahub.siglerh2o.org/data). To identify all volunteer monitoring projects across Montana with available data, the names of projects with Montana Volunteer (MTVOLWQM_WQX) and Montana Watershed (MTWTRSHD_WQX) Organization Identifiers (IDs) were downloaded from WQX via the WQP (2021) using the “Project Data” comma-separated data profile (National Water Quality Monitoring Council 2026). Following the screening methodology detailed in Figure 2, I evaluated the 233 water quality monitoring projects by manually searching for keywords in the project description, project name, and organization formal name columns of the WQP data file. I then cross-referenced the groups that passed the keyword filter against the MWCC (2025) Monitoring Resource Library’s volunteer monitoring list to verify list completeness. Tribal, State, and Federal volunteer projects listed on

the MWCC Monitoring Resource Library were excluded to maintain focus on volunteer monitoring efforts.

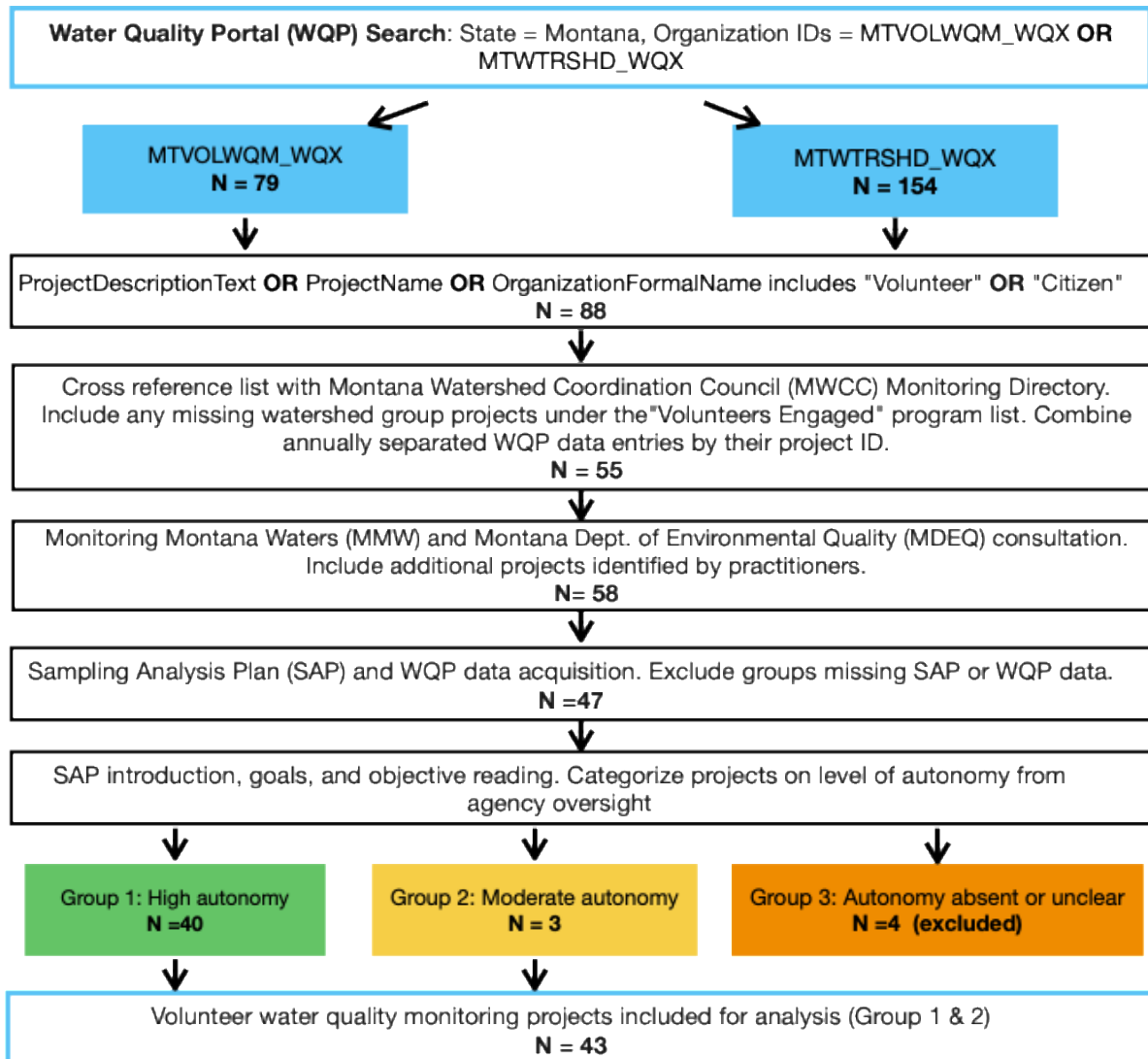


Figure 2. Volunteer Group Filtering Flowchart. Flow chart summarizing how volunteer monitoring groups were identified and filtered during the initial search, showing how I arrived at the final 43 groups. The bolded number at the bottom of the text boxes represents the number of projects remaining after the action described in the text boxes occurred.

Several monitoring projects uploaded data into WQX under year-specific project identifiers (e.g., SQWI-2024, SQWI-2023). For each data profile, I assigned a simplified project

ID to group annually segmented efforts into 55 total water quality monitoring projects. The list of 55 statewide volunteer monitoring projects was circulated to MDEQ, MMW, and MSUEWQ technical service providers, and the consultation identified three missing volunteer groups. I then leveraged connections to the 58 isolated projects from MSUEWQ (Montana State University n.d.), MDEQ (n.d.), and MMW (Flathead Lake Bio Station 2025) to retrieve the projects' SAPs. If a project lacked an SAP or comparable planning document listing project goals and objectives, or if its data could not be found in the WQX, it was excluded from further analysis. After reading the SAPs' introduction, goal, and objectives, I assigned the projects into three groups based on the level of autonomy from agency oversight. Projects with moderate autonomy were those that either lacked SAPs formatted to match MMW and MDEQ templates or had unclear relationships with state and federal entities. Volunteer monitoring projects operating under direct government or agency supervision were excluded, leaving 43 independently operating projects for assessment (Figure 2). A noteworthy volunteer group that was excluded from analysis but identified by MDEQ was the Clark Fork Coalition. The Clark Fork Coalition has robust volunteer monitoring programs, but its water quality planning documents did not provide evidence that volunteers were responsible for setting the monitoring goals and objectives, and a SAP was not readily located to facilitate identification of goals and objectives. While several non-government organizations (NGOs) were involved in more than one of the 43 volunteer projects, each project had a unique combination of authors and involved volunteers. Therefore, herein, the terms "volunteer projects" and "volunteer groups" are synonymous.

Water Quality Data Acquisition and Processing

I compiled the physical and chemical water quality data for the 43 volunteer monitoring groups from the WQX via the WQP using RStudio (v4.1.2, R Core Team 2021) and the *dataRetrieval* package (v2.7.21, Laura DeCicco et al. 2025). After downloading group data by their “ProjectIdentifiers” retrieval element (National Water Quality Monitoring Council 2026 tbl. 5), I reassigned the data to their simplified project ID names to facilitate reproducible querying and data management. Simplified project IDs uniquely identify each volunteer monitoring group, ensuring consistency between water quality records segmented by year and their associated SAPs. Data from the *dataRetrieval* download were compiled in tabular format. I removed unnecessary data columns from the analysis, including administrative and metadata fields, while retaining columns related to project identification, sample date, sampling activities, characteristic name, result values, detection limits, and sample site locations and names. The data acquisition workflow is available in a public GitHub repository (Amongero 2026).

After data acquisition, I renamed retired water quality characteristics using the WQX-suggested substitutes to eliminate ambiguity caused by historical naming differences across projects (WQP 2021). I then classified the measured water quality characteristics into parameter types, as shown in Table 1. Water quality characteristics were first assigned to one of four parameter types: nutrients, sediment, metals, or biological. Remaining characteristics that did not clearly fit within these four initial parameter types were subsequently classified as either chemical or physical parameters. While several water quality characteristics could belong to multiple parameter groups, like nitrate being both a nutrient and chemical parameter, I prioritized classification based on how the parameter is typically interpreted in monitoring contexts and how it was described in the monitoring groups’ SAPs. The final parameter classification in Table 1

was reviewed by Dr. Sigler from MSUEWQ and cross-referenced with MDEQ's internal Excel spreadsheet for nutrient reclassification (Montana DEQ 2007). Each parameter was also designated as laboratory-derived or field-derived based on standard sampling practice and explicit descriptions of sampling and analytical methods provided in SAPs (Makarowski and Sigler 2020).

Table 1. Parameter Type Groups. Classification of the water quality characteristics measured by the identified volunteer monitoring groups into parameter types. Italicized characteristics are field-derived, and all other characteristics are lab-derived.

Parameter Type	Included Water Quality Characteristics
Biological	BacR DNA marker; Chlorophyll a; <i>Chlorophyll a (probe relative fluorescence)</i> ; <i>Chlorophyll a (probe)</i> ; Chlorophyll a, corrected for pheophytin; Chlorophyll a, uncorrected for pheophytin; <i>Macroinvertebrate Count</i> ; DG3 DNA marker; Enterococcus; Escherichia coli; Fecal Coliform; GFD DNA marker; HF183 DNA marker; Total Coliform
Chemical	Alkalinity, total; Bicarbonate; Biochemical oxygen demand, standard conditions; Boron; Bromide; Caffeine; Calcium; Calcium carbonate; Carbonaceous biochemical oxygen demand, standard conditions; Carbonate; Chemical oxygen demand; Chloride; <i>Dissolved oxygen (DO)</i> ; Ferrous ion; Fluoride; Hardness, Ca, Mg; Magnesium; Methane; Organic carbon; Oxidation reduction potential (ORP); Potassium; Salinity; Selenium; Silica; Sodium; Sodium adsorption ratio [(Na)/(sq root of 1/2 Ca + Mg)]; Sulfate; Sulfide; Sum of anions; Sum of cations; Total dissolved solids; <i>pH</i>
Metals	Aluminum; Antimony; Arsenic; Barium; Beryllium; Cadmium; Chromium; Copper; Iron; Lead; Manganese; Mercury; Molybdenum; Nickel; Silver; Thallium; Uranium-238; Zinc
Nutrients	Ammonia; Ammonia-nitrogen; Inorganic phosphorus; Kjeldahl nitrogen; Nitrate; Nitrate + Nitrite; Nitrite; Nutrient-nitrogen (retired); Orthophosphate; Phosphate-phosphorus; Phosphorus; Soluble Reactive Phosphorus (SRP); Total Kjeldahl nitrogen (Organic N & NH ₃); Total Nitrogen, mixed forms; Total Phosphorus, mixed forms
Physical	<i>Apparent color</i> ; <i>Barometric pressure</i> ; <i>Conductivity</i> ; <i>Depth</i> ; <i>Secchi disk depth</i> ; <i>Depth, bottom</i> ; <i>Dissolved oxygen saturation</i> ; <i>Flow</i> ; <i>Height, gage</i> ; <i>Precipitation during activity (choice list)</i> ; <i>Specific conductance</i> ; <i>Stream flow</i> ; <i>Stream width measure</i> ; <i>Temperature, air</i> ; <i>Temperature, water</i> ; <i>Velocity - stream</i> ; <i>Water level reference point elevation</i> ; <i>Weight</i>
Sediment	Particle size (size/range in result Particle Size Basis); <i>RBP Turbidity Code (choice list)</i> ; Suspended Sediment Concentration (SSC); Total suspended solids; Total volatile solids; <i>Turbidity</i> ; Volatile suspended solids

Using R (v4.1.2, R Core Team 2021), I downloaded *treemapify* (v2.5.6, David Wilkins 2023) and *tidyverse* packages (v2.0.0, Wickham et al. 2019) along with their dependencies. The

data packages were then used to summarize and graph the dataset to quantify the scope and structure of volunteer water quality monitoring data across Montana. Summary metrics included the total number of volunteer monitoring groups, the number of sampling activities, the geographic coordinates of sampling sites, the number of sampling sites, the number of active projects per year, the diversity of analytes measured, and the duration and frequency of monitoring for each project. These summaries allowed me to evaluate the overall extent of volunteer-collected water quality data and changes in monitoring efforts over time.

Qualitative Coding of Goals and Objectives from Sampling Analysis Plans

All 43 SAPs were uploaded into MAXQDA (v.24.11.0, VERBI Software 2021), where the goals and objectives were extracted into a tabular format to facilitate consistent coding and comparison across groups. Goals represent an overarching monitoring intention, while objectives specify how the goal will be achieved (Makarowski and Sigler 2020). Before coding, I reviewed the tabulated goals and objectives to familiarize myself with the range of monitoring designs and common terminology. I first coded SAP goals using a deductive qualitative coding approach, in which codes were based on the predefined five goal types outlined in the MSUEQW, and MDEQ coauthored MMSG (Makarowski and Sigler 2020). I assigned each SAP goal statement to one or more of the five goal types described in the MMSG (Makarowski and Sigler 2020), depending on whether monitoring was intended to establish baseline conditions, assess trends, identify pollution sources, evaluate the project's effectiveness, or support education and outreach. Deductive methods were informed by Bingham (2023) and, by using the five MMSG goal types, the coding ensured classifications were grounded in an established educational document widely

circulated in the Montana water quality community. The deductive coding scheme for classifying monitoring goals is outlined in Table 2.

Table 2. Monitoring Goal Type Coding Scheme. The qualitative coding scheme applied to the volunteer monitoring group Sampling Analysis Plan (SAP) goals, derived through deductive qualitative coding. The five goal types were predetermined and are detailed in Markarowski and Sigler's (2020) Monitoring Methods Selection Guide. Code descriptions include how the qualitative codes were applied and examples from SAPs quote language, pulled directly from the SAPs, for each specified code type.

CODE	DESCRIPTION	EXAMPLES FROM SAPs
BASELINE	Applied to goals monitoring baseline conditions, assessing current water quality status, or characterizing environmental conditions.	"To determine current conditions..." "...establish a baseline for future comparisons..."
TREND ANALYSIS	Applied to goals describing monitoring intended to assess temporal changes, trends, or patterns in water quality. Including references to trends over time, seasonal comparisons, or detecting changes due to management actions or environmental conditions	"...to analyze long term trends in water quality..." "Evaluate spatial and temporal trends..." "...to detect temporal trends in water quality..."
SOURCE ASSESSMENT	Applied to goals identifying the source, concentration, load, or spatial distribution of pollutants, including goals focused on measuring contaminant inputs across different land uses or river reaches.	"To investigate nutrient concentrations and source areas..." "To investigate nutrient concentrations ... as they relate to location in the basin, agricultural activities, and development."
EDUCATION	Applied to goals emphasizing education, outreach, community engagement, stewardship, or public awareness; including goals related to training volunteers, informing the public, or supporting educational programming through monitoring activities.	"Develop a sustainable community science program..." "...providing a method for education and outreach about water resources." "Generate interest and engagement..."
EFFECTIVENESS MONITORING	Applied to goals focusing on the success of specific projects or comparing before and after a project's implementation.	"To assess the successes and failures of the ...Project and future restoration projects..."

Project objectives were then coded using an inductive qualitative coding approach, which allows codes to emerge directly from the content of the SAPs rather than taking an a priori

approach (Bingham 2023). The MDEQ (2022a) and MMW (2025) Volunteer Monitoring SAP templates specify four elements that should be included in all objectives: the parameters, monitoring locations, monitoring timing, and monitoring context. Based on these elements, I read all objective statements across the SAPs to identify recurring patterns in how the objectives describe the *what* (parameters), *where* (locations), *when* (timing), and *why* (context) of monitoring. I developed objective codes within the parameter, monitoring location, and monitoring timing objective elements to capture the range of SAP monitoring designs. Monitoring context was excluded from the inductive approach. I then coded objective statements to capture explicit references to each element, with objectives allowed to receive multiple codes within a category. The resulting inductive coding scheme for monitoring objectives is summarized in Table 3. Of the coded parameters, engagement is the only parameter that is not a water quality characteristic but rather an outreach and education-associated measure.

I applied thematic codes to categorize the monitoring context of objectives following Braun and Clarke's (2006) thematic analysis framework and a supplementary checklist provided by Ahmed et al. (2025). Context codes were developed by identifying recurring themes by drawing from the deductive and inductive codes, SAP keywords, introductions, and problem descriptions. This thematic coding approach allowed me to differentiate among objectives with similar technical characteristics but distinct underlying motivations. Context codes were likewise non-mutually exclusive, and the thematic coding scheme used to categorize monitoring context is presented in Table 4. All objective codes were applied such that they co-occurred with the goal types under which they were nested. Therefore, the hierarchical relationship between monitoring goals and objectives was preserved. This coding structure allowed for subsequent analysis of

relationships among goals, objectives, parameters, and monitoring design characteristics across groups. The hierarchical structure of qualitative codes used in this study is shown in Figure 3.

Table 3. Monitoring Objectives Coding Scheme. The qualitative coding scheme applied to the volunteer monitoring group's Sampling Analysis Plan (SAP) objectives. Codes were created using an inductive qualitative coding approach, where the codes emerged after reading all SAP objectives. Code descriptions detail the language or content that must be present in an objective for the code to apply. Each objective explicitly references at least one code for each of the three parent objective elements, breaking down the when (Monitoring Timing), where (Monitoring Location), and what (Parameters) for each water quality monitoring objective.

