Water Quality Data in Long Island Sound: Stakeholder Needs and Opportunities for Collaborative Data Management, Sharing, and Visualization

Final Report, September 2020

Jamie M.P. Vaudrey[1] and Sarah C. Crosby[2]
[1]Department of Marine Sciences, University of Connecticut, Groton CT 06340
[2]Harbor Watch, Earthplace, Inc., Westport CT 06880

Earthplace

HARBOR WATCH

UCONN
COLLEGE OF LIBERAL ARTS AND SCIENCES
DEPARTMENT OF MARINE SCIENCES

Funding Generously Provided by the Long Island Sound Stewardship Fund

Table of Contents

Acknowledgements

Funding for this work was generously provided by the Long Island Sound Funders Collaborative through the Long Island Sound Stewardship Fund, and administered by the Long Island Community Foundation. We sincerely appreciate their support. The work presented has been part of a larger collaborative effort, and we are grateful to our partners at Save the Sound and Chesapeake Commons, especially Tracy Brown and John Dawes. Thank you to Nikki Spiller and Devan Healy for their assistance with research and data analysis. We would also like to thank the many stakeholders and partners who participated in our surveys and interviews and enabled the success of this effort.

Executive Summary

Long Island Sound is an ecologically diverse and valuable ecosystem, with many different stakeholders engaged in environmental data collection on the Sound and its watershed. As more data are collected, our ability to effectively understand and protect the Sound will be improved. However, this will only be true if those data are effectively managed, stored, and shared with decision makers. Here, a needs assessment of the community of "data generators" (*i.e.,* groups engaged in monitoring) and "data end users" (*i.e.,* managers, decision makers, and groups engaged in data translation) around the Sound was conducted to establish the current status of groups' data-related processes and their potential willingness to engage with a centralized database in the future.

Surveys of monitoring groups show they have a wealth of data and they want to share those data with the public, but many are not currently doing so. Among the groups collecting data around the Sound, we found that 75% of data generators want to share their data with the public (and yet only 22% said that their data are readily available for public download), 68% of data generators would consider using a centralized database, and 49% of groups are not happy with their current process for entering, storing, and sharing their data (out of n = 39). Additionally, we found that most data end users are currently acquiring data through informal, person-to-person networks; as such, potential sources of available data are not being used to their full potential if their existence is not known by a given end user.

An opportunity exists to support management, promote collaboration, and elevate the impact of community-based monitoring data. With 55 different groups currently engaged in monitoring around the Sound, an increased effort to promote data sharing and stewardship will improve the management and protection of the Sound into the future. Here, we present the results of a series of surveys and interviews with key local stakeholders, an assessment of currently available tools for data generators, and provide a summary of recommended next steps based on the goals and needs identified by the community.

Recommendations for developing a database and visualization tools from data generators:

- Groups are happy with data entry in Excel; groups would like to use this familiar format.
- Guidelines on best practices for groups on data management and quality control are needed.
- Support and tools are needed to increase engagement with the EPA Water Quality Exchange (WQX).
- Public data sharing is a priority for groups and not currently easy to do.
- Interest is high in visualization tools (mapping and graphing of individual groups' data), including real-time display options, and social media and ArcGIS compatibility.
- Groups need a clear protocol for metadata development and a list of what information is required for inclusion of their data in the database, in order to meet the needs of different end users.

- Groups expressed concern over resource limitations that are unrelated to technology (*e.g.,* self-reported insufficiencies in staff time and expertise); thus, keeping required effort low will be key for group engagement.
- Need to define the geographic area to be served by this effort.
- All parameters should be supported; the database should allow groups store all types of data.
- Ability to retrieve group-specific data directly from a group's own website is highly desired.
- The database should be easy to search and data should be easily downloadable.
- The database program needs a designated point of contact for questions and support.

Recommendations for developing a database and visualization tools from end users:

- Must have a well-developed search function for finding relevant data.
- Robust documentation of data (i.e., metadata) to allow other users to decide if the data can be used in a given analysis or application.
- Built-in display and analysis tools. Ability to display both regional summaries as well as drill down on specific, local, raw data.
- Interoperable with other databases.
- Make it a "win-win" and give groups something for providing their data, such as reporting and analysis tools. WQX/WQP does not currently "give anything" for the data provided. Whatever is built should be attractive to groups and make them want to submit data.
- Establish a high burden on constraints to uploading data, requiring groups to report out on methods, units, and assorted quality assurance metrics. The data intended for upload does not need an EPA-approved QAPP nor do analyses need to be conducted in an ELAP-certified lab, but it needs to be clear which standards included in these quality assurance frameworks have been met.
- End users want to be able to filter for data that have met certain thresholds or standards for quality of the data. Some specific concerns were expressed about how to filter or assess data already in WQX/WQP.
- Look at user-generated data and assess the easiest path to get all data into a central database at launch; explore whether spreadsheet templates, R code, or another method would facilitate initial entry into the system.
- Inclusion of data collected as part of state MS4 permit reports would be valuable.

Concerns expressed by data generators and end users included:

- Some end users expressed concern about maintaining the ability to present the data on their organizational website and not be exclusively redirected to a third-party website.
- Finding a design that allows others to upload data and make it an area for shared data.
- Water quality data are complex, and concern was expressed about further complexity that would be added through the inclusion of biological data such as zooplankton, macroinvertebrates, eelgrass, and other relevant data types.
- Need to consider how to address continuous versus discrete data records (e.g., every time stamp or hourly averages or some other time interval).

- Would one master database be preferable or multiple databases that can communicate easily with each other and facilitate an exchange of data?
- Ensure that the solution is adaptive and whatever gets built should be relevant for future methods and data collected at assorted spatial scales and time intervals.
- Need to determine the potential audience for both the database itself and its outputs.
- Need to determine how data will be cited or attributed to the groups collecting the data.

1 Introduction

Long Island Sound is a tremendously valuable natural resource that provides numerous ecosystem services on which we depend. Recognizing its importance, dozens of local stakeholder groups work diligently toward the Sound's protection. These groups span governmental agencies and non-governmental organizations, from academics to very localized community groups. Some of these stakeholders have focused on specific pollutants (like microplastics or sewage) or specific environmental concerns (like hypoxia), while others approach the resilience of the Sound, as a whole, to large-scale perturbations like sea-level rise. These efforts have two things in common: (1) a shared goal of a healthy and resilient estuary, and (2) the generation of an enormous amount of environmental data.

This growing wealth of data for Long Island Sound presents opportunities and challenges. As the volume of data grows, especially long-term records, we are better prepared to understand and effectively manage this natural resource. However, because of the large number of data generators involved, these data are being collected in disparate ways and shared inconsistently among potential end users. For some organizations, the data being collected may only be shared via technical reports to a relatively small distribution list, directly with State or local government, or not at all. This presents a missed opportunity for scientists, managers, and the public to access the full body of data available to support informed decision-making.

Data management and sharing is a complex problem across all disciplines. A 2009 issue of the journal *Nature*[1] directly addressed the "neglect" of scientific data: "All too many observations lie isolated and forgotten on personal hard drives and CDs, trapped by technical, legal and cultural barriers — a problem that open-data advocates are only just beginning to solve." The issue of neglected data impacts the conservation and stewardship of the Sound; our efforts will only be as good as the information available to inform them. Over the past decade, there has been an increasing emphasis on the importance of data stewardship. For example, both the National Science Foundation and the National Oceanic and Atmospheric Administration (NOAA) require all prospective grantees to submit a "Data Management Plan" detailing how their data will be stored and shared. Despite this, many barriers remain that render data sharing inconvenient or even prohibitive.

A 2013 report by Vaudrey and colleagues explored the existing network of community-based monitoring groups around the Sound[2]. Twenty-one groups were found to have active monitoring programs at the time of that survey, and that number has grown since the Unified Water Study (led by Save the Sound) began in 2016. One of the suggested priorities that emerged from that extensive analysis was: "A centralized database in a standardized format is

[1] Nelson, B. (2009). Data sharing: Empty Archives. *Nature*, *461*, 160–163

[2] Vaudrey, J. M. P., Alonzo, J., Esposito, A., Johnson, C., Murphy, M. D., & Yarish, C. (2013). Evaluation of current community-based monitoring efforts and recommendations for developing a cohesive network of support for monitoring Long Island Sound embayments. Final report submitted to the New England Interstate Water Pollution Control Commission

essential to a cohesive network of community-based monitoring in Long Island Sound. The database should include some built-in reporting features as well as the ability for end users to download the data." In fact, it was identified as one of the top 5 priorities from the entire analysis. Another of those top 5 priorities was to increase communication among stakeholders (which would further facilitate data and knowledge sharing). That need is echoed throughout the report. For the many groups collecting data around the Sound, a need exists to store, share, and visualize data in new and collaborative ways.

Here, we built off of the work done for that 2013 report in an effort to (1) determine what the ideal database would entail such that it could meet the needs of the maximum number of data generators and end users, and (2) provide a mechanism for communication among Long Island Sound data generators such that redundancy of effort is minimized, efficient use of potentially shareable resources is maximized, collaboration is promoted, and new groups can be bolstered by a network of groups with more experience. This report will inform the development of a new Community of Practice around this issue. To meet these goals, we conducted surveys and one-on-one interviews with stakeholders from around Long Island Sound, as well as with relevant experts from other geographic regions. In the following sections, an aggregation of the needs and priorities of these multiple stakeholders are presented.

2 Data Collection Around Long Island Sound

In this study, we sought to explore the needs of groups collecting data in the greater Long Island Sound area as it relates to the management, sharing, and visualization of their data to assess the demand for and potential role for a shared, collaborative solution. The results shown in this section indicate that monitoring groups have a wealth of data, they want to share those data with the public, but many are not currently doing so.

Below are the findings of two online surveys (n = 55 and n = 39) and one set of in-depth key stakeholder interviews (n = 7) of people representing organizations that collect data around Long Island Sound. Our surveys and interviews indicated that 55 different groups are actively involved in water quality monitoring in freshwater sites, coastal salt water sites, as well as in Long Island Sound.

2.1 Survey of Monitoring Status (Survey #1, n = 55)

GOAL: One of the first steps in developing a Long Island Sound data platform for water quality monitoring is determining who is doing the monitoring, what parameters they are monitoring, and how they conduct their monitoring. These data inform the development of the interface and the structure of the database.

METHODS: In the fall of 2019, a survey was distributed to 349 people in the coastal watershed of Long Island Sound, as well as the southern half of Connecticut, and all of Long Island, to also capture people who solely conduct freshwater monitoring. The full list of the survey questions is available in Appendix 6.4 (page 80). We had a response rate of 23%, with not all of those respondents actively conducting monitoring; the survey was also used as a way to find people interested in being interviewed at a later stage of the project. This return rate is considered excellent for our study: we were not conducting a random survey to gather opinions. Instead, we cast a wide net across many people and organizations who might have been engaged in monitoring activities. Non-responders were people we were already pretty sure were not actively monitoring. Our approach did capture a number of people new to us as active monitoring groups.

Of the 81 people answering the survey, we had 55 responses from unique groups (we deleted replicate responses submitted by multiple people from the same group, working in the same site, using the same methods). Responses indicate these 55 respondents are working in a minimum of 124 salt water areas, 55 freshwater areas, and 13 of the respondents also work in the open waters of Long Island Sound. Our estimates are minimum numbers as some respondents were not specific in their responses (e.g., “many tributaries in the Housatonic watershed” versus providing specific locations).

In analyzing the survey responses, the focus was on the number of sites monitored under each program. If the same group participated in two different monitoring programs in the same site, the site was listed twice but with the associated parameters specific to each program. Additionally, if a respondent indicated they monitored in three sites, their responses were replicated so that all three sites were represented in the final tallies. Sites were divided into freshwater and salt water systems for some parameters. A total of 192 places/programs are represented, where "places/programs" refers to multiple programs and/or organizations per location or multiple sites monitored by an organization under one program.

Many of the sites are sampled by organizations who are members of State-funded monitoring programs or supported through regional efforts; a few highlights are noted here, see Appendix 6.1 (page 54) for a full list of sites and organizations.

- The Interstate Environmental Commission conducts monitoring in three Long Island embayments.
- Of the salt water systems included in the data analysis, 38 participate in Tier 1 of the Unified Water Study and 12 participate in Tier 2 of the Unified Water Study; many of these systems are also studied under other programs.
- In the freshwater areas, 22 sites participate in CT DEEP's Riffle Bioassessment by Volunteers (RBV).
- The Maritime Aquarium participates in CT DEEP's alewife monitoring program in four sites.

2.1.1 Results on Parameters (Survey #1 – Monitoring Status, n = 55)

Water quality monitoring and bacterial monitoring are the most popular type of programs (Figure 1). Temperature, salinity, and dissolved oxygen are monitored in most water quality programs. Bacteria monitoring programs often collect temperature data as well, but may collect no other water quality parameters. Turbidity (NTU turbidimeter, Total Suspended Solids), pH, and chlorophyll-*a* are also popular parameters. Light (Secchi depth, light meter) and nutrients are collected in about half of the sites and programs. All other parameters listed are rarely collected; in this case, most of the rarer parameters were collected by one group.

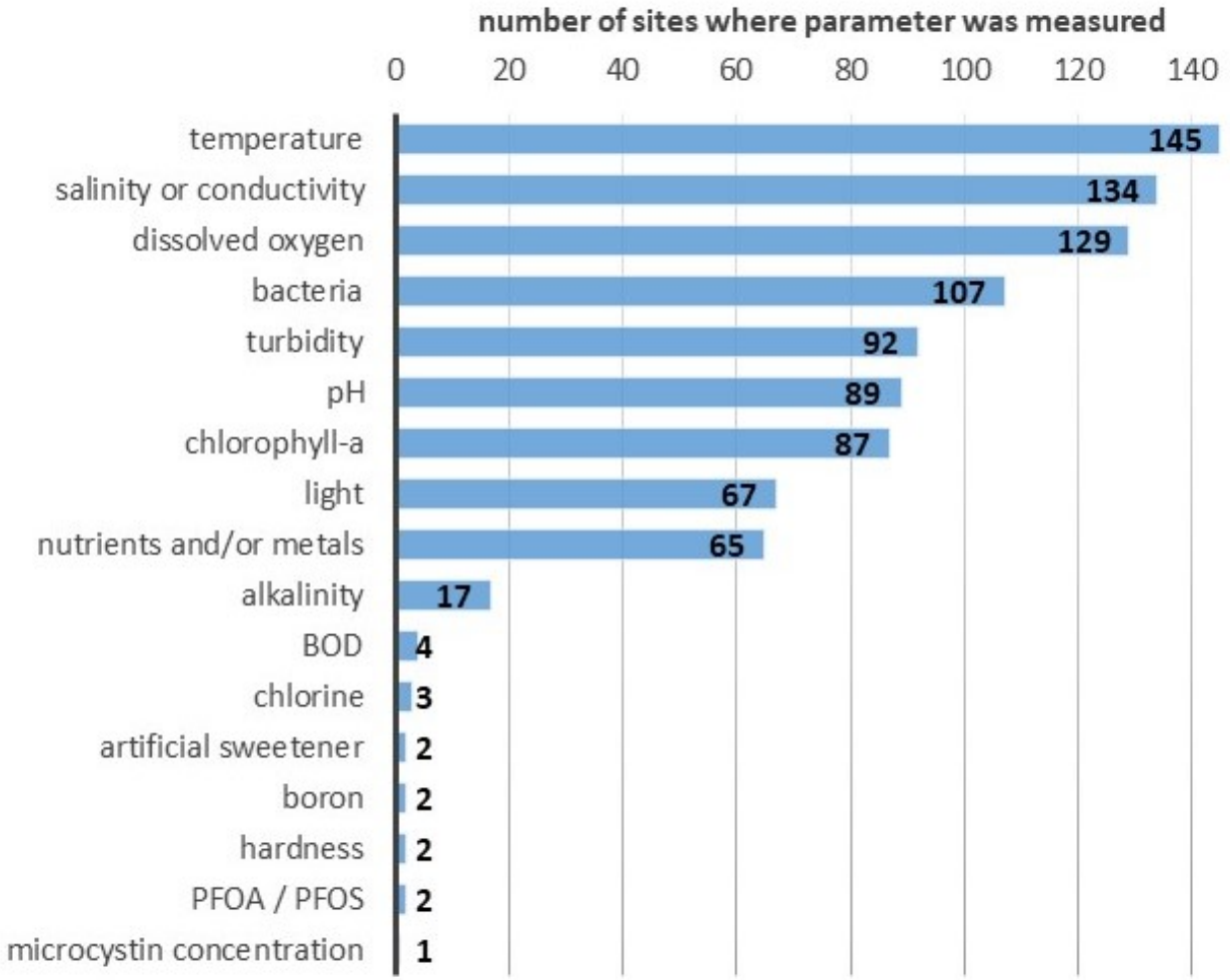


Figure 1. Parameters monitored in 192 sites/programs.

Other types of sampling include characterizing the sediment (grain size, organic content), and assessing the community of organisms living in the environment, from seaweed and eelgrass (macrophytes), to the microscopic organisms that live in the water column (plankton) to larger organisms (e.g., seals, fish, birds, insects, etc.; Table 1).

Table 1. Sampling types being conducted by groups other than water quality. Respondents (n = 55) were asked about other types of sampling they conduct. Responses are divided into salt water and freshwater environments, with a column for responses where the type of water was not specified. Most of these parameters do not have nationally or State-standardized protocols. Agency standardized protocols include: CT DEEP's Riffle Bioassessment by Volunteers (insects) and CT DEEP's alewife monitoring program. Grey areas indicate that the organism is not present in that type of water.

Non-Water Quality Sampling	salt water	fresh-water	type of water was unspecified	organizational or program (e.g. UWS) protocol	agency standardized protocol
sediment	19	2	8	29	
macrophytes	57	3	8	68	
plankton	3	1		4	
benthic organisms (not insects)	23	3	13	38	
aquatic insects		28		8	20
fish	10	3	13	22	4
birds	6	1	6	13	
seals	4			4	
Diamondback Terrapin & prey			3	3	

2.1.2 Results on Methods (Survey #1, n = sites/programs as listed for each parameter)

Understanding the methods used for each parameter is key to developing a database that meets the needs of end users.

- The database should provide users with the option of entering their data in the form in which it was collected, including a record of the unit and the method used.
- The end users of the data likely have preferences for units or may change a display choice based on differing methods. The ability to accommodate these options in the database interface and in the visualization tools is key to making these products truly useful – recognizing that organizations have different needs is important.

The following series of figures provides a breakdown of the methods used for the most popular parameters, along with typical units. For all parameters, use of a hand-held digital device accounts for about 50% of the data. The remainder of the data are collected using other methods.

Temperature
Monitored in 146 sites/programs.
Typical Units: °C, °F

Hand-held digital device refers to a probe (YSI, Eureka, etc.) with a digital read-out. The user submerges the probe and can read the results immediately on a hand-held unit. This is usually a point reading = a single reading at a specific depth.

Deployed digital device is an instrument that is left in the water for hours to months. The unit logs data internally and the user downloads the data once the unit is retrieved. Data are collected in one location at one depth, but collected at set time intervals (e.g. every 15 minutes).