OBJECTIVE ELEMENT	CODE	DESCRIPTION
MONITORING TIMING	Unspecified / NA	Unspecified temporal window or the parameter of interest (like engagement) is not derived from field samples.
	Pre Run-off (Jan–March)	Any monitoring between January and March.
	Run-off (April–June)	Any monitoring between April and June.
	Summer Baseflow (July–September)	Any monitoring between July and September or when peak growing season is explicitly motioned.
	Winter Baseflows (October–December)	Any monitoring between October and December.
MONITORING LOCATION	Rivers / Streams	Monitoring sites are in lotic systems including rivers, streams, or creeks.
	Lakes	Monitoring sites are in lentic systems, including lakes, reservoirs, or ponds.
PARAMETERS	Nutrients*	Monitoring nutrient-related parameters.
	Sediment*	Monitoring sediment-related parameters.
	Metals*	Monitoring metals and metal concentrations.
	Biological*	Monitoring biological-related parameters.
	Chemical *	Monitoring chemical measures or general water chemistry indicators.
	Physical *	Monitoring physical water quality parameters.
	Engagement	Measuring volunteer participation, community involvement, information used in decision making, or other engagement-based monitoring metrics.

* See Table 1 for a list of the water quality characteristics classified into the nutrients, sediment, metals, biological, chemical, and physical parameter group categories.

Table 4. Objective Context Coding Scheme. The qualitative coding scheme used to categorize the context, or motivations, for each volunteer monitoring group objective. Codes were created using a thematic qualitative coding approach, in which codes emerged after reading all the Sampling Analysis Plan (SAP) objectives, and recurring themes were identified from both deductive and inductive codes. Code descriptions detail the language or content that must be present in an objective for the code to apply.

OBJECTIVE ELEMENT	CODE	DESCRIPTION	EXAMPLE FROM SAPs
CONTEXT	Condition	Monitoring overall ecosystem or waterbody condition, health, or status.	“Document the general condition...”
	Pollution/ Invasives	Monitoring pollution concerns, invasive species, contamination threats, or environmental degradation.	“...monitoring water quality exceedances of trace metals...”
	Recreation	Monitoring efforts or interests linked to the recreational use of water bodies, such as fishing, swimming, boating, or aesthetic enjoyment.	“To collect water samples at sixteen public swimming areas...when most people are swimming.”
	Land Use	References to land use influences (e.g., agriculture, urbanization, forestry) as reasons for monitoring water quality.	“Evaluate nutrient concentrations from Forest Service land and agricultural land...”
	Education	Monitoring activities emphasizing education, outreach, improved decision making, or public awareness.	“To track involvement and raise awareness about watershed health...”

In MAXQDA (v.24.11.0, VERBI Software 2021), I used the code matrix and code relation browsers under the Visual Tools ribbon to explore qualitative code co-occurrences and relationships within the SAPs. I also used the complex code patterns functions to evaluate how often specific goal types and object elements of interest co-occurred in SAPs. Where appropriate, all descriptive statistics and statistical analyses comparing projects’ WQX data and SAP qualitative codes were conducted in R (v4.1.2, R Core Team 2021).

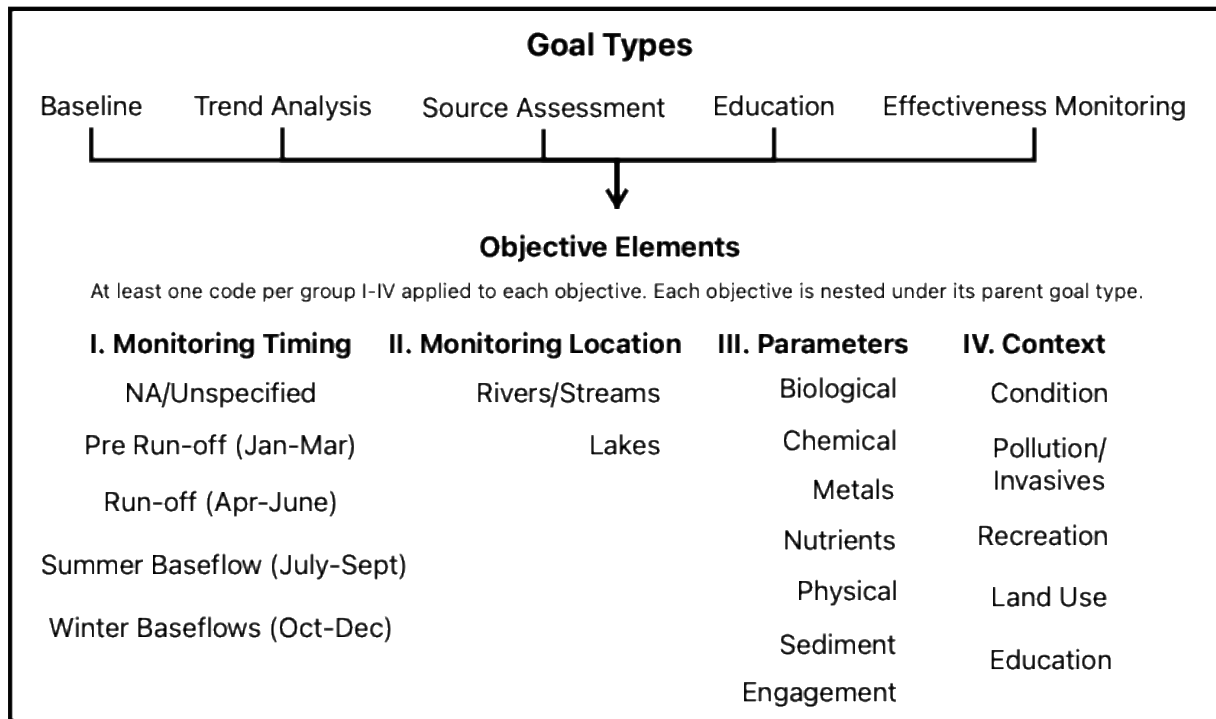


Figure 3. Goals and Objectives Hierarchy Flowchart. The qualitative coding hierarchy of goals and objectives used across the volunteer water quality monitoring groups' Sampling Analysis Plans (SAPs). Objective codes co-occur with the goal types under which they were nested to preserve the hierarchical relationship between monitoring goals and objectives. All codes are non-mutually exclusive, and objectives can have multiple codes within an objective element category.

Objectives Evaluation Framework

Using qualitative SAP and quantitative WQX project information, I developed an evaluation framework to assess how volunteer water quality monitoring projects in Montana structured their monitoring efforts relative to their stated monitoring goals. A breakdown of the evaluation framework is outlined in Figure 4. Qualitative-coded goal types derived from SAPs were exported from MAXQDA Pro (version 24.11.0) to an Excel (Microsoft Corporation 2024) spreadsheet. I arranged the spreadsheet so that each project had its own row for each parameter it aimed to measure, and the columns were binary-coded to indicate the presence (1) or absence (0)

of each of the five goal types for a given project and parameter. The Excel spreadsheet was converted to a CSV file and imported into R (v4.1.2, R Core Team 2021), where I merged the goal-type dataset with the physical and chemical water quality data obtained from the WQX using simplified Project IDs shared across datasets.

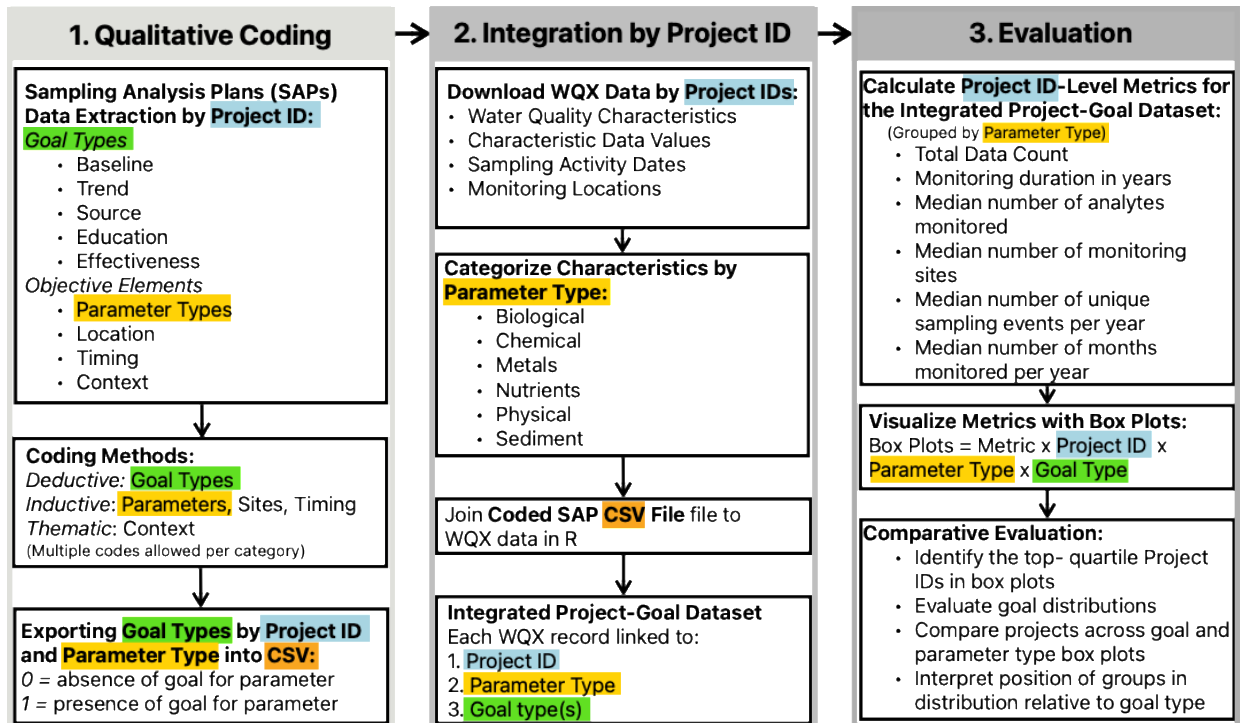


Figure 4. Qualitative Coding and Quantitative Water Quality Data Integration Framework. A graphical representation of the integration framework showing the methodological flow from qualitative coding of project Sampling Analysis Plans (SAPs), to integrating qualitative codes into Water Quality Exchange (WQX) data by project ID. At the evaluation stage, integrated project data is then investigated by parameter and goal type. Color coding indicates the type of project information that passes through each stage of the framework.

The data integration allowed each water quality record to be linked to one or more monitoring goal types, enabling evaluation at the project-goal level. After integration, I grouped the dataset by project, parameter type, and goal type to calculate project-level metrics characterizing the monitoring effort. If a project measured a parameter that was not detailed in its

SAP, the SAP was re-evaluated in its entirety to ensure complete coding. If the collected characteristic supported another parameter type's interpretation, like measuring discharge as a physical parameter to help interpret baseline sediment loads, the missing parameter was coded under the goal type it supports (baseline in this example). If the measure did not support the interpretation of another parameter type detailed in the SAP, I listed all goals as absent for all five goal types. Similarly, if a parameter type was coded for in a project's SAP but did not match any WQX data, then it was not associated with any goal type in the subsequent analysis.

Evaluation metrics were selected to reflect project monitoring design and objective elements rather than the data values and included total data count, sampling duration in years, median number of analytes measured per year, median number of monitoring sites, median number of unique sampling events per year, and median number of months monitored per year (Figure 4). All metrics were calculated for each project-parameter type combination so that a project with parameters falling under different goal types could contribute data to each goal category. For example, a project monitoring nutrient and metal trends would have project metrics calculated for both nutrient-only and metal-only data, since these two parameters may have different monitoring designs. Engagement was the only parameter type not recorded in the WQX dataset and, therefore, not integrated into the evaluation framework.

To compare projects with the same monitoring goals, I plotted the project-level metrics using side-by-side boxplots, grouped by goal type, and subset the boxes by parameter type. I used the *ggh4x* (v0.3.1, van den Brand 2025) package to create faceted boxplots. Each boxplot panel shows the distribution of a given metric and parameter type across all projects associated with a given goal type, with individual projects plotted as points overlaid on the boxes to

preserve project-level resolution. Since metrics were calculated at the parameter-type level, a project may appear multiple times per panel if it monitored multiple parameters. This box-plot approach allowed me to compare both central tendencies and variability in monitoring efforts within and across goal types, all while retaining insight into individual monitoring projects. The effectiveness monitoring goal type had only two projects, so point plots were used rather than boxplots.

I chose to use medians in boxplots to summarize and compare monitoring metrics because volunteer water quality monitoring datasets are heterogeneous in duration and design. Project-level metrics, such as sampling frequency, the number of analytes, and the number of monitoring sites, were right-skewed and thus sensitive to outliers. Medians provide a robust measure of central tendency that better reflects characteristic monitoring efforts within each goal type while retaining the ability to interpret metrics across heterogeneous projects. Boxplots also allow visualization of variability and extreme values, which helps in understanding the breadth of monitoring strategies within volunteer programs. Overlaying individual project points maintains transparency at the project level and supports the identification of groups whose monitoring characteristics stand out from the distribution. Together, using medians and boxplots offers a clear, high-level approach that is tailored to exploratory, right-skewed, goal-based evaluation of citizen science monitoring programs.

Within each set of boxplots, I highlighted a subset of noteworthy projects to facilitate cross-plot comparison. The identified projects were selected by examining distributions across goal types and selecting groups that consistently fell within the upper quartile of one or more metrics. My selection criteria also emphasized representation across multiple goal types and

inclusion of groups monitoring both lakes and streams to capture a range of monitoring contexts. Final selections were reviewed in consultation with Dr. Sigler of MSUEWQ to ensure that highlighted projects were representative of volunteer monitoring efforts. Overall, this framework provides a structured, reproducible approach to evaluating water quality monitoring programs by linking monitoring goals to measurable characteristics of monitoring designs, encompassing several objective elements. By organizing project evaluation around goal types and using comparable project-level metrics, this approach supports assessing whether monitoring programs are well positioned relative to peers to meet their stated purposes. The framework serves as an evaluative lens for understanding variation in monitoring strategies across a diverse volunteer water quality monitoring landscape rather than serving as a performance ranking tool.

RESULTS

Breadth of Volunteer Monitoring and the Parameters of Interest

Of the 43 identified volunteer water quality monitoring projects, 40 demonstrate a high level of autonomy in their planning documentation, while three (Clearwater Nut, Gallatin WC, and Gallatin LWQD) rely more heavily on government or agency support (Figure 2). I excluded 11 projects from the analysis due to missing SAPs or WQX data, despite their involvement of volunteers and their listing in the MWCC library. Volunteer water quality monitoring locations across Montana are concentrated in the state's western portion (Figure 5), with 56.5% of sites west of the continental divide. There are 115 sub-basins (HUC-8) watersheds in Montana, and 59 are completely contained within the state's borders. Of the 822 monitoring locations across the 43 projects, 36.9% fall within the Blackfoot, Gallatin, and Flathead Lake HUC-8 watersheds,

which are shown in dark red and orange in Figure 5. Only 30.4% of Montana's HUC-8 watersheds (spanning 187,747 km²) contain any volunteer monitoring sites. Monitoring sites cluster around population centers (Figure 1), including Bozeman, home to Montana State University; Missoula, home to the University of Montana; Kalispell; and Billings.

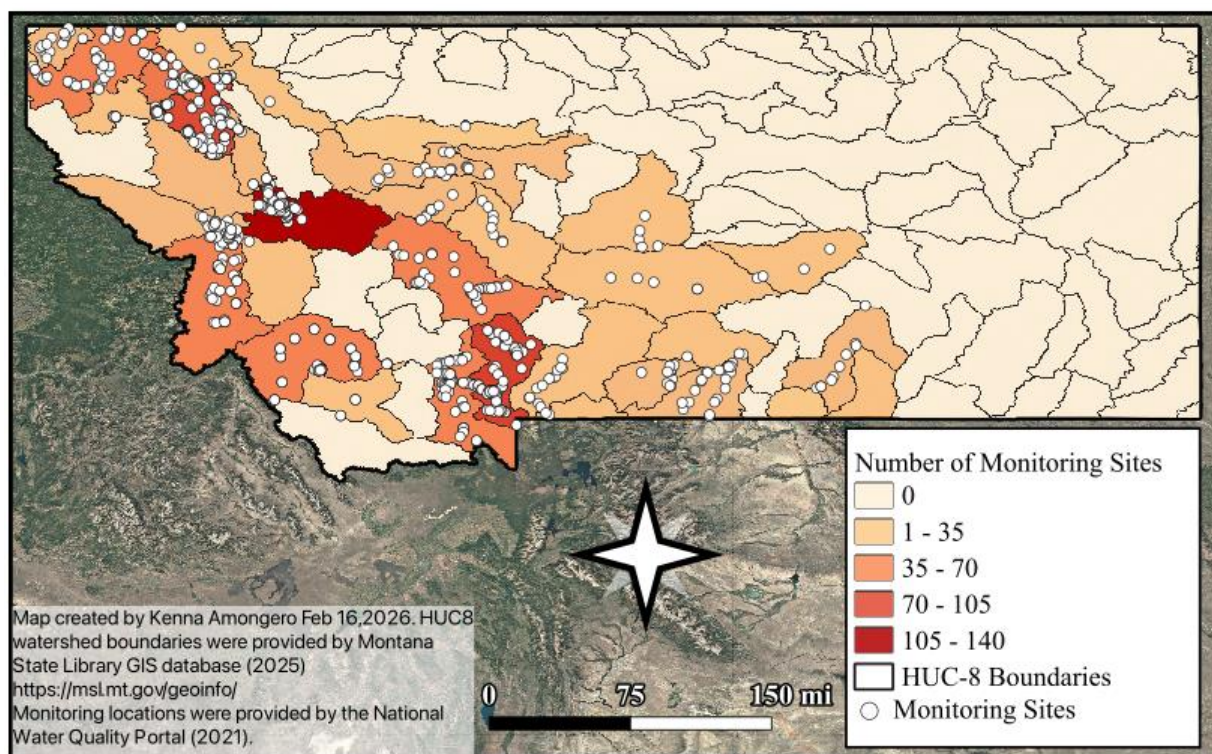


Figure 5. Map of Volunteer Monitoring Sites and Density. A HUC-8, or sub-basin level, heatmap showing the spatial distribution of volunteer water quality monitoring sites across Montana using the NAD83 coordinate system. Each white dot represents a monitoring location where at least one data point was collected.