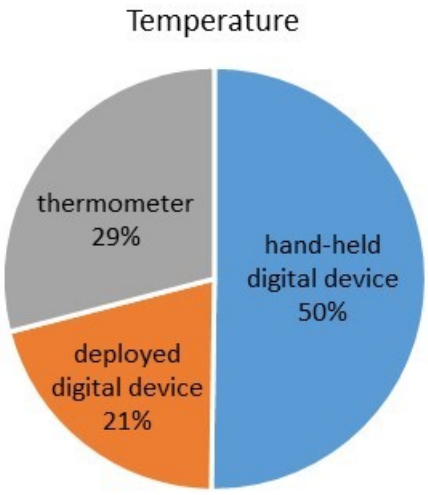


Figure 2. Summary of methods used for temperature data collection.

Thermometer refers to a digital or analog (alcohol) thermometer that is dipped into a collected sample and read immediately.

Salinity (salt water) or Conductivity (fresh water)
Monitored in 135 sites/programs.
Typical Units: ppt = psu, µS/cm

Hand-held digital device – see description for temperature.

Deployed digital device – see description for temperature.

Refractometer (for salinity) is a hand-held device, typically analog but may be digital. A drop of water is placed on a plate. The refraction of light through the thin layer of water is used to determine salinity.

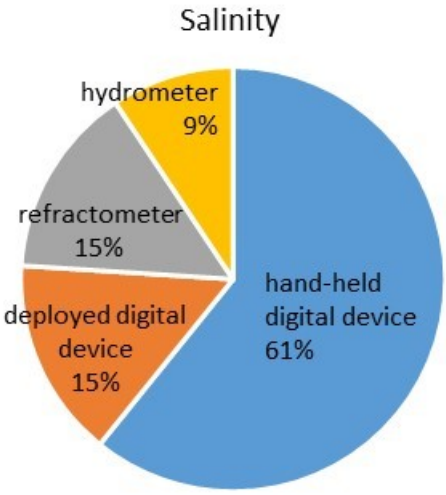


Figure 3. Summary of methods used for salinity or conductivity data collection.

Hydrometer (for salinity via specific gravity) is a calibrated sealed glass ampoule containing a scale for reading specific gravity (density relative to a standard). The sealed glass ampoule floats, displacing varying amounts of liquid depending on the salinity of the liquid. This can be converted to salinity.

Dissolved Oxygen Concentration or Percent Saturation

Monitored in 130 sites/programs.
Typical Units: mg/L = ppm, % saturation

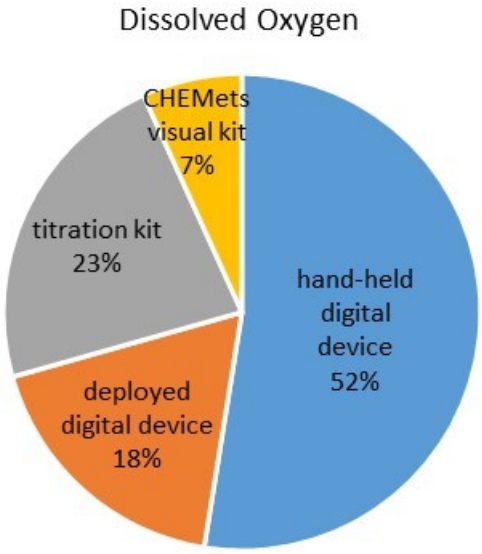


Figure 4. Summary of methods used for dissolved oxygen data collection.

Hand-held digital device – see description for temperature. Allows for readings every 0.01 mg/L.

Deployed digital device – see description for temperature. Allows for readings every 0.01 mg/L.

Titration kits use wet chemistry to lock dissolved oxygen into a solid compound that can be dissolved with acid to form iodine (yellow in color, stronger color means more oxygen was originally present). A second solution is used to titrate (adding drop-by-drop) until the yellow color is gone. The amount of titrant used is equivalent to the oxygen present in the sample. A number of manufacturers provide field kits for easily conducting the titration (Hach, LaMotte, etc.). Allows for readings every 0.2 mg/L.

CHEMets Visual kits operate using the same wet chemistry as the titration kits but instead of titrating the sample, the user compares the intensity of the yellow iodine color to a reference standard to estimate the concentration of dissolved oxygen present in the original sample. Less precise than other methods, as you round to the nearest 1 mg/L.

Chlorophyll-*a* Concentration

Monitored in 88 sites/programs.
Typical Units: µg/L = ppb (in salt water), mg/L = ppm (in freshwater)

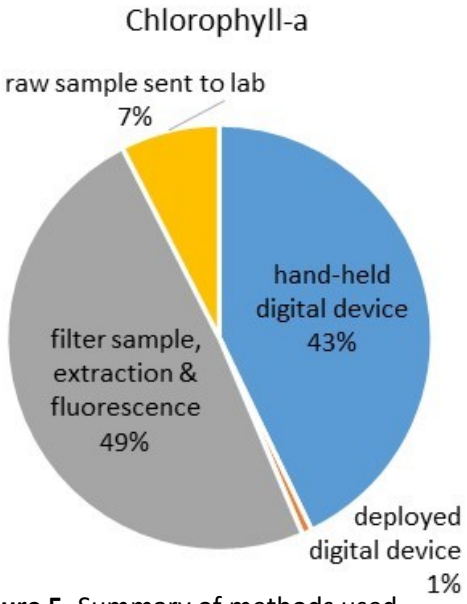


Figure 5. Summary of methods used for chlorophyll-*a* data collection.

Hand-held digital device – see description for temperature. Relative fluorescence units (RFU) or fluorescence standard units (FSU) must be converted to concentration using a reference sample of known fluorescence and concentration.

Deployed digital device – see description for temperature. See note under chlorophyll-*a* hand-held device regarding conversions.

Filter sample, extraction & fluorescence refers to collecting a water sample, passing it though a filter

with known pore size (typically 0.7 µm), extracting the chlorophyll-*a* pigment by soaking the filter in acetone, reading the intensity of the sample in a fluorometer, and relating the fluorescence to a concentration using the volume sampled and a reference standard. The filter may be frozen prior to extraction, allowing for the sample to be held for 28 days before analysis.

Raw sample, sent to analytical lab refers to the procedure described in "*Filter sample…"* This method differs in that the raw sample is immediately delivered to the lab, where the sample is filtered and further processed. Most labs use the fluorometric technique as it is more sensitive at low concentrations, though some may still use the spectrophotometric technique.

Water Clarity & Turbidity
Monitored in 92 sites/programs.
Typical Units: turbidity (NTU, FNU), total suspended solids (mg/L, g/m^3), Secchi depth (m), light attenuation (m^{-1})

Water clarity may be described as the light penetrating through the water column (Secchi disk, turbidity tube, light meter), as turbidity determined optically (NTU, FNU), or as total suspended solids (TSS). Turbidity is considered equivalent to water clarity. TSS and turbidity capture many of the same particles that decrease water clarity (clay, bacteria, plankton, non-settleable solids). Turbidity also captures dyes and colored dissolved organic matter not captured by TSS. TSS captures settleable solids, not captured by turbidity.

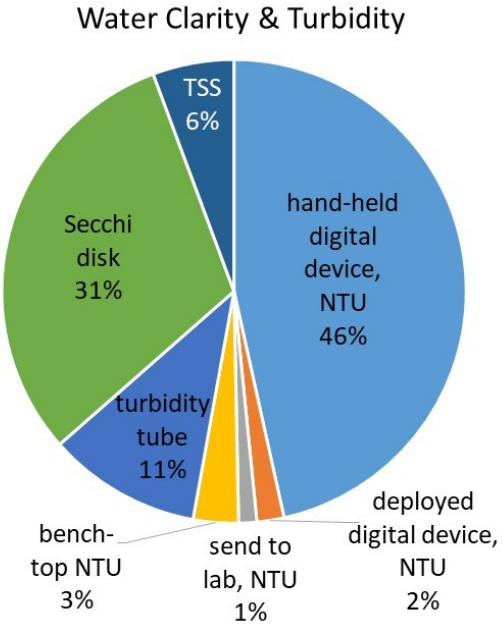


Figure 6. Summary of methods used for water clarity and turbidity data collection.

Hand-held digital device – see description for temperature. This category may also include a handheld turbidimeter designed for field use; see "*Send to lab, NTU,"* below.

Deployed digital device – see description for temperature.

Send to lab, NTU and *Bench-top NTU* both refer to the collection of a water sample and the analysis of the sample using a turbidimeter. A raw sample of water is placed in a device which assesses the transmittance of light through the sample and provides output, typically in units of NTU.

Secchi disk is a round, plate-like disk typically divided into four pie-shaped quadrants of alternating black and white. The disk is lowered into the water until it disappears from view.

This depth is the "Secchi depth" and is considered equivalent to the depth where 1% of the surface light remains.

Turbidity Tube is a method for gathering Secchi depth data when the depth is too shallow to allow the Secchi disk to disappear from view. Water is placed in a long tube and the user compares a miniature Secchi disk at the end of the tube to reference images of a Secchi disk to determine an estimate of the turbidity in NTU or as Secchi depth.

TSS (total suspended solids) is measured by collecting a water sample, filtering it through a 1.5 µm filter of known weight, drying the filter, and weighing the particles retained on the filter. A further step may be conducted to estimate the volatile TSS (V-TSS). For this, the sample is burned at a high temperature, leaving behind only the non-volatile solids.

pH

Monitored in 90 sites/programs.
Typical Units: unitless

Hand-held digital device – see description for temperature. Precision is typically 0.01 units.

Deployed digital device – see description for temperature. Precision is typically 0.01 units.

pH strips – indicator strips dipped into a sample. The color change on the strip is used to indicate the pH. Precision will be 0.25 to 1 pH unit, depending on type of strip used.

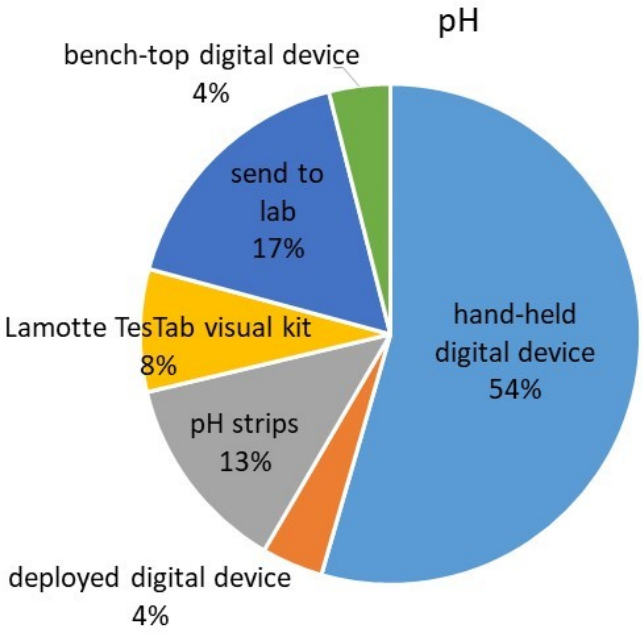


Figure 7. Summary of methods used for pH data collection.

LaMotte TesTab visual kit – a tablet of reagent is added to a sample. The color change corresponds to the pH, a visual reference is used to estimate pH. Precision is typically 1 unit.

Send to lab and *bench-top digital device* – a water sample is delivered to a lab and read with a bench-top pH meter. Samples must be kept at a certain temperature and delivered to the lab within a certain time frame (usually within hours of collection). Precision is typically 0.01 units.

Alkalinity
Monitored in 18 sites/programs.
Typical Units: mEq/L

Total alkalinity is measured by collecting a water sample, and determining the amount of acid needed to bring the sample to a pH of 4.2.

Bench top instrument implies the group conducts their own analysis of alkalinity.

Send to an analytical lab indicates the group sends the sample out for analysis.

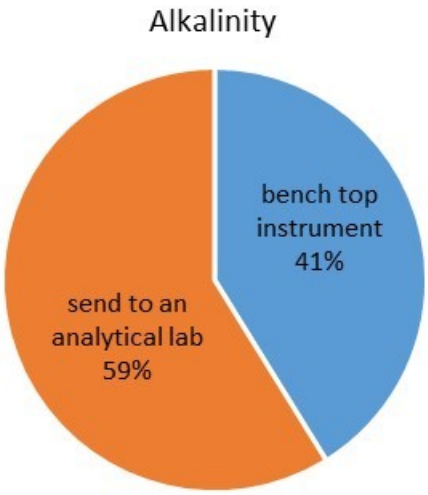


Figure 8. Summary of methods used for alkalinity data collection.

Nutrients and Metals
Monitored in 4 to 60 sites/programs, depending on type of nutrient.
Typical Units: mg/L = ppm, µg/L = ppb, mM, µM

The various types or "species" of nutrients routinely monitored are shown in the figure below. For example, species of nitrogen (N) include ammonium (NH_4^+), nitrate (NO_3^-), nitrite (NO_2^-), and organic N; with total N being the sum of these species. Samples may be filtered (thus only measuring the dissolved fraction of the species) or unfiltered (thus measuring the total for the species, including particulates). Various species of the major nutrients are typically monitored, where those major nutrients include: C = carbon, N = nitrogen, P = phosphorus, S = sulfur.

Analytical lab – analysis of collected water samples. Precision is typically < 1 µM (~0.02 mg/L).

Field Kit – a pre-packaged chemistry kit for use in the field. Manufacturers include LaMotte, Hach, CHEMetrics. Precision is typically 0.2 mg/L (~15 µM). The sulfate field probe is included in this category and has a greater precision, at about 0.01 mg/L.

Test Strip – a tab that is dipped into the sample. Blocks on the tab change color and that color is compared to a reference. The precision varies by analyte but is usually low accuracy: for example, nitrate can be read at specific levels: 0, 10, 25, 50, 100, 250, and 500 mg/L.

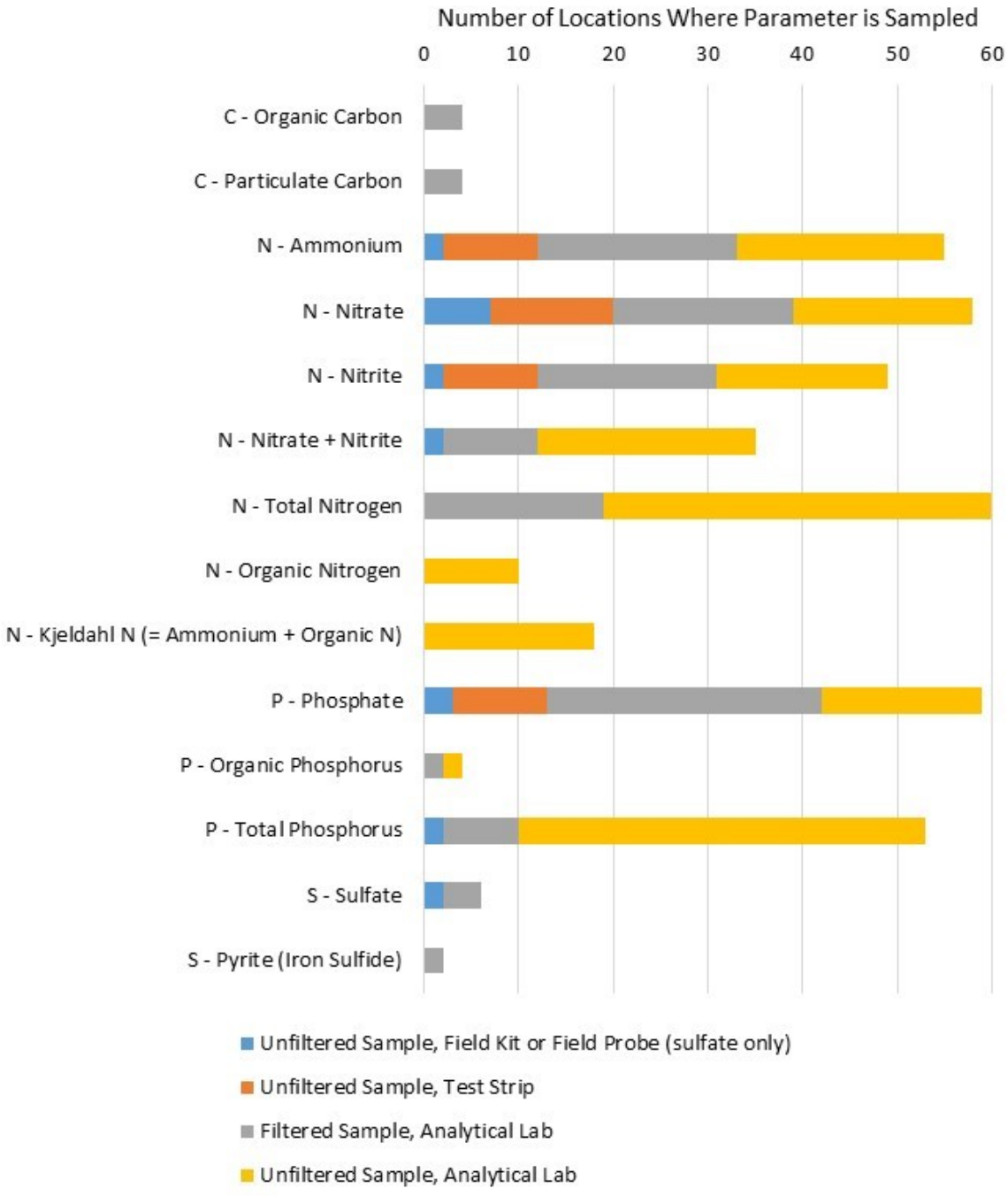


Figure 9. Summary of sampling methods for nutrients and metals and number of locations monitored for each parameter.

Bacteria
Monitored in 104 sites/programs.
Typical Units: CFU (colony forming units) = MPN (most probable number)

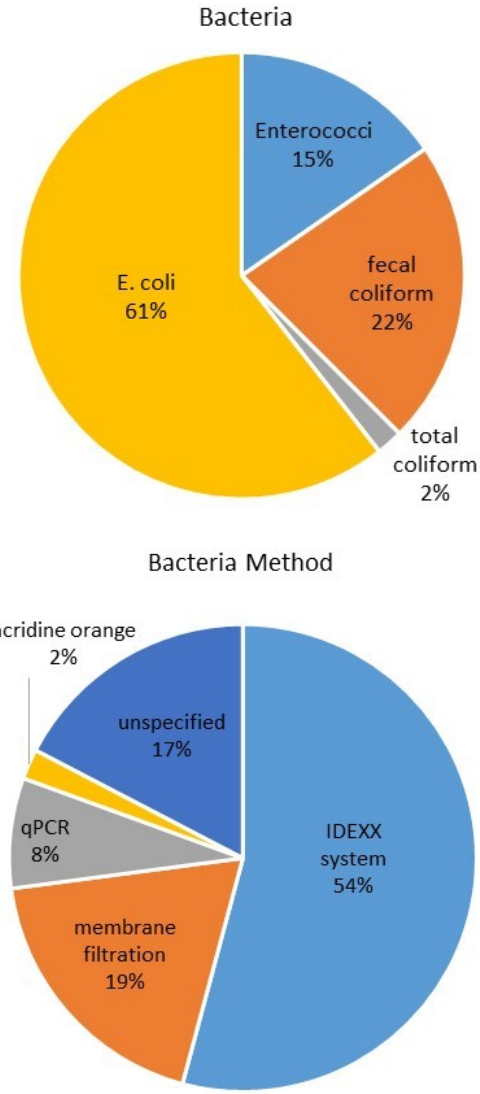


Figure 10. Summary of bacteria monitoring type and methods.