Volunteer water quality monitoring projects with publicly accessible WQX data began in 2001 and run through present (Figure 6 & Figure 7). Several monitoring projects are no longer active or have been absorbed into other monitoring efforts. For example, NWM Lakes 1 and NWM Lakes 2 were run by the same monitoring group, but the two projects initially operated under different SAPs. In 2013, NWM Lakes 1 was absorbed into the NWM Lakes 2 project in

the WQX (Figure 6, Whitefish Lake institute 2023). Similarly, Gallatin TF VM is the legacy group for Gallatin TF Nut. In 2019, Gallatin TF VM received a grant through MDEQ and redrafted its SAP, which led to the group's transition into the Gallatin TF Nut monitoring effort (Gardner et al. 2024). Sampling activities occur intermittently throughout the year, and some groups, such as the Flathead Lakers and Gallatin WC, have data gaps of one or more years between sampling events (Figure 6). Seven monitoring projects have data for a single year, while five projects span more than a decade, with the Sun being the longest-running volunteer monitoring project. The median number of years monitored across all groups is four years (*IQR* = 2-7 yrs).

Active volunteer monitoring projects have continued to rise since 2001, with 21 active projects recorded in 2024 (Figure 7). Since 2008, the number of active projects has steadily climbed, with a brief plateau between 2012 and 2014. In 2022, the number of active projects dipped when Whitefish LI, Clearwater Lakes, NWM Lakes 2, and Big Hole WC Elkhorn projects became inactive in the WQX, with the Flathead Lakers and Big Horn RF also missing a year of data collection (Figure 6 & Figure 7). Seeley CRC, Rock CCRC, and Kootenai began their monitoring efforts in 2022, resulting in a net decrease of three projects (Figure 6). Only the Sun and Madison projects were active during the first seven years of volunteer water quality monitoring in Montana (Figure 6 & Figure 7). Of the 21 monitoring projects active in 2024, three had WQX data available for 2025 (Figure 6), ahead of the February 2026 submission deadline (Flathead Lake Bio Station 2026b), when the data for this project was downloaded in December of 2025.

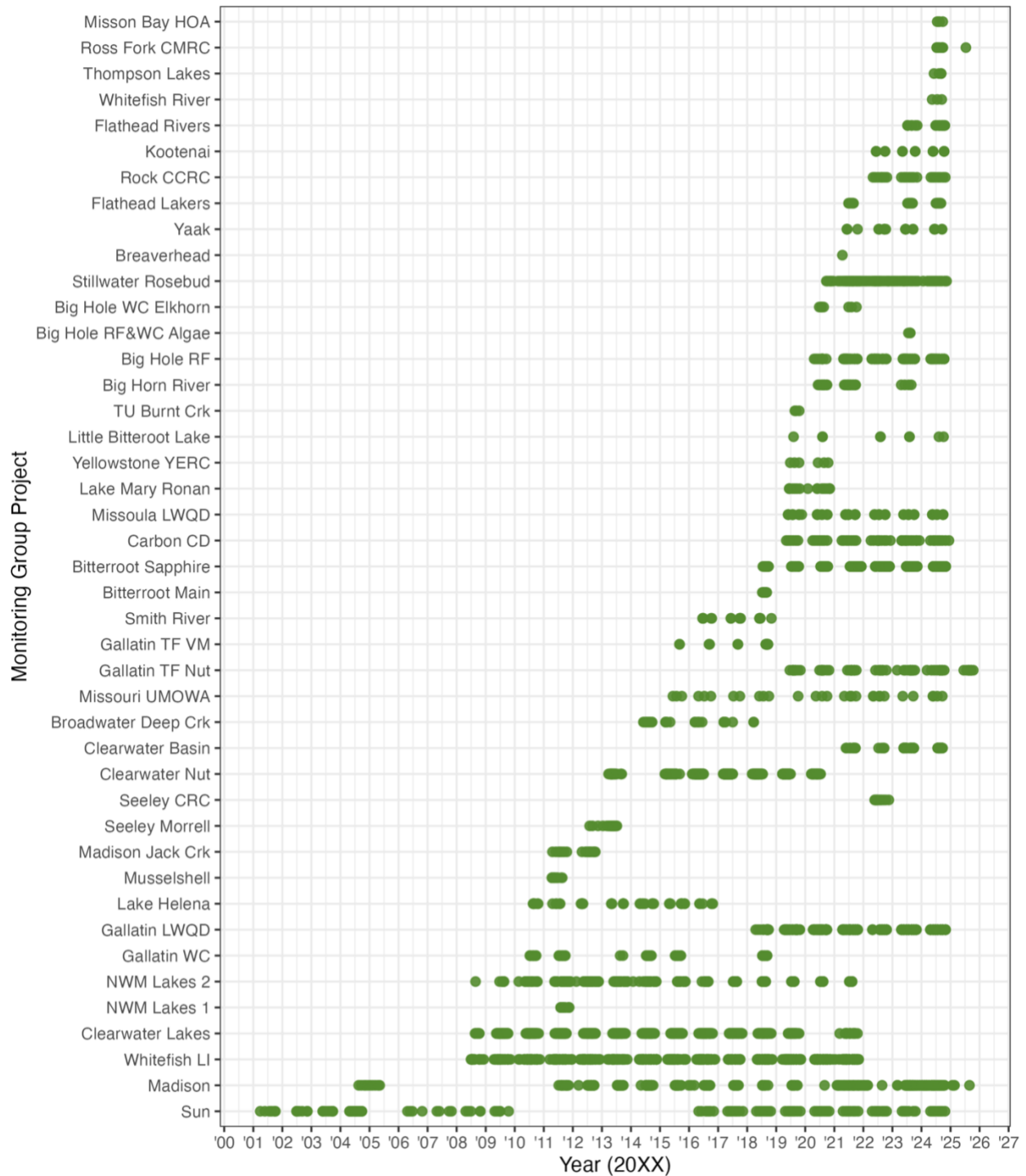


Figure 6. Monitoring Group Data Collection Timeline. Timeline of water quality sampling activities for each volunteer monitoring project in Montana. Sampling activities represent a unique event during which water quality data were collected.

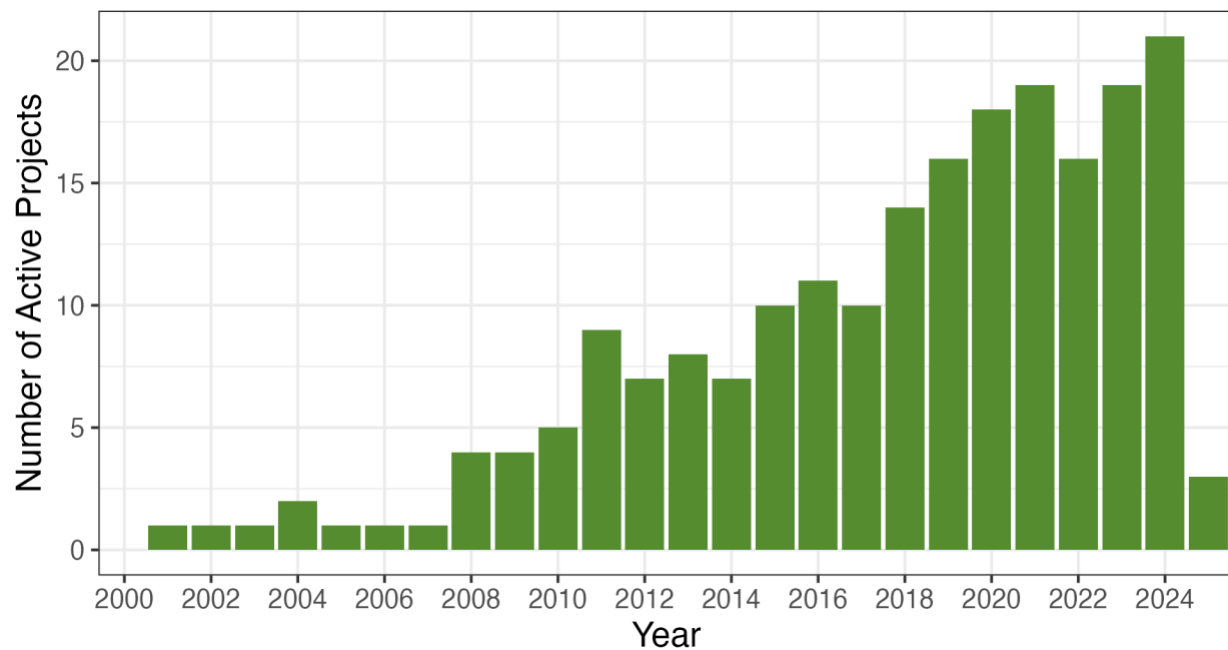


Figure 7. Monitoring Group Count by Year. Number of active volunteer monitoring projects in Montana per year since 2001. A project is considered active if water quality data is collected during a given year.

The most frequently cited parameters of interest across volunteer monitoring project SAPs are physical and nutrient parameters. Yet, physical and chemical parameters account for the greatest number of data points, with physical counts 2.5 times greater than chemical counts (Figure 8). Specific conductance, water temperature, and dissolved oxygen saturation make up 80.2% of the physical water quality characteristics (Figure 9). All physical parameters were collected in the field, whereas all nutrient parameters were lab-derived (Figure 8 & Table 1). Dissolved oxygen and pH are the only chemical parameters collected in the field, accounting for 76.7% of the chemical data (Figure 8 & Figure 9). The remaining 30 chemical parameters are lab-derived (Table 1). The nutrient parameter type is primarily composed of phosphorus (mixed forms, 29.35%), nitrate plus nitrite (26.72%), and total nitrogen (mixed forms, 22.45%) (Figure 9). Turbidity and total suspended solids account for 81.9% of sediment data counts and 6.46% of

the 196,224 total data counts. Only 9% of projects aimed to collect metal parameters in the SAPs, and metals account for 1.54% of the data across projects. Of the 101 water quality characteristics collected by volunteer monitoring groups, ten characteristics account for over 77% of the data (Figure 9).

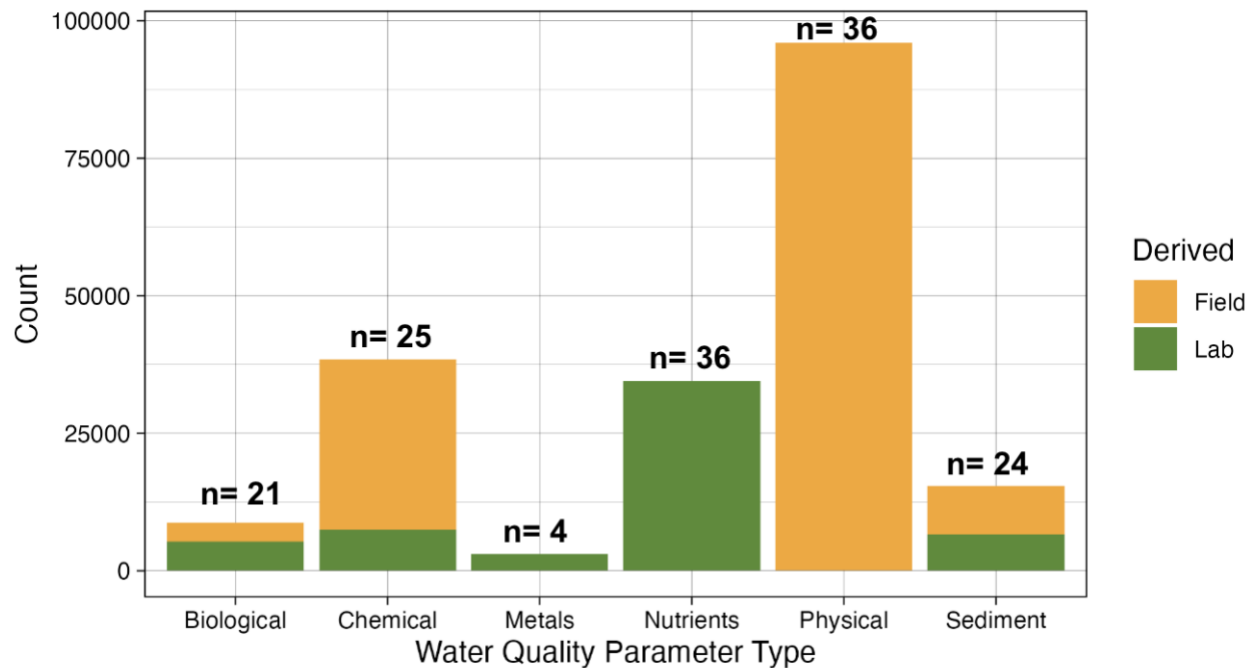


Figure 8. Data Collection Quantified by Parameter Type. Amount of data collected across parameter types by volunteer monitoring groups, categorized as field- or laboratory-derived. Numbers above bars indicate the number of volunteer monitoring projects measuring each parameter type. For a breakdown of which water quality characteristics are included in the parameter types, see Table 1.

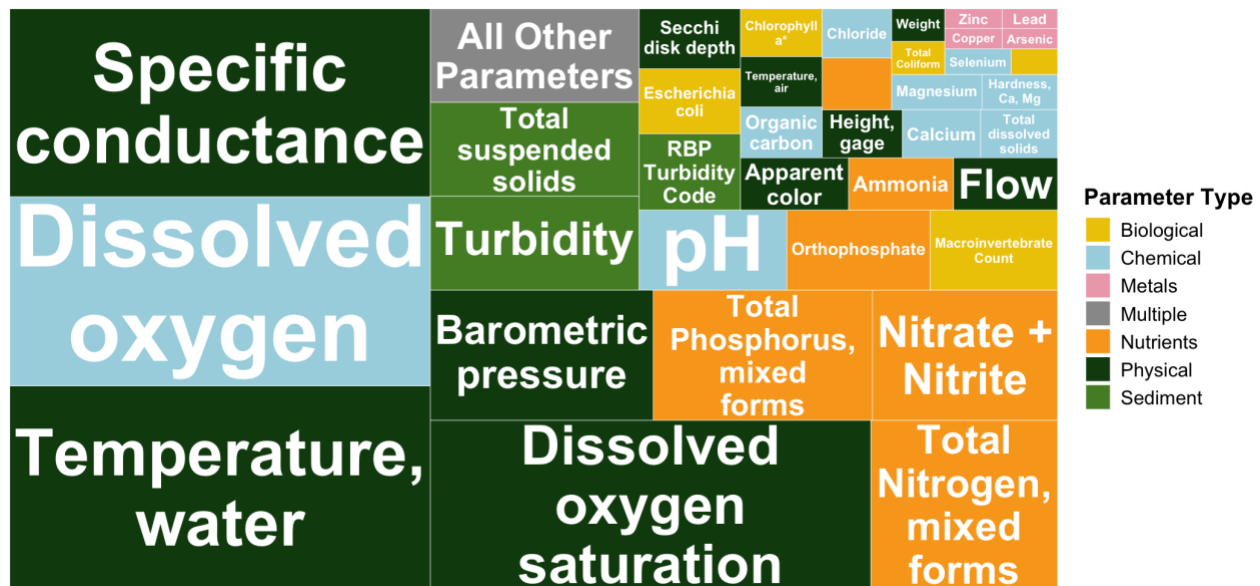


Figure 9. Top Parameter Collection Tree Map. Tree map showing the most frequently measured parameters across all groups of the total 101 characteristics collected. Tree map boxes are proportional to the number of data points collected for the labeled water quality characteristic. Water quality characteristics are color-coded by parameter type according to Table 1 classifications. The 64 parameters with counts under 350 were grouped into the “All Other Parameters” class.

Qualitative Codes Across Sampling Analysis Plan Goals and Objectives

Across water body types, baseline goals are most abundant in the volunteer monitoring project SAPs. Of the 33 projects with baseline goals, 96.7% are monitoring during summer baseflows, and 84.8% have a condition context (Table 5 & Figure 10A). Baseline goals are the only goal type that overlap with a recreational monitoring context (Table 5). For all five goal types, the most common monitoring timing is during summer baseflows, though one monitoring project (Lake Helena) did not collect data during this period (Figure 10B & Figure 6). The second most common sampling period for projects is during spring runoff from April to June. Only five projects monitor water quality year-round. Additionally, 80% of projects with a trend analysis goal monitor under a condition context, yet 20% of the projects sample exclusively

during summer baseflows (Table 5, Figure 10). Neither of the two projects with an effectiveness monitoring goal collects data year-round or has event-based sampling, yet both aim to address the condition of the rivers they monitor after sediment-reduction restoration efforts (Kreiner et al. 2015, Gardner et al. 2018). Over three times as many projects monitor rivers and streams as monitor lakes in Montana (Figure 10D).