Bacteria are most often monitored as an indicator of waters harboring fecal matter with the potential to be harmful to humans, typically defined as causing gastrointestinal (GI) distress. The indicator bacteria are not the only harmful bacteria and in fact, most strains of the indicators do not cause human illness; rather, they indicate the presence of fecal contamination. The basis for recommending criteria that use bacterial indicators of fecal contamination is that pathogens often co-occur with indicators of fecal contamination[3]. In an unimpaired system, only a very small fraction of all bacteria in sea water or freshwater are harmful to humans.

Multiple indicator bacteria are used:

- *Escherichia coli (E. coli)* – EPA recommended fecal indicator bacteria in freshwater.
- *Enterococci* – EPA recommended fecal indicator bacteria in freshwater and marine waters.
- *Fecal coliform* – previously recommended by the EPA as a fecal indicator in recreational waters. Replaced in the EPA guidelines in 1986 by *E. coli* and enterococci.
- *Total coliform* – previously recommended by the EPA as a fecal indicator in recreational waters. Replaced in the EPA guidelines in 1986 by *E. coli* and enterococci.

The methods (all EPA approved) used for determining bacteria indicators include:

- Simple, semi-automated IDEXX methods
- Membrane filtration followed by incubation and colony counts
- qPCR, a fast, DNA-based test
- Filtering followed by staining with a fluorescent acridine orange and microscopic counting of individual bacteria (only used for counting ALL bacteria, not indicative of fecal contamination).

[3] EPA. 2012. Recreational Water Quality Criteria. OFFICE OF WATER 820-F-12-058. https://www.epa.gov/sites/production/files/2015-10/documents/rwqc2012.pdf

2.2 Survey of Data Generator Needs (Survey #2, n = 39)

GOAL: As described above, this study sought to determine what the ideal data solutions would entail, such that the maximum number of data generators and end user needs could be met. Data and opinions gathered through this project support the development of a database to bolster the accessibility of data around the Sound and engagement of a network of stakeholders.

METHODS: Following the first survey, a set of follow up questions were developed and distributed in Spring 2020. While the first survey focused on capturing the scope of data (e.g., what types of data were being collected, where, and how), this second survey sought to better capture how groups are using their data and what needs they see in terms of management, sharing, and visualization of data. The survey was broadly distributed to those stakeholders identified through the first survey, with a good representation of respondents from around the Sound (Figure 11).

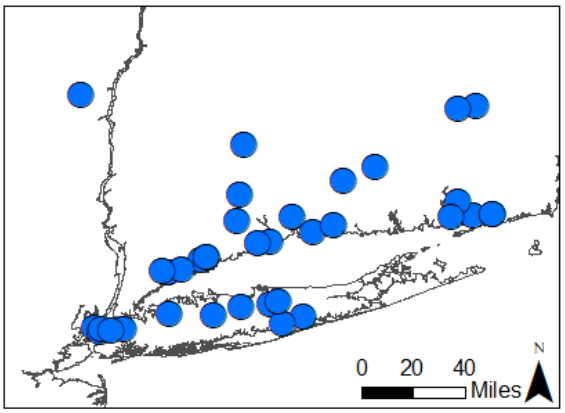


Figure 11. Thirty-nine people responded to the second "data generator" survey (distributed to 61 people), with good geographic representation from around Long Island Sound. Please note that the survey respondent location was collected by the survey software from the IP address of the respondent and thus represents the location of the respondent when they took the survey and not necessarily where they conduct monitoring on Long Island Sound.

2.2.1 Results Within the Themes (Survey #2 – Data Generator Needs, n = 39)

Respondents were asked to rank their agreement with a series of statements broken up into themes. The exact phrasing of each question can be found in Appendix 6.4 (page 80). The following subsections provide responses to each of these questions, visualized in the pie charts.

2.2.1.1 General Issues (Survey #2 – Data Generator Needs, n = 39)

Data management presents technical challenges for monitoring organizations, as well as monopolizing time that could be spent on other goals. Out of 39 respondents, only 8% of respondents "strongly agreed" they were happy with their process for entering, storing and sharing their data, while 49% indicated they were not happy. There was interest in making changes in these areas; 65% of respondents were interested in streamlining their process for storing and sharing data. Despite the wealth of data being collected, these data may not currently be used to their full potential in the stewardship and management of the Sound. Most respondents (62%) indicated that if they had more time, they would do more with their data. Importantly for the needs assessment goals in the present study, 68% of respondents would consider using a centralized database for Long Island Sound data if one existed.

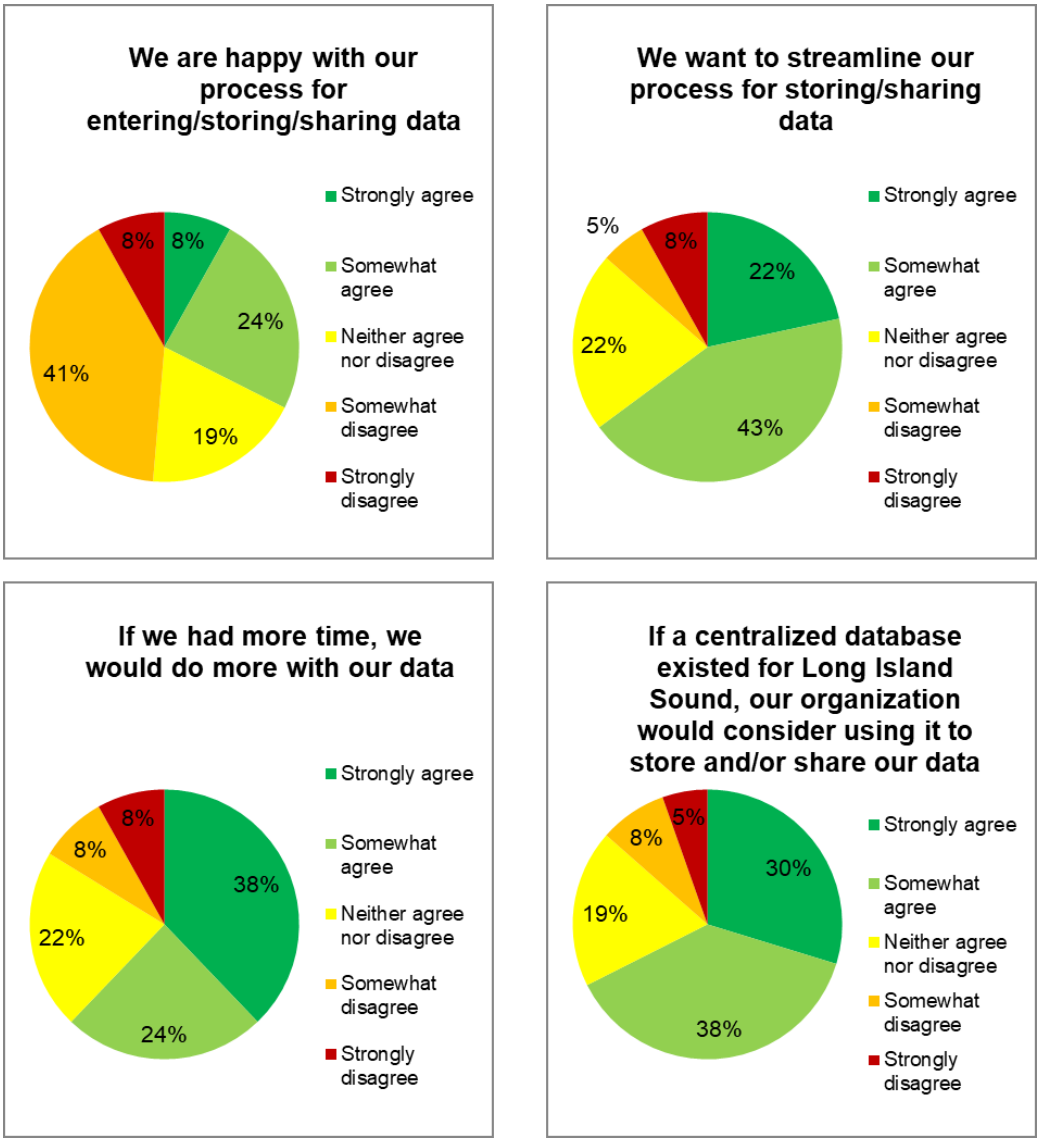


Figure 12. Percentage of responses (n = 39) to general data-related statements as paraphrased at the top of each graph.

2.2.1.2 Data Entry (Survey #2 – Data Generator Needs, n = 39)

Fewer concerns were expressed specifically about data entry. 81% of respondents indicated they are comfortable with entering their data with most preferring Excel, and a few (8%) who felt strongly that they wanted an option other than Excel for data entry (however, this may be the result of either not feeling comfortable with Excel, or conversely wanting to use a more complex set of tools). There was some interest in using phones or tablets for mobile data collection (60%), though some groups shared logistical constraints they have faced in their efforts to transition to digital data entry in the field.