Table 5. Goal Type and Objective Co-occurrence Matrix. A qualitative coding co-occurrence matrix for all the goal types and objective elements for the 43 monitoring projects. Numbers represent the total co-occurrences of the qualitative codes across the Sampling Analysis Plans (SAPs) and are color-coded to highlight the codes that frequently overlap. Objective elements are nested under goal types; therefore, each objective element across SAPs must have at least one occurrence of a goal type. Co-occurrences indicate how many SAPs the two terms overlap in.

OBJECTIVE ELEMENTS		GOAL TYPE				
		Baseline	Trend Analysis	Education	Source Assessment	Effectiveness Monitoring
MONITORING TIMING	Unspecified/ NA	5	1	10	3	1
	Pre Run-Off (Jan-Mar)	4	1	2	2	0
	Run-Off (Apr -June)	26	7	8	13	1
	Summer Baseflow (July-Sep)	32	15	8	17	2
	Winter Baseflow (Oct-Dec)	12	6	6	8	1
MONITORING LOCATION	Rivers/Streams	26	12	12	18	2
	Lakes	10	4	3	3	0
PARAMETERS	Nutrients	24	13	6	16	1
	Sediment	14	7	6	10	2
	Metals	3	1	0	1	0
	Biological	15	9	4	6	2
	Chemical	19	7	3	9	0
	Physical	27	11	7	8	1
	Engagement	5	1	10	2	0
CONTEXT	Condition	28	12	5	6	2
	Pollution/Invasives	8	1	1	3	0
	Recreation	3	0	0	0	0
	Land Use	12	6	2	15	0
	Education	4	2	12	2	0

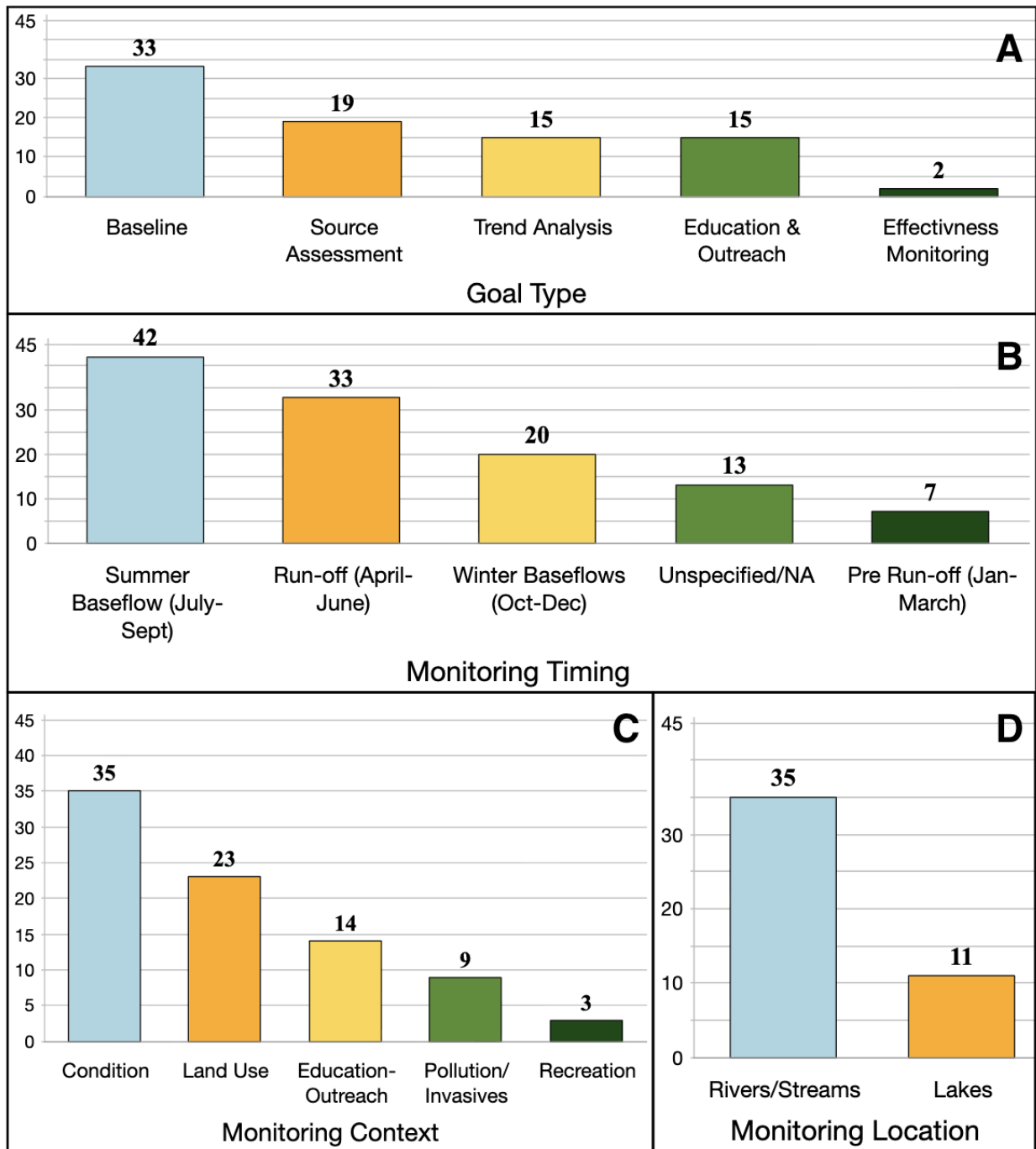


Figure 10. Goal and Objective Type Counts. The (A) frequency of different goal types and objective elements, including (B) monitoring timing, (C) monitoring context, and (D) monitoring location across the 43 Sampling Analysis Plans (SAPs). Bar charts show how many SAP documents contain each qualitative code. Codes are not exclusive, and more than one code may be applied to goals and objectives.

While an equivalent number of projects monitor nutrient and physical parameters, 27 baseline projects monitor physical parameters, and 24 monitor nutrients (Table 5). The source assessment goal type shows the greatest co-occurrence with nutrient parameters, land use contexts, and river monitoring locations, with 13 of the 19 source assessment projects exhibiting all four co-occurrences. Unspecified monitoring timing is 92.3% attributed to the engagement parameter, which primarily co-occurs with education and outreach monitoring contexts (Figure 11). Of the fifteen projects with education and outreach as a monitoring goal, fourteen listed it as a reason for monitoring in their objectives (Figure 10C).

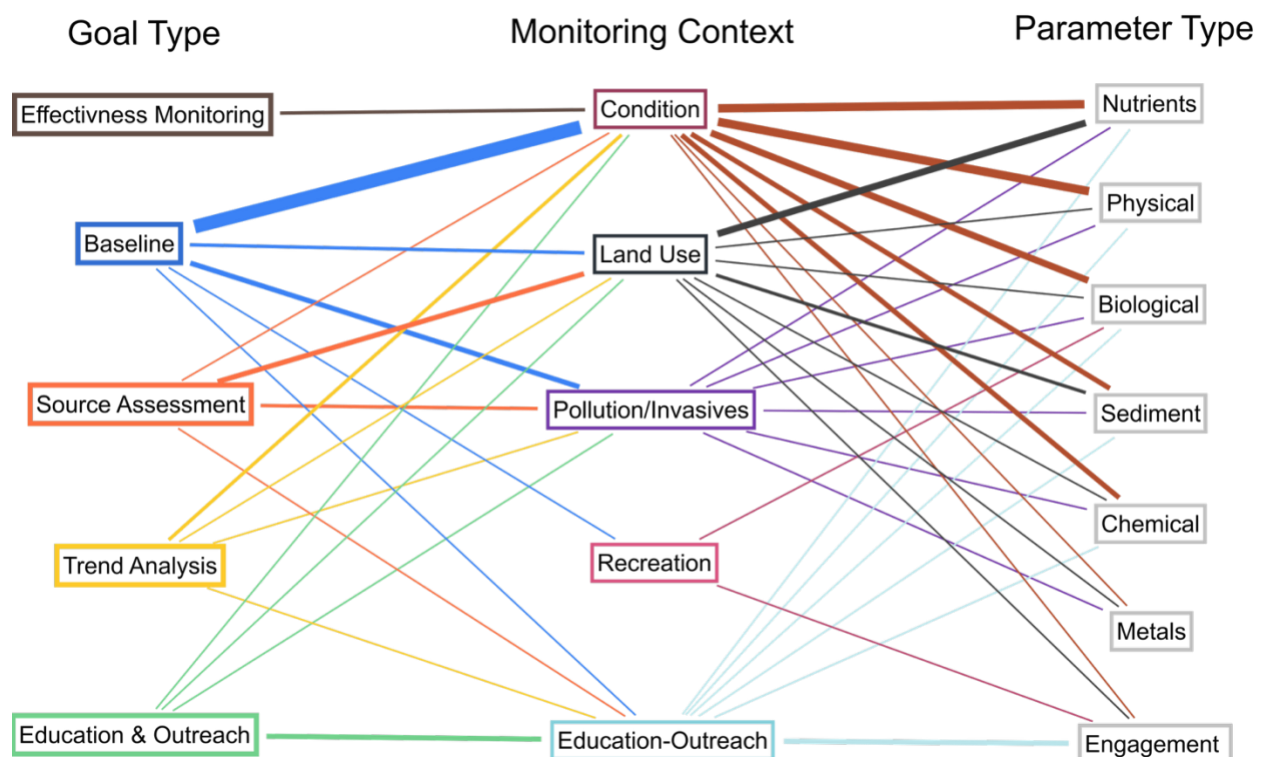


Figure 11. Goal and Objective Co-occurrence Network Diagram. Network diagram generated using the MAXQDA code-mapping feature, illustrating the qualitative coding co-occurrences among monitoring goal types, objective contexts, and objective parameter types. All codes were derived from Sampling Analysis Plans (SAPs), and line thickness between codes reflects the frequency of code co-occurrence across SAPs. Line colors correspond to goal types and objective contexts.

A SAP keyword search revealed that of the 28 projects without education as a monitoring goal, two mention education or outreach in their SAP narrative. The only goal type not represented in the education and outreach context is effectiveness monitoring, and all but the metals parameter type co-occur with this context category. The recreation context has the fewest co-occurrences with other goal and objective codes (Figure 11). In contrast, the Figure 11 co-occurrences that stand out are between baseline goals and condition contexts for physical, nutrient, and biological parameters, and between source assessment goals and land-use context for nutrient parameters. Finally, the biological parameter type is the only parameter to co-occur with all five monitoring contexts.

Goal and Objective Evaluation Framework

When the parameter types of each of the 43 volunteer water quality monitoring projects were listed and paired with their goal types, there were a total of 159 different project-to-parameter-to-goal(s) assignments in the integrated project-goal dataset (Figure 4). All the projects included more than one parameter in their SAP objectives, and all WQX data aligned with the groups' SAP parameters of interest. Eight parameters of interest, as indicated in SAPs, were not collected by the motioning groups and uploaded to the WQX. The remaining 151 rows in the integrated project-goal dataset (Figure 4) are visualized by goal type in Figures 12-16 and grouped by parameter.

Groups with a Baseline Goal Type

There are 103 project-to-parameter-to-goal assignments under the baseline goal type, making baseline goals most abundant (Figure 12). The most common parameter type monitored

under baseline is nutrients, with 78.8% of groups collecting nutrient data. The number of points collected ranges from ten to 10^4 across groups and parameters and is heavily right-skewed. Physical parameters have the largest IQR, ranging from 1 to 37,006. However, the median number of data points across parameter types was similar, ranging from 183 to 600. All but three projects have been collecting data for under ten years, with the Madison and Whitefish LI projects standing out as outliers. For all parameter types, 50% of the projects have been operating for five years or less. Duration boxplots are similar in shape and IQR across parameter types, with metals being the only parameter with all groups monitoring for less than ten years. In contrast, the boxplot of the number of analytes measured for metals shows higher values than for all the other parameters, except for one project's chemical outlier. The median number of metal analytes measured per year is eight, which is double the next-highest median in the physical parameter category. The sediment category in Table 1 lists the fewest number of WQX characteristics ($n = 7$), while the chemical category lists the greatest number of characteristics ($n = 32$). Yet, for baseline goals, the median number of analytes measured per year clustered within a narrow range for most parameter types. Projects with physical and nutrient parameters have higher upper quartiles for median analyte counts than those with sediment and biological parameters.

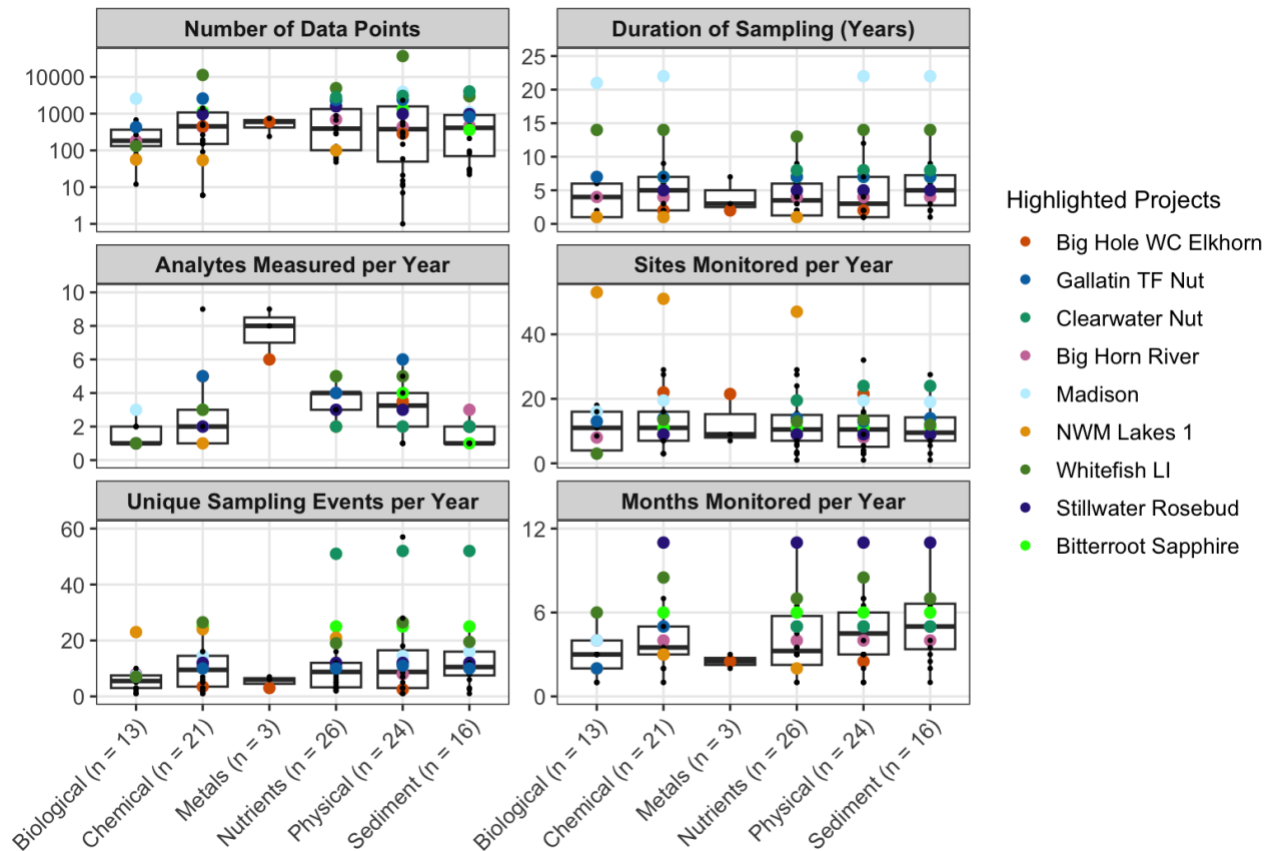


Figure 12. Baseline Goal Projects Data Collection Boxplots. Side-by-side boxplots of data counts, sampling duration in years, median number of analytes measured per year, median number of monitoring sites, median number of unique sampling events per year, and median number of months monitored per year for projects with a baseline goal type, grouped by parameter type. Projects of interest are highlighted in color to facilitate cross-panel comparison, and the number of projects falling within each parameter type is included in x-axis labels.

The median number of sites monitored per year is generally consistent across parameters, with most projects monitoring 15 sites annually. The distributions are right-skewed, with NWM Lakes 1 monitoring substantially more sites than the median. Projects like Clearwater Nut and Madison are positioned above or near their parameter's upper quartiles for the median number of sites monitored per year, indicating broader spatial coverage than groups like Stillwater Rosebud, which have fewer annual monitoring sites. Unique sampling events and the number of monitored months per year show similar patterns across parameter types. Only three baseline projects have

sampling frequencies consistently above the 75th percentile. Among the highlighted projects, Whitefish LI and Bitterroot Sapphire fall in the upper range for both unique events and months monitored and have a comparatively higher temporal resolution. Overall, the median number of unique sampling days per year for all baseline projects, regardless of parameter type, is eight (*IQR* = 3–13.25 *days/yr*; *max* = 57 *days/yr*), and the median number of months monitored per year is 3.5 (*IQR* = 2.5–5.5 *months/yr*; *max* = 11 *months/yr*). Approximately one-quarter of parameters are monitored for more than 5.5 months each year.

Groups with a Source Assessment Goal Type

There are 52 project-to-parameter-to-goal assignments under the source goal type (Figure 13). Only one project with a source goal monitored metals; therefore, the boxplot cannot aid in interpreting the distribution of metal parameters across metrics. Source projects show a more negative skew in the number of data points and less dispersion in sampling durations across parameter types compared to baseline, trend, and education goal types (Figure 13). The sampling duration for all but nutrient parameters is less than ten years (*IQR* = 2–6 *yrs*; *max* = 14 *yrs*), indicating mid- to short-term investigations. Median analytes per year remain below seven across projects, with physical and metal parameter types showing higher analyte values than sediment (*Mdn* = 2 *analytes*, *IQR* = 1–2 *analytes*) and biological (*Mdn* = 1 *analyte*, *IQR* = 1–1 *analyte*) parameter types. The *IQR* (2–5.5 *analytes*) is broader for the median number of nutrient analytes per year than for the baseline goal type (*IQR* = 2–4 *analytes*), indicating greater heterogeneity in source project study designs. The median number of sites monitored per year across parameters follows a pattern similar to that observed in Figure 12 & Figures 14–15, where all upper quartiles hover at or below 20. However, the source goal type is the only one with no extreme outliers in

the number of sites monitored per year. Temporal metrics, including unique sampling events and the number of months monitored per year, reveal that most source projects conduct monitoring for fewer than six months annually, with 50% of points clustered below four months. However, projects like Clearwater Nut demonstrate extended temporal coverage and higher sampling frequencies for nutrient, physical, and sediment parameter types.

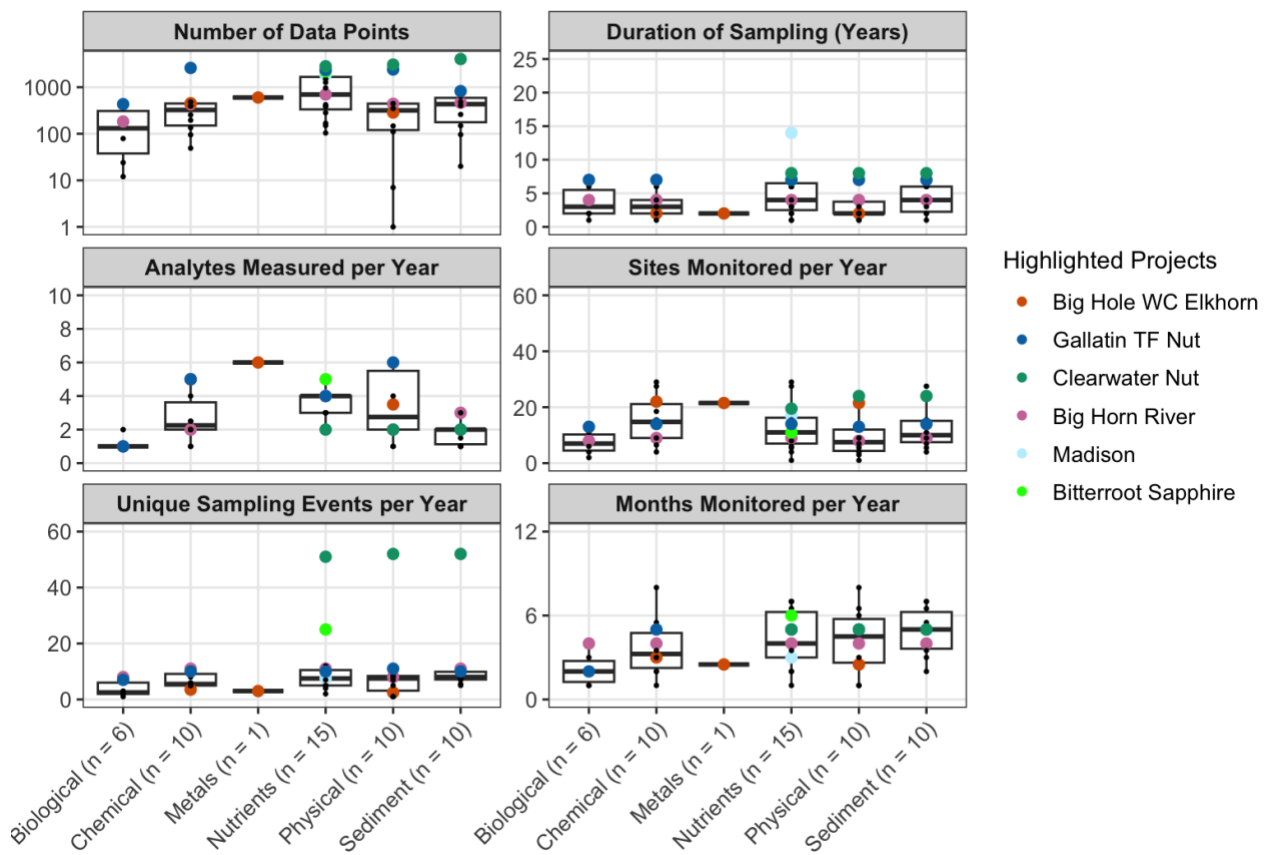


Figure 13. Source Assessment Goal Projects Data Collection Boxplots. Side-by-side boxplots of data counts, sampling duration in years, median number of analytes measured per year, median number of monitoring sites, median number of unique sampling events per year, and median number of months monitored per year for projects with a source assessment goal type, grouped by parameter type. Projects of interest are highlighted in color to facilitate cross-panel comparison, and the number of projects falling within each parameter type is included in x-axis labels.

Groups with a Trend Goal Type

Like the source goals, there are 52 project-to-parameter-to-goal assignments under the trend goal type (Figure 14). Compared to projects with a baseline goal, trend project goal types generally exhibit longer sampling durations, with multiple parameter types showing higher median and maximum years of monitoring. The distribution of data points across parameter types is right-skewed, yet all parameter-type medians fall between 100 and 1,000. The median duration of sampling across all parameters was four years (*IQR = 1-9 yrs; max = 24 yrs*), and five (25%) of projects with trend goals have been monitoring for at least nine years. Extreme duration outliers occur more frequently in this goal type than in any other. As in Figure 13, one project monitored metals; therefore, the interpretations of metals under the trend goal type are limited.

There are no outliers under the median number of analytes measured, and the nutrients boxplot shows positive skew, which is the opposite of the skew observed in the comparable plot in the baseline and source figures. Half of the median number of sites monitored per year data points equal nine or fewer, and the upper quartiles are more variable than observed under baseline goal types. The median number of unique sampling days per year for all trend projects is six (*IQR = 3-11 days/yr; max = 24 days/yr*), and medians across parameter types all cluster around seven days. The median number of months monitored per year is three (*IQR = 2-3 months/yr; max = 6 months/yr*), which is less than the baseline, source, and education goal types. The Madison project is in or above the upper quartile across nutrient, physical, and sediment parameters for all six trend metrics. Compared to other goal types, trend projects appear to emphasize extended duration rather than high within-year frequency. The distributions of temporal metric box plots also appear more compressed, with fewer extreme high-value outliers.

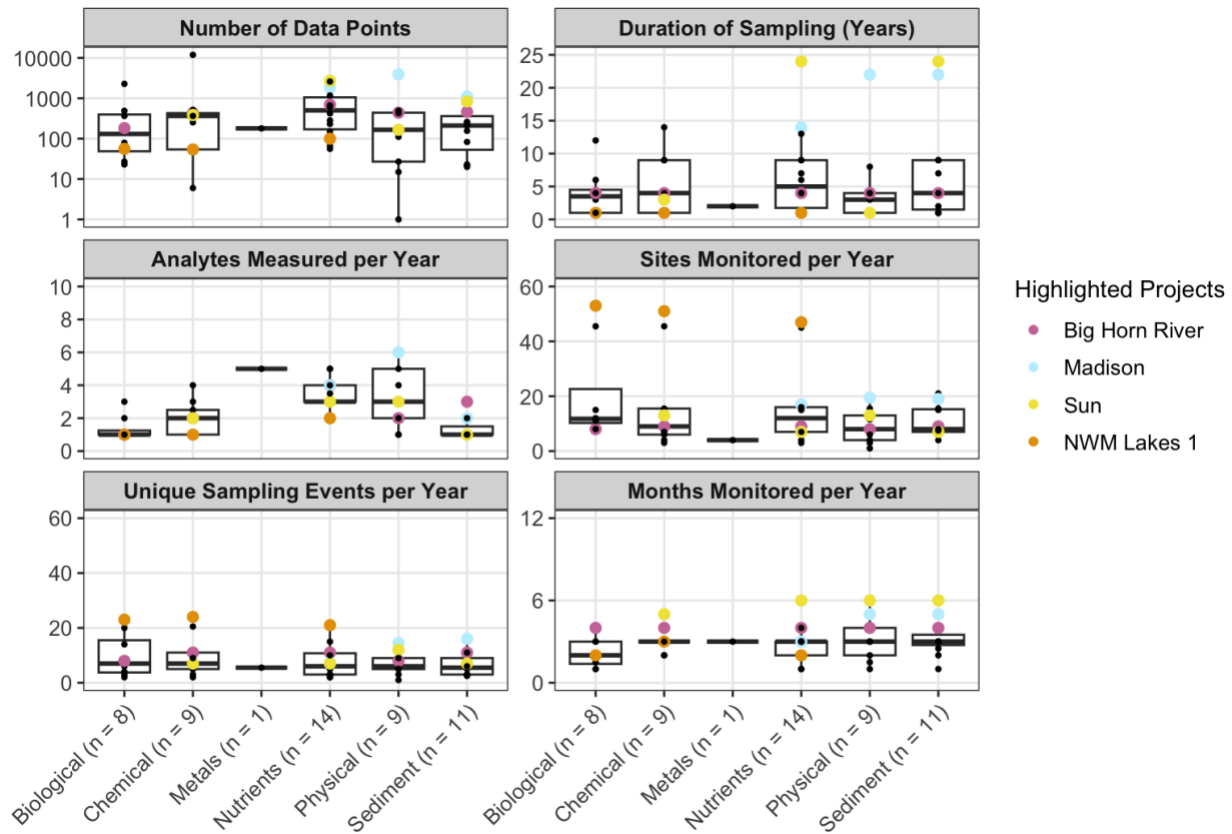


Figure 14. Trend Analysis Goal Projects Data Collection Boxplots. Side-by-side boxplots of data counts, sampling duration in years, median number of analytes measured per year, median number of monitoring sites, median number of unique sampling events per year, and median number of months monitored per year for projects with a trend goal type, grouped by parameter type. Projects of interest are highlighted in color to facilitate cross-panel comparison, and the number of projects falling within each parameter type is included in x-axis labels.

Groups with an Education Goal Type

There are 37 project-to-parameter-to-goal assignments under the education goal type (Figure 15). Education projects have parameter-specific data point count IQRs that range up to a factor of ten, as similarly observed in Figures 12-14. Sampling durations have a median of two years across all parameters ($IQR = 2-12$ yrs; $max = 24$ yrs). Half the project-to-parameter combinations span two years or fewer. However, the upper quartile for biological and physical parameters extends to 12 years. The Sun project is the longest-running volunteer monitoring

effort (Figure 6) and has both a trend and an education goal. The position of the Sun datapoints within the six-metric boxplots for the trend (Figure 14) and education (Figure 15) goal types is similar in all but the median months monitored per year metric plot. In Figure 14, the Sun project has the most months monitored for its project parameters, while in the education plot (Figure 15), the Sun project points fall at the parameter medians. Approximately 27% of education projects monitor more than six months annually.

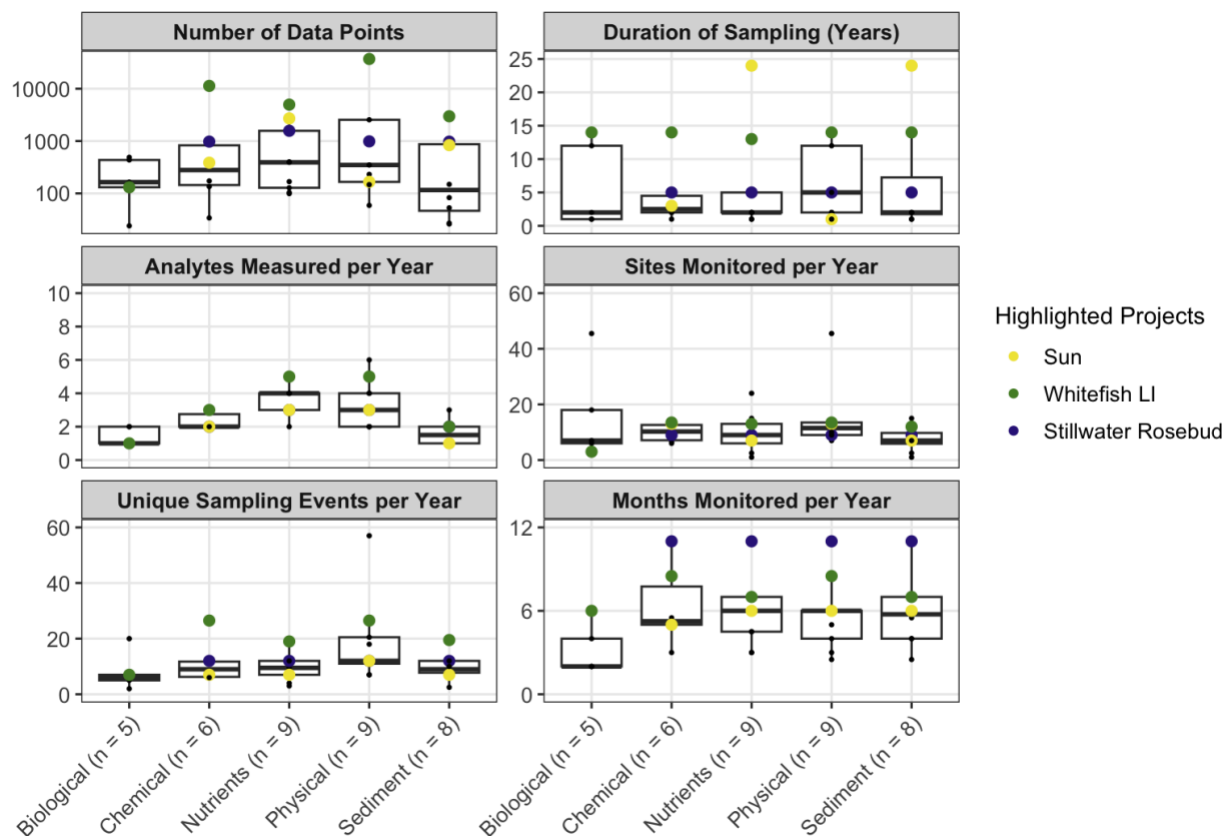


Figure 15. Education Goal Projects Data Collection Boxplots. Side-by-side boxplots of data counts, sampling duration in years, median number of analytes measured per year, median number of monitoring sites, median number of unique sampling events per year, and median number of months monitored per year for projects with an education goal type, grouped by parameter type. Projects of interest are highlighted in color to facilitate cross-panel comparison, and the number of projects falling within each parameter type is included in x-axis labels.

The median number of nutrient, biological, chemical, and sediment analytes measured per year ranges from zero to one across Figures 12-15. No volunteer monitoring projects with an educational goal monitored metals. There are 21 projects with an educational goal that monitor ten or fewer sites per year, and the IQR for the number of sites monitored per year across parameters is the narrowest of all goal types (*Mdn* = 9 sites/yr; *IQR* = 7-13 sites/yr). The median number of unique sampling days per year is ten (*IQR* = 7-12 days/yr; *max* = 57 days/yr), and the median number of months monitored per year is 5.5 (*IQR* = 4-7 months/yr; *max* = 11 months/yr), which is the highest cross-parameter median months monitored among all goal types.

Groups with an Effectiveness Monitoring Goal Type

Figure 16 displays the projects coded with an effectiveness monitoring goal type. This category includes two projects; therefore, boxplots are replaced by individual points due to the limited sample size. The variability observed in Figure 16 is primarily driven by differences among individual projects rather than broader distributional patterns. The number of data points for each parameter type falls at the lower end of the distribution compared to other goals, with the maximum biological data count of 435 being less than 21.0% of the biological data counts for the other four goal types. Sites monitored, and sampling durations between the projects vary across the measured parameters but remain within the ranges observed in the box plots for the other four goal types. The median number of unique sampling events per year for Gallatin TF VM and Broadwater Deep Crk is six (*IQR* = 3-7) across all parameters, which is in the 35th percentile compared to all other goal types. Both groups monitor for less than half the year. Gallatin TF VM has more sites monitored per year than Broadwater Deep Crk, indicating a broader spatial coverage. For the Broadwater Deep Crk project, only physical parameters were

sampled for more than one year, while Gallatin TF VM monitored both biological and nutrient parameters for multiple years.