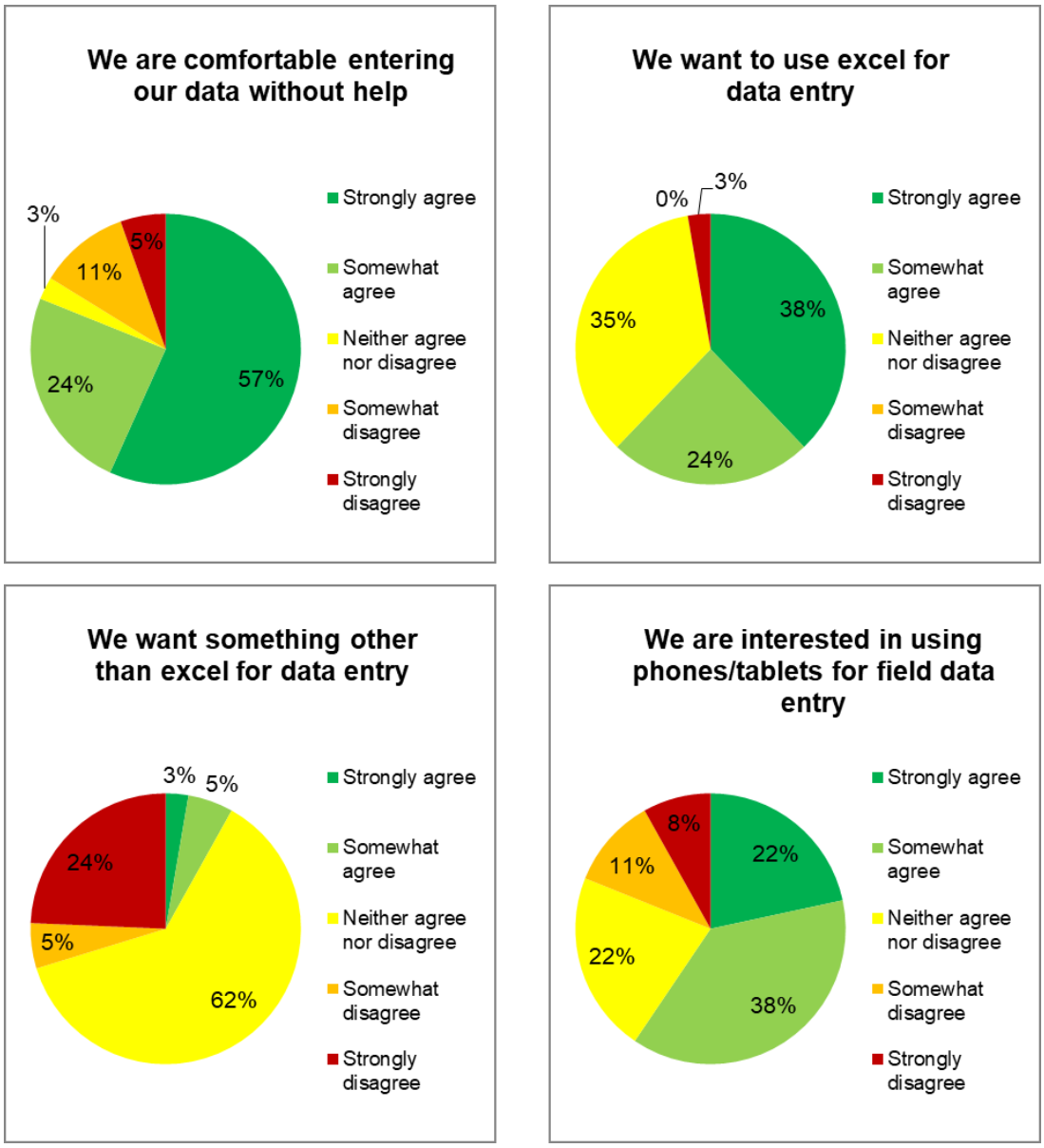


Figure 13. Percentage of responses (n = 39) to statements related to data entry as paraphrased at the top of each graph.

2.2.1.3 Data Management (Survey #2 – Data Generator Needs, n = 39)

A greater need was seen in the community related to the management of data. 57% of respondents agreed that data management is a burden on their time (Figure 14). Only 16% "strongly agreed" their data are stored digitally in a well-organized format (with 33% either disagreeing or strongly disagreeing). Most groups expressed that their data are safely backed up (73%) and that they do not need help with data storage (63%). However, there was greater interest in assistance with data analysis.

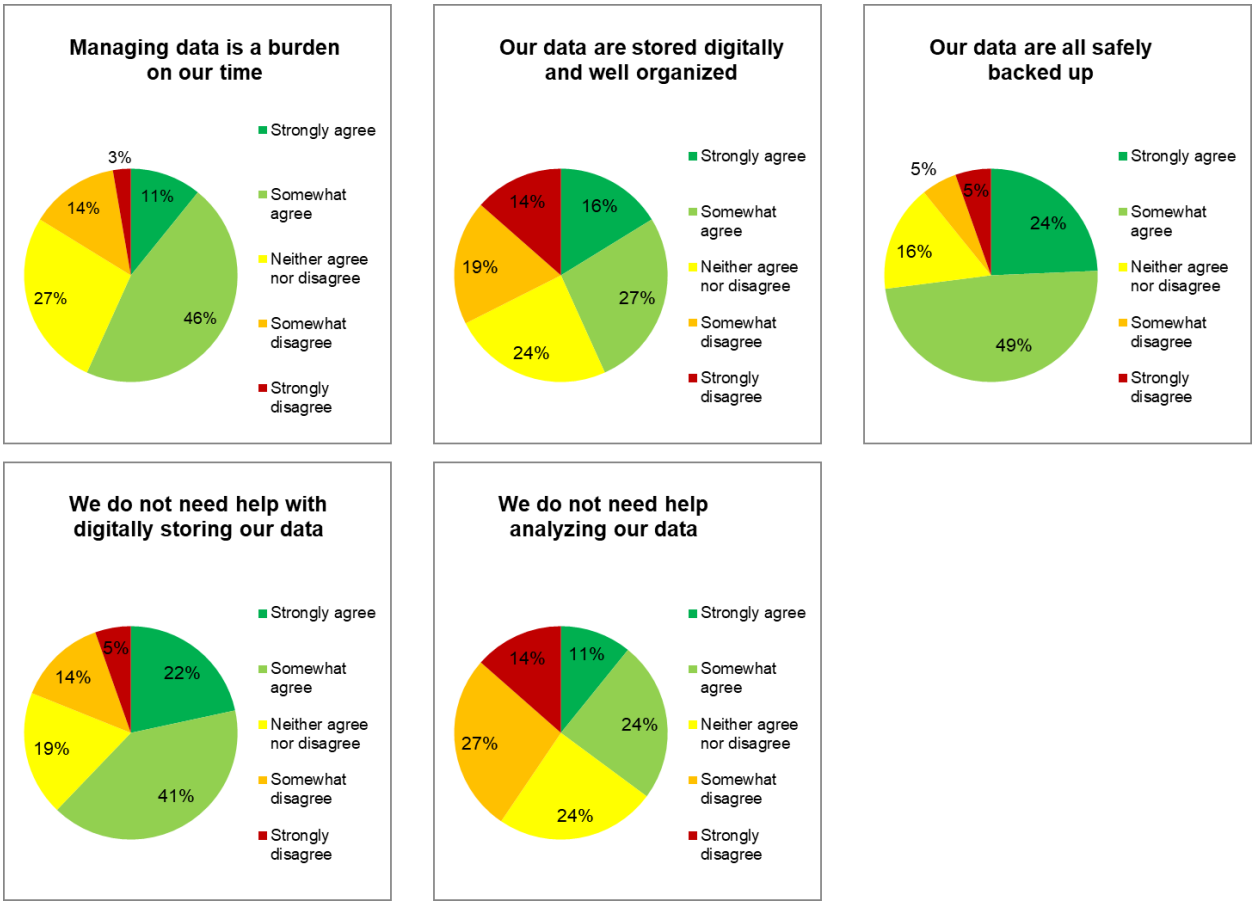


Figure 14. Percentage of responses (n = 39) to statements related to data management as paraphrased at the top of each graph.

2.2.1.4 Quality Control (Survey #2 – Data Generator Needs, n = 39)

Quality control for data entry was not universal, with only 30% of respondents "strongly agreeing" and 24% "somewhat agreeing" that they have a quality control process in place for their data; this may be an area with some room for expansion of those efforts and opportunities for shared learning on best practices. However, collecting data under an EPA Quality Assurance Project Plan (QAPP) was important to 59% of respondents.

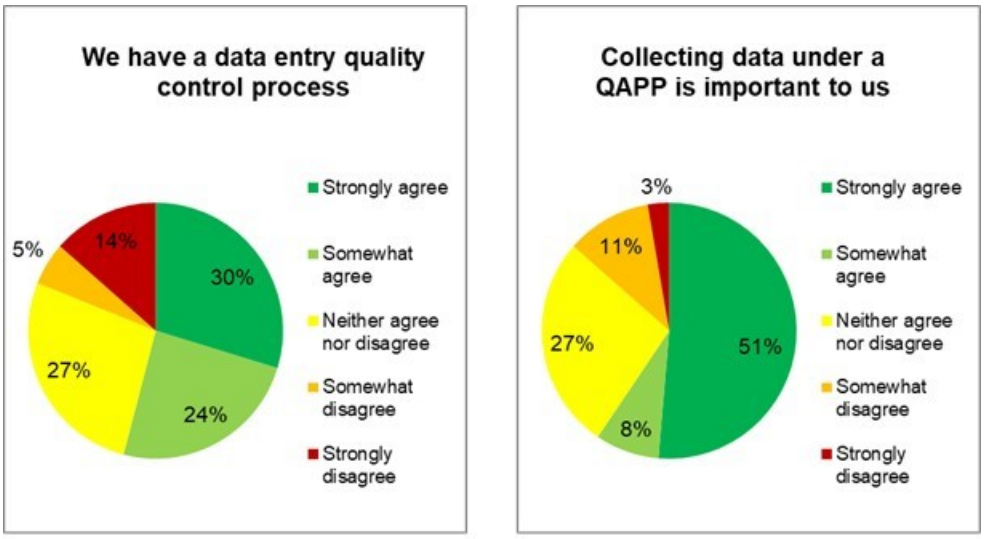


Figure 15. Percentage of responses (n = 39) to statements related to quality control as paraphrased at the top of each graph.