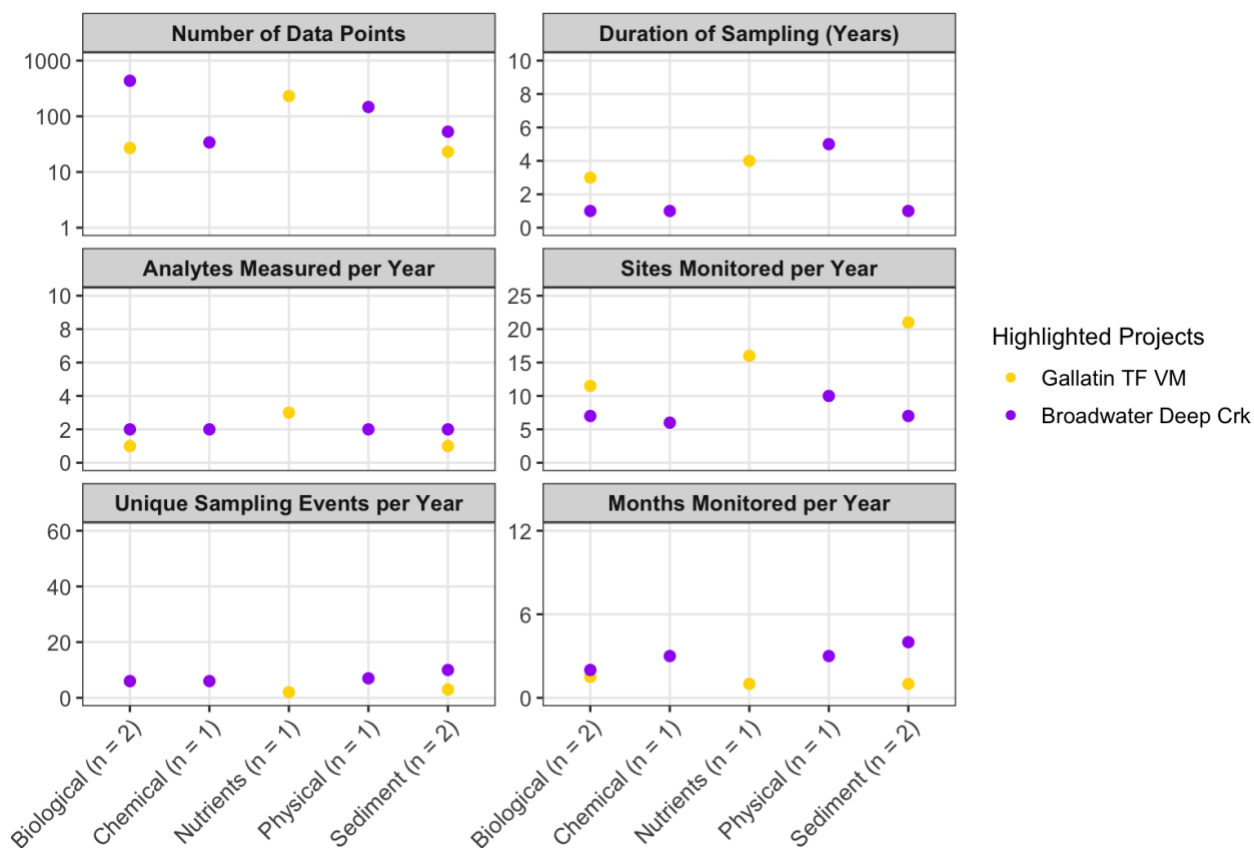


Figure 16. Effectiveness Monitoring Goal Projects Data Collection Boxplots. Side-by-side boxplots of data counts, sampling duration in years, median number of analytes measured per year, median number of monitoring sites, median number of unique sampling events per year, and median number of months monitored per year for projects with an effectiveness monitoring goal type, grouped by parameter type. Projects of interest are highlighted in color to facilitate cross-panel comparison, and the number of projects falling within each parameter type is included in x-axis labels.

Cross Goal Comparisons

Data counts are strongly right-skewed across all goal types (Figures 12-16). The number of data points collected by projects is higher for groups that either measure multiple analytes per parameter or are in the top 75th percentile for unique sampling events per year. Trend projects

(Figure 14) exhibit the longest upper-range durations ($Mdn = 4$ yrs; $IQR = 1-9$, $max = 24$) and baseline projects (Figure 12) show moderate upper-range duration ($Mdn = 4$ yrs, $IQR = 2-7$ yrs, $max = 22$ yrs). Baseline durations do not differ from trend durations at the median level, but have fewer long-duration (>10-year) projects. Source projects (Figure 13) have slightly shorter durations ($Mdn = 3$ yrs, $IQR = 2-6$ yrs, $max = 14$ yrs), and education projects (Figure 15) have the fourth shortest median duration of two years ($IQR = 2-12$ yrs, $max = 24$ yrs) but relatively high intra-annual sampling ($Mdn = 5.5$ months/yr). Effectiveness monitoring (Figure 16) is the shortest in duration ($Mdn = 1$ yr, $IQR = 1-3.5$ yrs, $max = 5$ yrs). When all parameter types are combined, the median number of analytes measured per year is similar across monitoring objectives. The number of median analytes monitored per year does not differentiate goal types, as the overall median across goal types is consistently two years, with the upper quartile never exceeding four years. Metals parameter types show moderate to high analyte counts, but fewer sampling days per year across all goal types.

About 25% of baseline and source projects exceed the median of 15 sites per year for monitoring. The median number of sites per year is consistent across goal types, ranging from nine to ten, indicating limited differentiation in spatial scale by goal type. The NWM Lakes 1, Bitterroot Sapphire, Clearwater Nut, and Whitefish LI groups are in the 75th percentile for the median number of unique sampling events per year across goal and parameter types. Education projects exhibit the highest median sampling frequency ($Mdn = 10$ days/year), followed by baseline ($Mdn = 8$ days/year), source ($Mdn = 7.25$ days/year), trend ($Mdn = 6$ days/year), and effectiveness ($Mdn = 6$ days/year). The median number of months monitored per year is also highest for the education goal type ($Mdn = 5.5$ months/ys) and lowest for effectiveness

monitoring (*Mdn* = 2 months/yr). Baseline (*Mdn* = 3.5 months/yr), source (*Mdn* = 3.75 months/yr), and trend goal types (*Mdn* = 3 months/yr) cluster in the mid-range. Effectiveness projects are both the smallest in scale and the shortest in duration.

The strongest differentiation between goal types occurs in the duration of sampling and the number of months monitored per year, which show the largest between-group separation in both central tendency and distributional spread relative to other metrics. For the duration of sampling, the difference between the highest and lowest medians for the five goal types is three years (1 yr - 4 yrs). Additionally, the upper-tail dispersion differs since trend projects extend to 24 years, whereas effectiveness monitoring projects do not exceed five years. The IQR for trend duration (*1-9 yrs; width* = 8 years) is substantially wider than for effectiveness monitoring (*IQR* = 1-3.5 years; *width* = 2.5 years), indicating greater variability and long-term commitment within projects with a trend goal type. For months monitored per year, median values range from two months (effectiveness monitoring) to 5.5 months (education). This represents a 175% increase in median annual monitoring between the lowest and highest goal type categories. Highlighted projects that are a mix of river and lake monitoring efforts, including Big Hole WC, Elkhorn, Gallatin TF Nut, NWM Lakes 1, Whitefish LI, Stillwater, Rosebud, Sun, Madison, and Bitterroot Sapphire. These projects frequently exceed 75th percentile thresholds for multiple metrics, showing higher monitoring intensity relative to category medians.

DISCUSSION

The main contribution of this study lies in its approach to shifting the evaluation of citizen science water quality monitoring from data credibility to data adequacy relative to monitoring intent. Most volunteer water monitoring studies focus on data accuracy and usability

(Albus et al. 2019). Studies on the reliability of citizen science have been crucial in positioning citizen science as a legitimate form of data acquisition and in helping address the research barriers described by Burgess et al. (2017). Similar syntheses by Stepenuck and Genskow (2018) and Shao and Bishop (2025) have characterized the breadth, participation, methods, motivations, and parameters of interest in volunteer water quality monitoring. Yet no available framework links volunteer-articulated goals to quantitative dataset characteristics to assess whether monitoring efforts are well poised to address project questions. Setting specific and measurable goals leads to higher project performance by focusing attention and sustaining efforts toward desired outcomes (Klassen et al. 2025). Establishing well-defined goals and objectives is critical because vague or undefined goals often lead to projects that fail to produce actionable results or sustained engagement (Klassen et al. 2025). Clear goals, paired with measurable objectives, provide a mechanism for accountability and align projects with their intended outcomes (Diamente et al. 2025). This study addresses the importance of sustainable project goals and begins to fill the gap in understanding how to integrate volunteer monitoring goals and objectives into program evaluation using qualitative coding alongside WQP data.

Volunteer Water Quality Monitoring Data Across Montana

The volunteer water quality monitoring locations are concentrated in areas of the state with the largest populations (Bryson and Johnson 2022), and 56.5% of monitoring locations fall west of the continental divide (Figure 5). The spread of sampling locations aligns with Sophia Bryson and Blair Johnson (2022) ArcGIS water quality synthesis, in that water quality monitoring occurs more frequently in the western half of the state. MDEQ reports that since 2019, volunteer monitoring efforts in Montana have accounted for up to 25% of DEQ water

quality monitoring (Montana DEQ 2022b), spanning across 30.4% of Montana's HUC-8 (Figure 5). It was expected that volunteer groups would sample areas near support networks, as prior research has shown that institutional support and localized sampling networks are strong predictors of volunteer program success (San Llorente Capdevila et al. 2020, Bonney et al. 2023). Montana's western region includes major universities and multiple watershed organizations, many of which support volunteer training and data management (Flathead Lake Bio Station 2025, Montana State University n.d.). The distribution of monitoring locations also reflects broader citizen science dynamics, in which monitoring tends to occur in socially valued or accessible locations (Devictor et al. 2010, Conrad and Hilchey 2011). Montana is the fourth-largest state and the second-least populous; therefore, I expected that volunteer groups would monitor locations with few logistical challenges and maintain a larger presence in watersheds with higher populations and greater community resources.

The first volunteer water quality monitoring project began in 2001, and by 2024, the number of projects had increased to more than 20 (Figure 6). The increase in volunteer water quality monitoring projects in Montana mirrors national and global growth in citizen science participation (Silvertown 2009, Shao and Bishop 2025). The number of publications on volunteer water quality monitoring began to increase after 2015 (Shao and Bishop 2025), a trend similarly observed in Montana, where the number of active projects has steadily increased since 2015. In 2009, MDEQ launched its lab Volunteer Monitoring Laboratory Analysis Support Program grant (A. Ebert, personal communication, April 13, 2026), and in Figure 7, 2008-2009 is the first period in which the number of active projects begins to climb. Similarly, MMW began funding volunteer water quality monitoring lab analyses and field gear in 2021 (Flathead Lake Bio

Station 2026c). Since 2021, twelve programs have begun, and ten are still active (Figure 7). The ongoing growth in monitoring efforts suggests that Montana's coordinated institutional support through MDEQ, MSUEWQ, and MMW has contributed to the expansion of programs. Funding cycles, internal transitions, shifting project goals, and volunteer capacity are repeatedly identified as central to program success (Conrad and Hilchey 2011, San Llorente Capdevila et al. 2020), and the turnover and data gaps observed in Figure 6 demonstrate how volunteer monitoring remains sensitive to these factors. The synthesis of volunteer monitoring activities (Figure 6 & Figure 7) suggests that Montana is fostering a robust volunteer monitoring landscape with institutional support and state funding; yet, program persistence can remain variable.

Physical and nutrient parameters dominated the WQX data and SAP parameters of interest. Water quality characteristics, including water temperature, specific conductance, dissolved oxygen, barometric pressure, turbidity, total nitrogen, and total phosphorus, accounted for the majority of water quality observations (Figure 9). The prevalence of these characteristics reflects their feasibility and relevance in volunteer monitoring. Nutrient and temperature data are important pieces for understanding water body impairment listings and eutrophication concerns across Montana, and their frequency in volunteer monitoring data suggests alignment with statewide water quality priorities (MDEQ 2025 p. 42). Over half of the volunteer water quality programs in Montana have monitoring objectives targeting the impacts of land use, with the majority focusing on nutrient and sediment loading (Figure 10 & Figure 11). However, metals and biological parameters are also of statewide concern (MDEQ 2025 p. 42), particularly in the Columbia River Basin (Eldridge and Hornberger 2023), where 52.6% of the monitoring sites are located. Yet, metals and biological data are less common in the volunteer datasets (Figure 8),

indicating that monitoring efforts remain focused on parameters that are either field-measurable or associated with active land uses. The Freshwater Research Lab at the Flathead Lake Bio Station offers discounted rates for analyzing total nitrogen and phosphorus together, making it more economical for groups to get both analytes measured in a single sample. (Flathead Lake Bio Station and Monitoring Montana Waters 2025). The lab's pricing incentive could influence the prevalence of these characteristics in volunteer datasets.

Insights from the Qualitative Coding of Group Sampling Analysis Plans

After completing the quality coding of project SAPs, it was clear that summer-baseflow sampling dominated across all goal types. In all trend analysis projects, 96.7% of baseline and 89.5% of source assessment projects collected data primarily during the summer. Montana has a snowmelt-dominated hydrological regime (Whitlock et al. 2017), and meaningful variation in water quality often occurs outside the summer sampling window. Spring snowmelt in the Rocky Mountains drives peak discharge and thus pulses of mobilized sediment and nutrients in watersheds (Pederson et al. 2011). Event-driven loading during runoff can substantially exceed baseflow concentrations, especially in agriculturally influenced watersheds (Correll et al. 1999, Berka et al. 2001). By focusing sampling in the summer low-flow periods, groups can gain insight into thermal stress, dissolved oxygen limitations, and impacts on recreation. Summer baseflows are also sensitive to drought conditions, with recorded declines in warm-season streamflow in the northern Rockies (Leppi et al. 2012). Therefore, summer monitoring is relevant for detecting climate-related stressors and chronic impairment. However, this timing may underrepresent hydrologically driven nutrient export and sediment-associated pollution transport during runoff events (Tate et al. 1999). Peak spring runoff is estimated to be shifting

earlier in the year due to climate change (Pederson et al. 2011, Xu et al. 2021), so groups that do not monitor between March and May could miss peak discharge events that were previously arriving in late May or June. Both Thornhill et al. (2016) and Hadj-Hammou et al. (2017) remark on how citizen scientists primarily sample in the spring and summer, introducing temporal biases discussed in Kosmala et al. (2016) and Rosário et al. (2025). These findings support the assumption that summer conditions are most pleasant for monitoring and that extreme weather limits volunteer participation.

The limited year-round sampling suggests that most volunteer datasets are better suited for baseline assessments and education efforts than for comprehensive source attribution or effectiveness monitoring. If groups consistently monitor during the summer, then trend analysis for summer conditions alone could be undertaken. For source assessment goals, event-based sampling during spring runoff or storm events may be necessary to capture peak loading (Jordan et al. 2007, Jones et al. 2011). The current emphasis on summer monitoring is marginally sufficient for detecting seasonal ecological stress (e.g., Dey et al. 2021) but insufficient to fully characterize nutrient dynamics and the driving factors of water quantity across a watershed (Tate et al. 1999). The bias introduced by monitoring over a single season can affect conclusions about a water body's condition (Pratt and Chang 2012), underscoring the importance of aligning sampling timing with the specific process, stressor, pollutant, or land use under investigation. Changing climatic regimes are shifting runoff timing and magnitude (Pederson et al. 2011); therefore, volunteer monitoring efforts must bear in mind that climatic change could begin to shift when targeted processes occur during a water year.

Qualitative coding revealed that most groups are interested in condition assessment, nutrients in water bodies, biological pollutants, and land-use impacts. This pattern aligns with findings that volunteer monitoring programs frequently prioritize community-relevant concerns and visible water quality issues (McCauley 2017, Ramírez et al. 2023). Additionally, the Montana Watercourse Volunteer Monitoring Guidebook, available on the MMW, MSUEWQ, and MDEQ websites, informs groups that “monitoring physical characteristics of the river will require the least amount of training, equipment, and time (Barnes 2015 p. 35).” The guide goes on to list a set of approachable physical, chemical, and biological monitoring parameters, including dissolved oxygen, pH, water clarity, and temperature (Barnes 2015). As expected, most groups measured the suggested parameters, and physical measurements accounted for the majority of the data counts (Figure 8 & Figure 9). The review of citizen science studies performed by Ramírez et al. (2023) likewise found that the top parameters monitored by citizen scientists were clarity, pH, dissolved oxygen, nitrites, and temperature. Moreover, Stepenuck and Genskow (2018) and Ramírez et al. (2023) found that more studies focus on rivers than lakes, a result echoed in Figure 10D.

Source assessment goals were heavily linked to nutrient parameters and land-use contexts (Figure 11). The co-occurrence of these goals and objective elements suggests that monitoring projects are actively trying to connect nutrient concentrations to agricultural or developed land uses to detect “hotspots” of pollutant loading, which aligns with well-documented nutrient land-use interests across citizen science efforts (Conrad and Hilchey 2011, Ramírez et al. 2023). Effectiveness monitoring goals were rare across volunteer monitoring projects (Figure 10), despite the prevalence of restoration activities across Montana (MDEQ 2025). The limited

number of effective monitoring goals was not surprising, as in the MMSG (Makarowski and Sigler 2020) effectiveness monitoring is presented as the most challenging and demanding goal type for groups to address. The order or approachability for the goal types suggested by the MMSG and Dr. Sigler of MSUEWQ is baseline, source assessment, trend analysis, and effectiveness monitoring (Makarowski and Sigler 2020). The MMSG ranking matches the order of the most common goal types across projects (Figure 10), excluding the more flexible and independent education and outreach goal type.

Education and outreach goals were associated with shorter sampling durations but higher median months monitored per year (Figure 15), suggesting that engagement-focused programs emphasize consistent participation rather than long-term continuity. However, it was surprising that fewer than half of the monitoring projects listed education and outreach as a goal or as a monitoring context in their SAPs. Hands-on involvement in citizen science fosters environmental stewardship and learning (Bennett et al. 2018, Peter et al. 2021, Spellman et al. 2021), and gaining knowledge about science, nature, and local watersheds is among the leading reasons volunteers report engaging in water quality monitoring in Montana (Warrenfeltz et al. 2025). Additionally, social connection and spending time in nature have been identified as powerful motivators for volunteers (Wright et al. 2015, Church et al. 2019, Warrenfeltz et al. 2025). Every monitoring project likely aims to achieve some degree of community engagement, education, or outreach. Nonetheless, project educational intentions may have been deliberately excluded from SAPs. SAPs are technical documents focused on how a project will collect and handle data (Barnes 2015 p. 97, Sigler and Wall 2017), and because of this, groups may feel discouraged from including less technical education and outreach goals. Communication between MSUEWQ

and MDEQ practitioners also revealed that MDEQ has actively discouraged some groups from including education and outreach goals because of the ambiguity about how they would meaningfully monitor progress towards those goals (W.A. Sigler, personal communication, April 13, 2026). Yet, Diamente et al. (2025) emphasize that choices about what to measure and publicly articulate shape overall program priorities. Meaning that education and outreach may be undervalued if they are not explicitly stated as goals in planning documents, even if the community-based nature of volunteer monitoring inherently advances education and outreach.

The Goal and Objectives Evaluation Framework

The evaluation framework demonstrated that certain projects consistently fall within the 75th percentile across multiple monitoring metrics and are therefore comparatively well positioned relative to peers to achieve their goals and objectives. The evaluation framework does not rank programs; instead, it contextualizes them within peer distributions to offer a lens for identifying structural strengths and limitations. Projects such as Sun, Madison, Whitefish LI, Bitterroot Sapphire, Clearwater Nut, and NWM Lakes 1 frequently appear at or near the upper quartile for monitoring duration, number of sites monitored per year, unique sampling events per year, or months monitored per year (Figure 12 - Figure 15). These projects combine extended temporal commitment with either broad spatial coverage or relatively high sampling intensity. Notably, the Clearwater Lakes and NWM Lakes 1 groups included education/outreach objectives alongside robust spatial and temporal coverage (Figure 15). In general, the common thread among well-positioned projects is a long-term commitment to monitoring, sampling at 15 or more sites, and sampling for more than three months per year. A review of 72 peer-reviewed studies on volunteer quality monitoring revealed a median of 24 monitoring sites, with 38% of

projects sampling monthly (Ramírez et al. 2023). The majority of projects discussed in Ramírez et al. (2023) were monitored for 1-5 years, mirroring the median monitoring duration of three years for Montana volunteer water quality projects. Yet, most monitoring durations may fall within this 1-5-year range because 23 of the 43 projects in Montana began and remained active between 2019 and 2024. Since not all 2025 data are included in this analysis, I could not account for how many projects operated past 2024. Knowing the number of active projects in 2025 could change the monitoring duration for over half of the projects in the 1-5-year monitoring period and raise the median monitoring duration across all volunteer monitoring projects in Montana.

Taking a closer look at cross-goal comparisons, nearly all the top projects included at least one baseline goal with a condition monitoring context. The top baseline projects are likely well poised to have the data needed to establish current water quality conditions for the time of year at which they were sampled. Not only is large-scale baseline volunteer monitoring shown to have high agreement with professional data (Albus et al. 2019), but it is also the easiest to approach for groups and requires little upfront investment in long-term planning (Barnes 2015, Makarowski and Sigler 2020). Sustaining participation and building a core group of experienced volunteers is essential to the success of long-term monitoring (San Llorente Capdevila et al. 2020, Lopez et al. 2024). However, building a core group of volunteers with sustained commitment is time-consuming and requires more upfront planning from monitoring leaders (McKinley et al. 2017, Lopez et al. 2024, Klassen et al. 2025). Groups with less sustainable volunteer pools are better positioned to measure current conditions rather than multi-year trends, which require consistent data collection and management and a communicated commitment to the future (Hirsch et al. 1982, Tate et al. 1999, Wellen et al. 2020).

Volunteer projects in Montana with a trend analysis goal were characterized by longer median sampling durations but not higher within-year sampling frequency (Figure 14). However, only the most recent SAPs for monitoring groups were analyzed. Therefore, some projects may have begun with baseline monitoring and set explicit trend goals only after sufficient years of data had been collected. This project evolution likely explains why projects with trend goals typically have longer monitoring durations. Analytically, it makes sense that trend objectives emerge after a sufficiently large baseline database capable of detecting change accumulates (Hirsch et al. 1982). In other words, the apparent success of trend projects at reaching longer median monitoring durations may reflect the fact that those groups have operated long enough to feel confident stating trend analysis as a goal. Additionally, trend analysis projects have fewer median unique sampling events per year than baseline, source assessment, and education goal types. Deutsch and Ruiz-Córdova (2015) found over the course of a 20-year volunteer monitoring project in Alabama that projects initially had robust engagement that eventually stabilized at a lower level of activity and monitoring. The challenges in sustaining long-term monitoring efforts described by Deutsch and Ruiz-Córdova (2015), along with the lower sampling frequencies observed in projects with a trend goal, imply that the most sustainable programs are those with realistic scopes and clear, achievable goals. Groups focusing on fewer sites or reduced sampling intensity per year can likely sustain their efforts longer than those attempting maximal coverage. Groups also need adaptable goals (Diamante et al. 2025) because a group might initially set broad, ambitious objectives but then consciously pare back to a core set of sites to maintain volunteer interest. Education projects exhibit the highest within-year temporal coverage. (Figure 15). It is not surprising that projects prioritizing education can sustain

higher within-year volunteer interest, as education is widely recognized as integral to maintaining active engagement in citizen science (Conrad and Hilchey 2011, Wright et al. 2015, McCauley 2017, Church et al. 2019, Warrenfeltz et al. 2025).

One pattern that emerged between goal types and monitoring metrics, shown in Figures 12-16, is that the source assessment and effectiveness monitoring projects frequently lacked the intensive or event-driven detail needed to fully inform these goal types. For example, Big Hole WC Elkhorn pursued source assessment goals without high-frequency (Figure 13), event-based, or inter-seasonal sampling that might capture runoff-driven nutrient pulses (Tate et al. 1999, Jones et al. 2011, LaPorte and Hurley 2021). Effectiveness monitoring is also an ambitious goal for volunteers to undertake (Tate et al. 1999), and the two effectiveness monitoring projects have short monitoring durations and limited monitoring scopes (Figure 16). Thus, preventing Gallatin TF VM and Broadwater Deep Crk from robustly evaluating their management outcomes unless monitoring is shown to be ongoing. As both projects are no longer active (Figure 6), their monitoring designs were not aligned with their intended monitoring objectives. The lack of strong differentiation in the evaluation frameworks for monitoring metrics by goal type indicates that monitoring design choices are influenced as much by volunteer capacity and local monitoring context as by formal goals and expressed monitoring intent. Ultimately, the cross-goal evaluations suggest that while many projects are appropriately structured for baseline characterization, education/outreach, and seasonal trend analysis, fewer are designed with the temporal or systematic depth needed for robust source attribution and effectiveness monitoring.

Study Limitations

The results of this study are exploratory and descriptive, and there are multiple limitations. First, not all volunteer water quality data in Montana have been uploaded to the WQP, and some groups with data in the portal have incomplete records (Jawer 2026). Groups like the Sun show data gaps in Figure 6, but some of these data gaps are erroneous, and efforts are underway to upload missing data under the correct Project IDs (Jawer 2026). Therefore, the projects detailed in this analysis provide a snapshot of the volunteer water quality landscape in Montana rather than a comprehensive inventory of all monitoring efforts. My analysis is biased towards volunteer projects that have their data uploaded correctly to WQP, a vital but technically complicated step that groups undertake. A second limitation of the study is that the evaluation framework treated all 43 projects uniformly for comparability (Figure 4), despite the groups being heterogeneous in their monitoring designs, organizational structures, volunteer capacities, funding, and analytical methods. Project diversity is one of the strengths of citizen science (San Llorente Capdevila et al. 2020), but it also complicates attempts to evaluate projects using a single framework. Applying the same evaluation metrics used in Figures 12-16 oversimplifies monitoring strategies. It fails to account for project-specific metrics needed to determine whether a dataset is sufficient to meet monitoring goals. For example, the spatial consistency of monitoring sites was not evaluated, but detecting trends in water quality requires repeated sampling at the same locations over time (Hirsch et al. 1982, Tate et al. 1999).

A methodological limitation is that I relied heavily on project SAPs, resulting in monitoring contexts being inferred solely from SAP project descriptions when monitoring intentions were not clearly stated in goals or objectives. SAPs provide a structured description of

monitoring efforts, but they do not always capture how a project has evolved. I expect volunteer programs to shift their monitoring goals as groups gain experience, develop partnerships, access additional educational resources, or respond to emerging environmental concerns. Groups like Big Hole RF and Gallatin LWQD included a table in their SAPs listing annual changes to their monitoring designs, but most SAPs did not include this information. Since I relied on the most recent version of each project's SAP in this study, it was not always possible to reconstruct or account for changing monitoring goals and intentions for each project. It is also worth noting that I was the only qualitative coder, so my biases are not controlled for in the dataset (O'Connor and Joffe 2020), and some code choices could be subjective. Although I established clear coding rules in Tables 2-4, qualitative coding inevitably involves some interpretive judgment.

Another limitation of this study is that the dataset does not fully meet the assumptions required for formal statistical independence. The unit of analysis in the evaluation framework is the combination of Project IDs by parameter type, grouped by each goal type. However, a single project can have multiple parameter combinations, each of which can be further assigned to multiple goal types. Therefore, evaluation framework observations are not strictly independent across goal categories, violating assumptions typically required for statistical hypothesis testing (Ramsey and Schafer 2013 chap. 3). Even using nonparametric tests to determine whether metrics differ by goal type or parameter is likely to lack statistical and inferential power due to small sample sizes and difficulty of translating practical interpretations (Ramsey and Schafer 2013 chap. 4, Serdar et al. 2021). Thus, the evaluation framework was limited in its statistical applications and instead had to focus on identifying patterns in how volunteer programs structure their monitoring efforts relative to their stated objectives.

The most important limitation of this study is the complexity of determining whether a dataset can address a monitoring goal or question. Assessing the adequacy of monitoring data is nuanced because different research questions require different combinations of spatial coverage, temporal frequency, and monitoring duration (Wellen et al. 2020, McDowell et al. 2024). For example, identifying nutrient sources may require high-frequency sampling during storm events or snowmelt periods, whereas detecting long-term water quality trends may require decades of consistent sampling at the same locations. Ultimately, there is no universal threshold at which a dataset becomes “sufficient” for answering a question, as what is “sufficient” is a constantly moving target specific to a monitoring program's objectives. How good is good enough is also an exercise of how much uncertainty is tolerable in a dataset.

Future Research Directions

This study represents an initial step towards developing a framework for evaluating the alignment between volunteer monitoring data and their stated goals and objectives. Multiple opportunities exist to expand this framework in future studies. One opportunity would be to incorporate additional monitoring metrics beyond those included in Figures 12-16. The current framework focuses on dataset characteristics, including data counts, monitoring duration, the number of analytes monitored, the number of monitoring sites, and within-year sampling frequency. While these metrics are useful for describing volunteer monitoring efforts, other factors are equally important for determining whether datasets are well poised to address specific monitoring questions. Expanding the framework to consider the spatial consistency of monitoring locations, the timing of sampling relative to hydrologic events, analytical precision,

and quality assurance procedures would provide a more comprehensive assessment of monitoring programs.

In future studies, direct engagement with volunteer monitoring groups through surveys or interviews could help illuminate how project goals have evolved. Interviews could confirm whether groups started with baseline goals and later adopt trend goals, and how choices are made to balance monitoring ambition with long-term project sustainability. Interviewing project leaders would provide valuable insights into how groups balance their resources and into the strategies that enable long-term program sustainability. Conversations with project leaders could also clarify whether educational goals are intentionally omitted from SAPs, and how education and outreach influence volunteer motivations. Water quality practitioners could also provide guidelines to project leaders on how to quantitatively monitor progress toward education and outreach goals by tracking metrics like attendance at meetings and website visits. Lastly, requiring groups to include a revision history table in their SAP would also set up future research efforts for success and eliminate guesswork on how projects have developed.

Future research should explore how volunteer monitoring programs contribute to community engagement and environmental stewardship. While few SAPs explicitly listed education and outreach as a monitoring goal, hands-on participation in environmental monitoring has been widely shown to strengthen public understanding of scientific processes and foster local stewardship of natural resources (McKinley et al. 2017, Bennett et al. 2018). Citizen science initiatives play an important role in connecting communities to environmental issues, increasing scientific literacy, and encouraging collaborative watershed management. Evaluating these

broader social outcomes alongside scientific monitoring objectives would provide a more complete understanding of the value of volunteer water quality monitoring programs.

To strengthen the study's qualitative aspects, future work should include at least two coders for SAP goal analysis to reduce bias (O'Connor and Joffe 2020). Establishing inter-coder reliability would lend robustness to the results about common group goal types and objective elements. Once the qualitative coding is fortified, the evaluation framework could be extended into a more comprehensive decision-support tool for future monitoring programs. Rather than evaluating monitoring programs relative to their peer's existing WQP datasets, future work could establish goal-specific thresholds that define the minimum data requirements for addressing types of questions within a defined confidence interval. Developing such thresholds would likely require collaboration between scientists, monitoring coordinators, and regulatory agencies, as well as require statistical power analyses to determine the sampling intensity needed to detect meaningful environmental changes.

Future volunteer water quality monitoring efforts could enhance their designs by focusing on sites with robust, existing high-frequency datasets and by using statistical methods discussed by Dunnette (1980) to estimate the number of samples needed at comparable sites to detect trends and distinguish signal from noise. Investigating the use of statistical simulations, such as Monte Carlo simulations, could also enable the evaluation tool developed in this study to advance from a descriptive assessment to a predictive decision-support system. Monte Carlo simulations can be used to repeatedly generate synthetic monitoring datasets under different sampling scenarios and evaluate how likely those datasets are to detect meaningful environmental signals (Grayman et al. 2012). Instead of comparing existing datasets, future

evaluation frameworks could assess hypothetical monitoring designs and estimate the probability that they will successfully detect environmental changes.

CONCLUSION

In this study, I developed and applied a framework to evaluate the extent of volunteer water quality data across Montana and whether those data are well poised to address the monitoring goals and objectives articulated in monitoring groups' SAPs. I found that volunteer monitoring efforts have expanded since 2001 and focus on characterizing physical and nutrient parameters across the state. Baseline and trend analysis goals are the most common among the 43 evaluated volunteer water quality monitoring projects. Surprisingly, there is a low frequency of education and outreach goals, despite their being central to sustaining long-term volunteer efforts. The strongest differentiation among goal types occurs in the temporal design, with monitoring duration and the number of months sampled per year setting groups apart rather than the differences in monitoring spatial extent or analyte diversity. Multiple monitoring efforts are well positioned compared to their peers to address their goals; however, I identified several mismatches between the monitoring design and intent for groups with effectiveness monitoring and source assessment goal types.

Reframing how citizen science is evaluated recognizes that volunteer monitoring projects are independently directed scientific efforts whose success should be evaluated against their internally determined objectives, not only external state or agency standards. Therefore, the goal and objective evaluation framework contributes to the broader literature by offering a replicable method for aligning monitoring intention and quantitative data criteria in citizen science monitoring contexts. By integrating additional quantitative monitoring metrics, gaining

qualitative program insights from interviews, encouraging groups to state their educational and outreach goals explicitly, and estimating data requirements using statistical models, future studies could develop tools to help volunteer groups design monitoring programs that position groups to address their goals and meet project objectives.

REFERENCES CITED

- Ahmed, S. K., R. A. Mohammed, A. J. Nashwan, R. H. Ibrahim, A. Q. Abdalla, B. M. M. Ameen, and R. M. Khedhir. 2025. Using thematic analysis in qualitative research. *Journal of Medicine, Surgery, and Public Health* 6:100198.
- Albus, K., R. Thompson, and F. Mitchell. 2019. Usability of Existing Volunteer Water Monitoring Data: What Can the Literature Tell Us? *Citizen Science: Theory and Practice*.
- Amongero, K. 2026. *LRES575-MTVolWaterMonitoring*. GitHub repository. (Available from: <https://github.com/KennaWhite/LRES575-MTVolWaterMonitoring>)
- Barnes, S. (Ed.). 2015. *Montana Watercourse Volunteer Water Monitoring Guidebook*. , 2nd edition. Montana Watercourse of the Montana Water Center.
- Bennett, N. J., T. S. Whitty, E. Finkbeiner, J. Pittman, H. Bassett, S. Gelcich, and E. H. Allison. 2018. Environmental Stewardship: A Conceptual Review and Analytical Framework. *Environmental Management* 61:597–614.
- Berka, C., H. Schreier, and K. Hall. 2001. Linking Water Quality with Agricultural Intensification in a Rural Watershed. *Water, Air, and Soil Pollution* 127:389–401.
- Bingham, A. J. 2023. From Data Management to Actionable Findings: A Five-Phase Process of Qualitative Data Analysis. *International Journal of Qualitative Methods* 22:16094069231183620.
- Bonney, P. R., B. D. Hansen, and C. Baldwin. 2023. Citizen Science and Natural Resource Management: A Social Network Analysis of Two Community-Based Water Monitoring Programs. *Society & Natural Resources* 36:600–621.
- van den Brand, T. 2025. ggh4x: Hacks for “ggplot2.” R package version 0.3.1.
- Braun, V., and V. Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3:77–101.
- Bryson, S., and B. Johnson. 2022. Spatial Patterns in Water Quality Portal Data. Spatial patterns in Water Quality Portal data. ArcGIS Online Dashboard, Duke University’s Nicholas School of the Environment and blmsuInternet of Water. (Available from: <https://www.arcgis.com/apps/dashboards/0dbe111a2c1542a4a1ff01387b037d13>)
- Burgess, H. K., L. B. DeBey, H. E. Froehlich, N. Schmidt, E. J. Theobald, A. K. Ettinger, J. HilleRisLambers, J. Tewksbury, and J. K. Parrish. 2017. The science of citizen science: Exploring barriers to use as a primary research tool. *Biological Conservation* 208:113–120.