2.2.1.5 Environmental Protection Agency's (EPA) Water Quality Exchange (WQX) (Survey #2 – Data Generator Needs, n = 39)

Unfortunately, we found that WQX is poorly understood and currently, not widely used; this presents an opportunity to increase data sharing through this Federal database and portal (Water Quality Portal, or WQP). 68% of respondents indicated they did not understand what WQX is and/or how to use it (Figure 16), and only 19% of respondents indicated that most or all of their data are currently in WQX. This low engagement with WQX could result from a variety of drivers and our survey did not seek to assess the barriers to using WQX. While we did not specifically ask whether all respondents were collecting data suitable for WQX, results of the first survey indicate that all groups collect water quality data which could be uploaded to WQX, as well as collecting other types of data that may be harder to submit to WQX. Some of the low engagement may be related to technical hurdles, with 74% of respondents indicating they could use some help getting data into WQX. Beyond that, 81% of respondents would consider using a tool to prepare data for WQX upload and all other responses were neutral (and *none* of the respondents disagreed with a statement asking if they would consider such a tool).

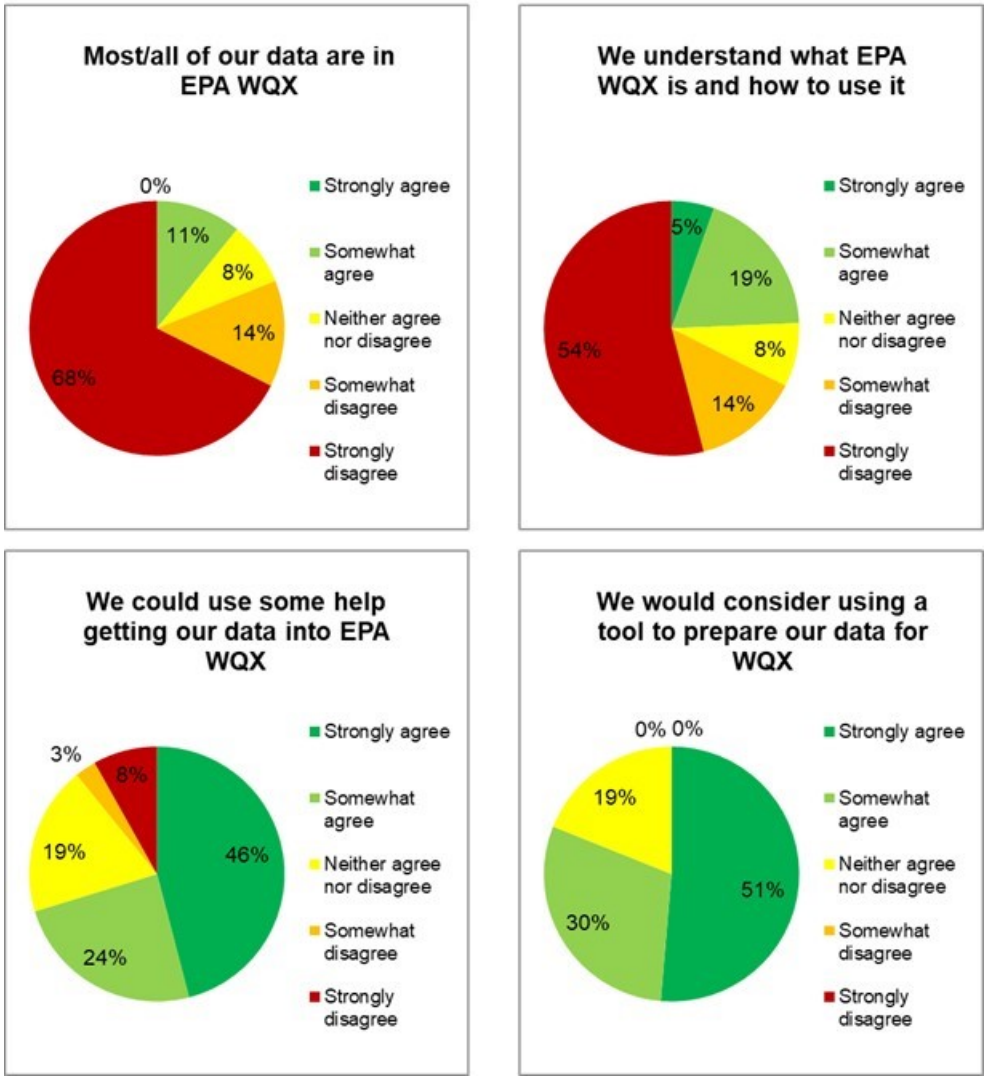


Figure 16. Percentage of responses (n = 39) to statements related to EPA WQX as paraphrased at the top of each graph.

2.2.1.6 Data Sharing (Survey #2 – Data Generator Needs, n = 39)

Engagement with the public and government partners was important to many of the survey respondents; 75% of respondents want to share their data with the public, and yet only 22% said their data are readily available for public download at the present time. While most (61%) said they could share data without much effort upon request, this activity may present an additional burden on staff or volunteer time at these organizations, or it may be a barrier to potential data end users. 69% of respondents indicated that sharing data with government partners is important to them, but again, only 19% have most or all of their data in WQX. This appears to be a core issue in the water quality data-generating community; most groups want to share their data, but lack a process through which to easily do so.

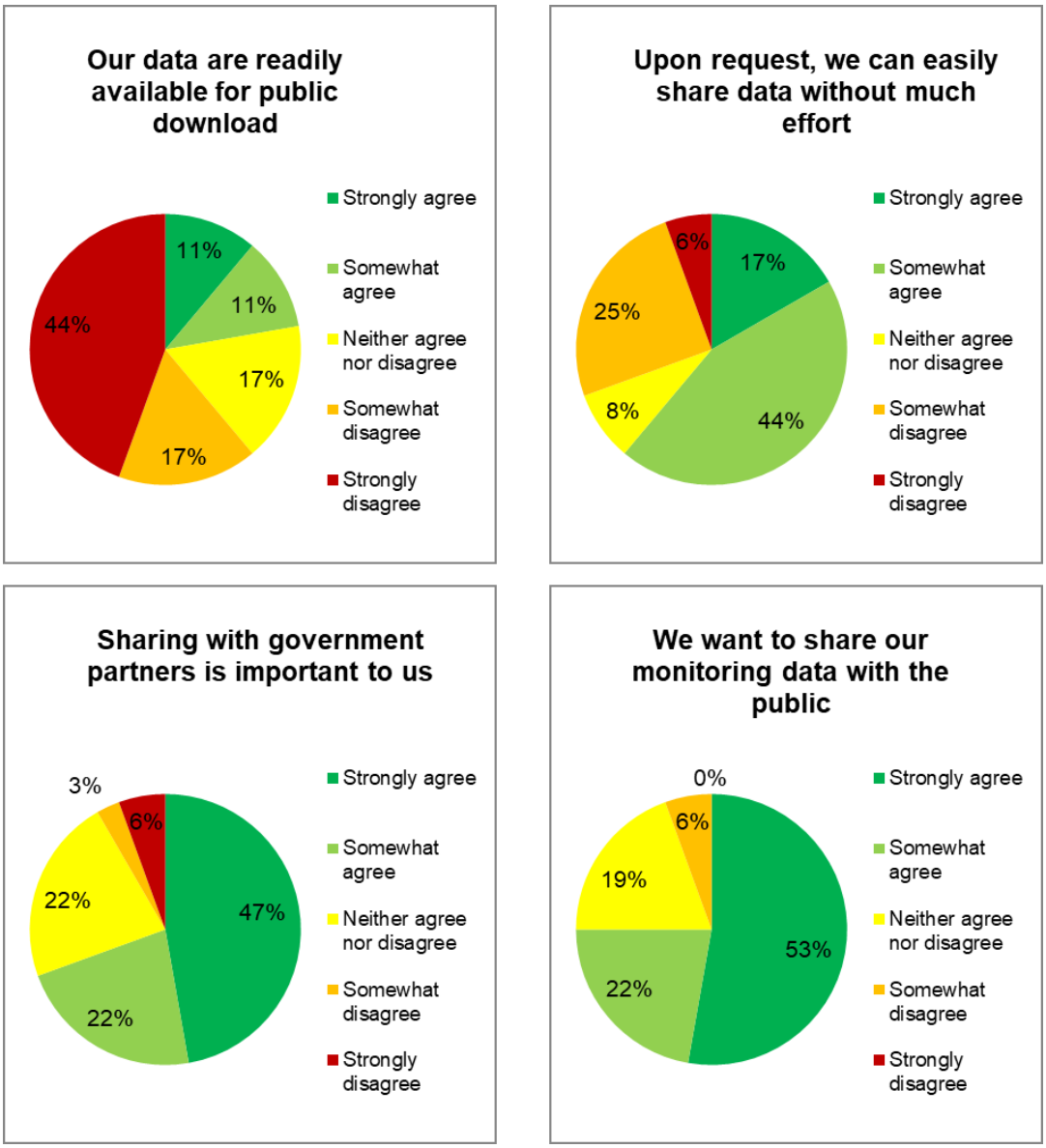


Figure 17. Percentage of responses (n = 39) to statements related to data sharing as paraphrased at the top of each graph.