- Church, S. P., L. B. Payne, S. Peel, and L. S. Prokopy. 2019. Beyond water data: Benefits to volunteers and to local water from a citizen science program. *Journal of Environmental Planning and Management* 62:306–326.
- Conrad, C. C., and K. G. Hilchey. 2011. A review of citizen science and community-based environmental monitoring: Issues and opportunities. *Environmental Monitoring and Assessment* 176:273–291.
- Correll, D. L., T. E. Jordan, and D. E. Weller. 1999. Transport of nitrogen and phosphorus from rhode river watersheds during storm events. *Water Resources Research* 35:2513–2521.
- Das, T., D. W. Pierce, D. R. Cayan, J. A. Vano, and D. P. Lettenmaier. 2011. The importance of warm season warming to western U.S. streamflow changes. *Geophysical Research Letters* 38.
- David Wilkins. 2023. treemapify: Draw Treemaps in “ggplot2.” R package version 2.5.6.
- DeCicco, L., R. Hirsch, D. Lorenz, J. Read, J. Walker, L. Platt, D. Watkins, D. Blodgett, M. Johnson, A. Krall, L. Stanish, J. Zemmels, E. Hinman, and M. Mahoney. 2025. dataRetrieval: R packages for discovering and retrieving water data available from U.S. federal hydrologic web services. U.S. Geological Survey, Reston, VA.
- Deutsch, W. G., and S. S. Ruiz-Córdova. 2015. Trends, challenges, and responses of a 20-year, volunteer water monitoring program in Alabama. *Ecology and Society* 20.
- Devictor, V., R. J. Whittaker, and C. Beltrame. 2010. Beyond scarcity: Citizen science programmes as useful tools for conservation biogeography. *Diversity and Distributions* 16:354–362.
- Dey, S., S. Botta, R. Kallam, R. Angadala, and J. Andugala. 2021. Seasonal variation in water quality parameters of Gudlavalleru Engineering College pond. *Current Research in Green and Sustainable Chemistry* 4:100058.
- Diamante, C., L. Gitman, A. Cramer, M. Brent, and Paul Holdredge. 2025, June 25. Sustainability Goal Setting: A New Chapter. BSR. BSR. (Available from: <https://www.bsr.org/en/sustainability-insights/insights-plus/sustainability-goal-setting-a-new-chapter>)
- Dunnette, D. A. 1980. Sampling Frequency Optimization Using a Water Quality Index. *Journal (Water Pollution Control Federation)* 52:2807–2811.
- Eldridge, S. L. C., and M. I. Hornberger. 2023. Spatiotemporal variations in copper, arsenic, cadmium, and zinc concentrations in surface water, fine-grained bed sediment, and aquatic macroinvertebrates in the upper Clark Fork Basin, western Montana—A 20-year synthesis, 1996–2016. Scientific Investigations Report. 2023–5070, U.S. Geological Survey. (Available from: <https://pubs.usgs.gov/publication/sir20235070/full>)

- Flathead Lake Bio Station. 2025, September 10. Monitoring Montana Waters (MMW). University of Montana. (Available from: <https://storymaps.arcgis.com/stories/36421da96e6343db9ed8820edd03ad55>)
- Flathead Lake Bio Station. 2026a. Flathead Lake Facts. Flathead Lake. (Available from: <https://flbs.umt.edu/flathead-lake/flathead-lake-facts/>)
- Flathead Lake Bio Station. 2026b. Funding for Laboratory Analyses and Gear. Monitoring Montana Waters Funding. Flathead Lake Bio Station. (Available from: <https://flbs.umt.edu/monitoring/mmw/funding/>)
- Flathead Lake Bio Station. 2026c. Funded Groups. Funded Groups. Flathead Lake Bio Station. (Available from: <https://flbs.umt.edu/monitoring/mmw/funded-groups/>)
- Flathead Lake Bio Station and Monitoring Montana Waters. 2025. MMW Sample Suite and Price Sheet 2025. Word Document, Freshwater Research Lab, Polson, MT.
- Gardner, K., Gabriella Matzer, and Darrin Kron. 2024. Sampling and Analysis Plan: Nutrient and Algae Monitoring in the Upper Gallatin Watershed. Gallatin River Task Force and the Montana Department of Environmental Quality, Monitoring and Assessment Section.
- Gardner, K., J. Murray, and A. DeGroot. 2018, March 20. Gallatin River Task Force Community Water Quality Monitoring Program Sampling and Analysis Plan. Gallatin River Task Force.
- Grayman, W. M., V. L. Speight, and J. G. Uber. 2012. Using Monte-Carlo Simulation to Evaluate Alternative Water Quality Sampling Plans. 1–11.
- Hadj-Hammou, J., S. Loisele, D. Ophof, and I. Thornhill. 2017. Getting the full picture: Assessing the complementarity of citizen science and agency monitoring data. *PLOS ONE* 12:e0188507.
- Hirsch, R. M., J. R. Slack, and R. A. Smith. 1982. Techniques of Trend Analysis for Monthly Water Quality Data. *Water Resources Research* 18:107–121.
- Jawer, G. 2026, March 20. MissingDataTracker. Excel Spreadsheet, Montana State University OneDrive.
- Jones, A. S., D. K. Stevens, J. S. Horsburgh, and N. O. Mesner. 2011. Surrogate Measures for Providing High Frequency Estimates of Total Suspended Solids and Total Phosphorus Concentrations. *JAWRA Journal of the American Water Resources Association* 47:239–253.
- Jordan, P., A. Arnscheidt, H. McGrogan, and S. McCormick. 2007. Characterising phosphorus transfers in rural catchments using a continuous bank-side analyser. *Hydrology and Earth System Sciences* 11:372–381.

- Klassen, M., C. B. Dobni, and N. T. Sheehan. 2025. Measure what matters: A blueprint for a sustainability culture diagnostic. *Business Horizons* 68:793–805.
- Kosmala, M., A. Wiggins, A. Swanson, and B. Simmons. 2016. Assessing data quality in citizen science. *Frontiers in Ecology and the Environment* 14:551–560.
- Kreiner, H., W. A. Sigler, K. Kaylor, and A. Bratton. 2015. Deep Creek Watershed Sampling and Analysis Plan. Broadwater Conservation District.
- LaPorte, B., and P. Hurley. 2021. Big Hole Watershed Committee Elkhorn Mine & Mill Restoration Project: Water Quality Monitoring Sampling and Analysis Plan. Big Hole Watershed Committee, Montana DEQ.
- Leppi, J., T. DeLuca, S. Harrar, and S. W. Running. 2012. Impacts of climate change on August stream discharge in the Central-Rocky Mountains. *Climatic Change* 112:997–1014.
- Lopez, L. S., A. Basu, K. Blagrove, G. Bove, K. Stewart, D. Bazely, and S. Sharma. 2024. Establishing a long-term citizen science project? Lessons learned from the Community Lake Ice Collaboration spanning over 30 yr and 1000 lakes. *Limnology and Oceanography Letters* 9:99–111.
- Lyons, J., B. Weigel, L. K. Paine, and D. J. Undersander. 2000. Influence of intensive rotational grazing on bank erosion, fish habitat quality, and fish communities in southwestern Wisconsin trout streams. *Journal of Soil and Water Conservation* 55:271–276.
- Makarowski, K., and W. A. Sigler. 2020. Water Resource Monitoring Methods Selection Guide. Version 1.0., Montana Department of Environmental Quality & Montana State University Extension Water Quality.
- McCauley, J. 2017. “We Can Help the People Who Can Make a Difference”: Motivations, Goals, and Outcomes in Volunteer Water Quality Monitoring. *Sociological Focus* 50:125–137.
- McDowell, R. W., A. Noble, M. Kittridge, O. Ausseil, C. Doscher, and D. P. Hamilton. 2024. Monitoring to detect changes in water quality to meet policy objectives. *Scientific Reports* 14:1914.
- McKinley, D. C., A. J. Miller-Rushing, H. L. Ballard, R. Bonney, H. Brown, S. C. Cook-Patton, D. M. Evans, R. A. French, J. K. Parrish, T. B. Phillips, S. F. Ryan, L. A. Shanley, J. L. Shirk, K. F. Stepenuck, J. F. Weltzin, A. Wiggins, O. D. Boyle, R. D. Briggs, S. F. Chapin, D. A. Hewitt, P. W. Preuss, and M. A. Soukup. 2017. Citizen science can improve conservation science, natural resource management, and environmental protection. *Biological Conservation* 208:15–28.
- Microsoft Corporation. 2024. Microsoft Excel. Microsoft Corporation.
- Monitoring Montana Waters. 2025. MMW SAP Template. Template, Flathead Lake Bio Station.

- Montana Department of Environmental Quality [MDEQ]. 2025. Montana 2022-2024 Draft Water Quality Integrated Report. Montana Dept. of Environmental Quality, Helena, MT.
- MDEQ. 2007. MUST_READ_for_NUTRIENTS. Unpublished Excel File.
- MDEQ. 2022a. Volunteer Monitoring SAP Template. Template, Montana Department of Environmental Quality.
- MDEQ. 2022b. Volunteer Monitoring Program. Montana Department of Environmental Quality.
- MDEQ. (n.d.). Montana Department of Environmental Quality. Montana.Gov. (Available from: <https://deq.mt.gov/>)
- Montana Lake Book Steering Committee. 2017. Montana Lake Book. , 3rd edition. Whitefish Lake Institute, Whitefish, MT.
- Montana State University. (n.d.). MSU Extension Water Quality. Department of Land Resources and Environmental Sciences. (Available from: <https://waterquality.montana.edu/>)
- Montana Watershed Coordination Council. 2025. MWCC Monitoring Library. MWCC Monitoring Library. (Available from: <https://www.mtwatersheds.org/resources/water-monitoring-resource-library/>)
- National Oceanic and Atmospheric Administration. (n.d.). Climate Division Map: NOAA Physical Sciences Laboratory. Physical Sciences Laboratory. NOAA. (Available from: <https://psl.noaa.gov/data/usclimdivs/data/map.html#Montana>)
- National Water Quality Monitoring Council. 2026, January 27. WQP User Guide. Water Quality Portal User Guide. U.S. Geologic Survey, Environmental Protection Agency. (Available from: https://www.waterqualitydata.us/portal_userguide/)
- Norton, P. A., M. T. Anderson, and J. F. Stamm. 2014. Trends in annual, seasonal, and monthly streamflow characteristics at 227 streamgages in the Missouri River watershed, water years 1960-2011. Scientific Investigations Report. 2014–5053, U.S. Geological Survey. (Available from: <https://pubs.usgs.gov/publication/sir20145053>)
- O'Connor, C., and H. Joffe. 2020. Intercoder Reliability in Qualitative Research: Debates and Practical Guidelines. International Journal of Qualitative Methods 19:1609406919899220.
- Pederson, G. T., S. T. Gray, T. Ault, W. Marsh, D. B. Fagre, A. G. Bunn, C. A. Woodhouse, and L. J. Graumlich. 2011. Climatic Controls on the Snowmelt Hydrology of the Northern Rocky Mountains. Journal of Climate 24:1666–1687.
- Peter, M., T. Diekötter, T. Höffler, and K. Kremer. 2021. Biodiversity citizen science: Outcomes for the participating citizens. People and Nature 3:294–311.

- Pratt, B., and H. Chang. 2012. Effects of land cover, topography, and built structure on seasonal water quality at multiple spatial scales. *Journal of Hazardous Materials* 209–210:48–58.
- R Core Team. 2021. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing.
- Ramírez, S. B., I. van Meerveld, and J. Seibert. 2023. Citizen science approaches for water quality measurements. *Science of The Total Environment* 897:165436.
- Ramsey, F. L., and D. W. Schafer. 2013. *The statistical sleuth: A course in methods of data analysis*. 3rd ed. Brooks/Cole, Cengage Learning, Australia ; Boston.
- Rosário, I. T., P. Tiago, S. Chozas, and C. Capinha. 2025. When do citizen scientists record biodiversity? Non-random temporal patterns of recording effort and associated factors. *People and Nature* 7:860–870.
- San Llorente Capdevila, A., A. Kokimova, S. Sinha Ray, T. Avellán, J. Kim, and S. Kirschke. 2020. Success factors for citizen science projects in water quality monitoring. *Science of The Total Environment* 728:137843.
- Sando, S. K., N. A. Barth, and K. J. Chase. 2025. Peak Streamflow Trends in Montana and Northern Wyoming and Their Relation to Changes in Climate, Water Years 1921–2020. Page 129. chap. G of Ryberg, K.R., comp. Peak streamflow trends and their relation to changes in climate in Illinois, Iowa, Michigan, Minnesota, Missouri, Montana, North Dakota, South Dakota, and Wisconsin, U.S. Geological Survey Scientific Investigations Report 2023–5064. (Available from: <https://doi.org/10.3133/sir20235064G>.)
- Serdar, C. C., M. Cihan, D. Yücel, and M. A. Serdar. 2021. Sample size, power and effect size revisited: Simplified and practical approaches in pre-clinical, clinical and laboratory studies. *Biochemia Medica* 31:010502.
- Shao, G. P., and I. J. Bishop. 2025. Citizen science in river monitoring: A systematic literature review of the whys and hows. *Frontiers in Environmental Science* 13.
- Sharpley, A. N., S. C. Chapra, R. Wedepohl, J. T. Sims, T. C. Daniel, and K. R. Reddy. 1994. Managing Agricultural Phosphorus for Protection of Surface Waters: Issues and Options. *Journal of Environmental Quality* 23:437–451.
- Sigler, W. A., and E. Wall. 2017, April 27. Guidance for Starting a Volunteer Water Monitoring Program. MSU Extension Water Quality.
- Silvertown, J. 2009. A new dawn for citizen science. *Trends in Ecology & Evolution* 24:467–471.

- Spellman, K. V., D. Cost, and C. P. Villano. 2021. Connecting Community and Citizen Science to Stewardship Action Planning Through Scenarios Storytelling. *Frontiers in Ecology and Evolution* 9.
- Stepenuck, K. F., and K. D. Genskow. 2018. Characterizing the Breadth and Depth of Volunteer Water Monitoring Programs in the United States. *Environmental Management* 61:46–57.
- Tate, K., R. Dahlgren, M. Singer, and B. Allen-Diaz. 1999. Timing, frequency of sampling affect accuracy of water-quality monitoring. 53:44–48.
- Thornhill, I., S. Loiselle, K. Lind, and D. Ophof. 2016. The Citizen Science Opportunity for Researchers and Agencies. *BioScience* 66:720–721.
- VERBI Software. 2021. MAXQDA 2022. VERBI Software, Berlin, Germany.
- Warrenfeltz, B., S. Chruch, and W. Sigler. 2025. Volunteer Water Monitoring Survey: 2024 Statewide Report. People Places Water Lab and MSU Extension, Bozeman: Montana State University.
- Water Quality Portal [WQP]. 2021. National Water Quality Monitoring Council, United States Geological Survey (USGS), Environmental Protection Agency (EPA), Washington, D.C. (Available from: <https://doi.org/10.5066/P9QRKUVJ>.)
- Wellen, C., P. Van Cappellen, L. Gospodyn, J. L. Thomas, and M. N. Mohamed. 2020. An analysis of the sample size requirements for acceptable statistical power in water quality monitoring for improvement detection. *Ecological Indicators* 118:106684.
- Whitefish Lake institute. 2023. Northwest Montana Lakes Network 2022 Summary Report. Prepared for Montana Fish, Wildlife and Parks, Whitefish, MT. (Available from: https://nmln.info/wp-content/uploads/2025/04/2024-NMLN-Report_121024.pdf)
- Whitlock, C., W. Cross, B. Maxwell, N. Silverman, and A. A. Wade. 2017. Montana Climate Assessment. Page 318. Montana State University and University of Montana, Montana Institute on Ecosystems, Bozeman and Missoula MT. (Available from: <http://scholarworks.montana.edu/xmlui/handle/1/13584>)
- Wickham, H., M. Averick, J. Bryan, W. Chang, L. McGowan, R. François, G. Grolemund, A. Hayes, L. Henry, J. Hester, M. Kuhn, T. Pedersen, E. Miller, S. Bache, K. Müller, J. Ooms, D. Robinson, D. Seidel, V. Spinu, K. Takahashi, D. Vaughan, C. Wilke, K. Woo, and H. Yutani. 2019. Welcome to the Tidyverse. *Journal of Open Source Software* 4:1686.
- Wright, D. R., L. G. Underhill, M. Keene, and A. T. Knight. 2015. Understanding the Motivations and Satisfactions of Volunteers to Improve the Effectiveness of Citizen Science Programs. *Society & Natural Resources* 28:1013–1029.

Xu, D., V. Y. Ivanov, X. Li, and T. J. Troy. 2021. Peak Runoff Timing Is Linked to Global Warming Trajectories. *Earth's Future* 9:e2021EF002083.