2.2.1.7 Data Visualization (Survey #2 – Data Generator Needs, n = 39)

While 46% of respondents are currently "using tools" to visualize their data, 46% indicated they are not comfortable mapping their own data, and 40% indicated they are not comfortable graphing their data without help. This may show an opportunity for such groups to benefit from collaboration and regional leadership on this issue, and 63% of respondents expressed interest in receiving help with data visualization. Additionally, 65% of respondents agreed they would like to see their data on a map with other data from around the Sound.

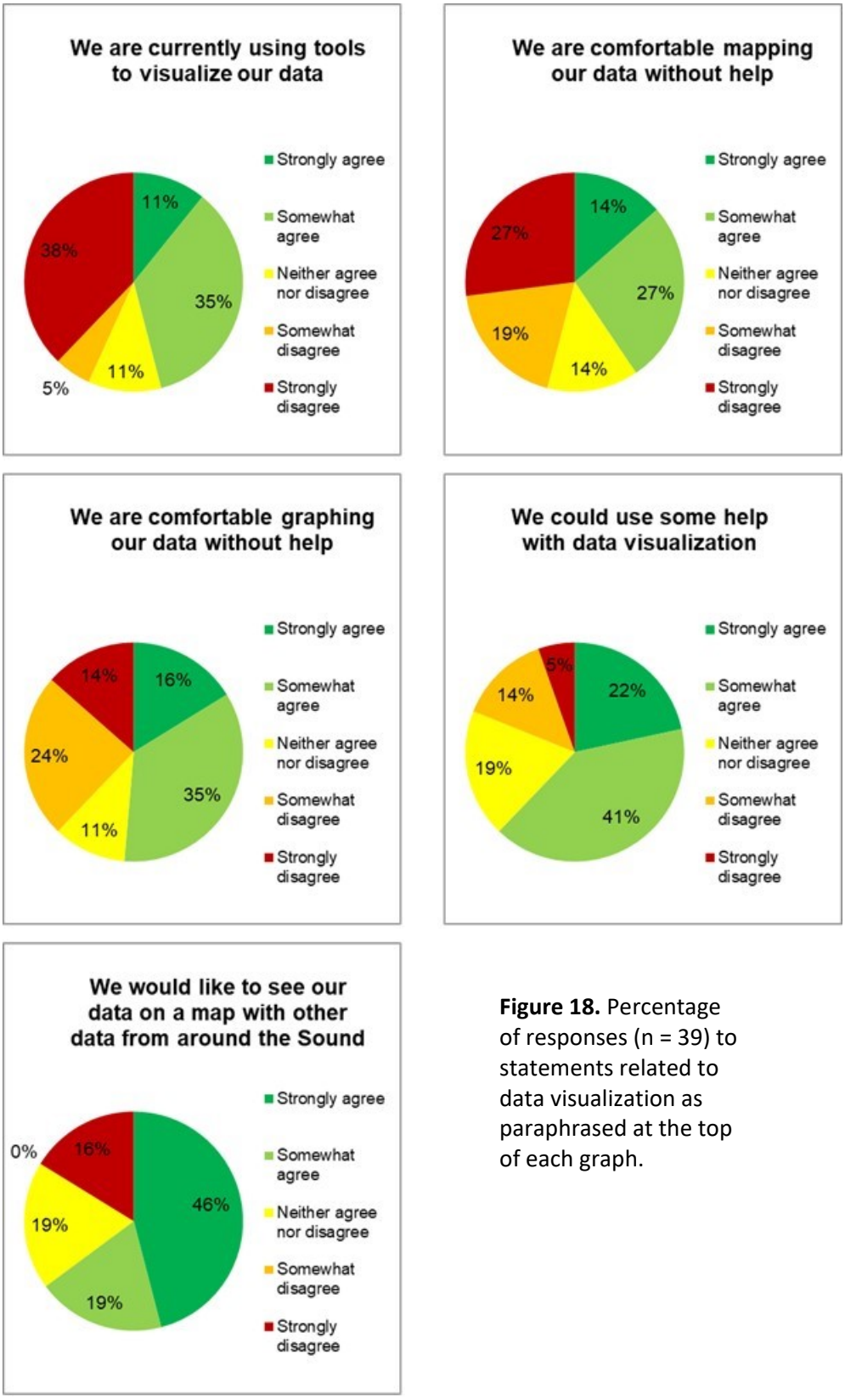


Figure 18. Percentage of responses (n = 39) to statements related to data visualization as paraphrased at the top of each graph.

2.2.1.8 Existing Tools and Programs (Survey #2 – Data Generator Needs, n = 39)

When asked about existing programs and tools, only 33% of respondents were familiar with the Sound Health Explorer, but 65% expressed interest in seeing their data on a map with other data from around the Sound. Most respondents (67%) were familiar with the Unified Water Study.

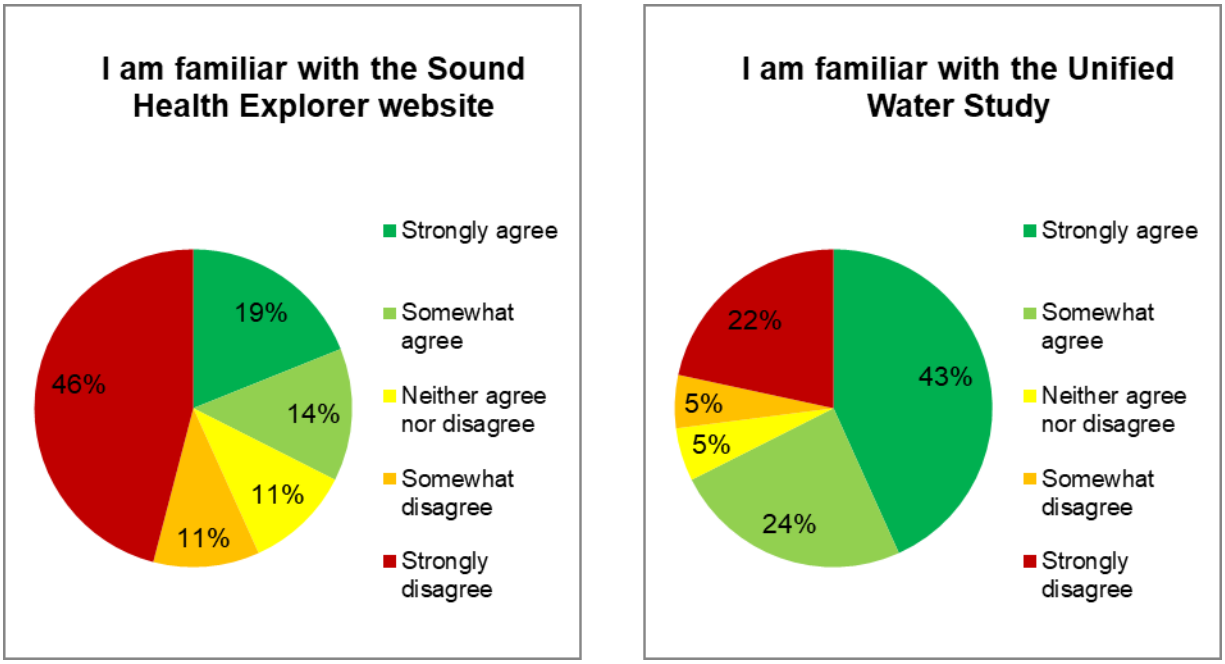


Figure 19. Percentage of responses (n = 39) to statements related to existing tools and programs as paraphrased at the top of each graph.

2.2.1.9 Summary of Write-In Comments (Survey #2 – Data Generator Needs, n = 39)

- Respondents expressed concerns about mobile apps for data entry related to phone data usage/signal access in remote field locations; also concerns about use of mobile devices in the field (e.g., hard to see screen on boat due to glare).
- Respondents expressed concerns about mobile apps for data entry related to the ability to conduct a quality check against a paper data sheet; redundancy needed between mobile entry and paper data sheet for QA/QC.
- Some respondents indicated they had tried mobile data entry without success.
- Most people seem happy with Microsoft Excel of Microsoft Access for data entry.
- Most people are comfortable with data management but are open to outside assistance.
- There exists a general need for better understanding of WQX/WQP.

2.2.2 Results on Consensus and Strength of Opinions (Survey #2 – Data Generator Needs, n = 39)

In an effort to explore the statements with which there was most agreement or disagreement and to explore the degree of similarity among respondents on these topics, we translated the 5 response categories onto a numerical scale from 1 to 4:

1. Strongly agree
2. Somewhat agree
3. Somewhat disagree
4. Strongly disagree

Because the "neither agree nor disagree" category was used by some groups or in some questions to imply "not applicable" and by/in others to indicate a neutral response, those were not translated into a numerical score. The data were thus scored according to this scale for the purposes of the findings shared below, with the complete set of responses reflected elsewhere (in pie charts and statements regarding the percentage of responses by category detailed in previous subsections).

Table 2. Statements for which respondents most strongly agreed (on average):

Statement	Mean Score (1=strongly agree, 4=strongly disagree)
"If a tool existed to help prepare our data for entry into EPA WQX, we would consider using it."	1.37
"We want to share our monitoring data with the public."	1.41
"We want to use Excel for data entry."	1.5
"Collecting data under an EPA-approved QAPP (quality assurance project plan) is important to our organization."	1.52
"Sharing our data with local (towns), State (CT DEEP, NY DEC) and Federal (e.g., EPA, NOAA) governmental groups is important to our organization."	1.57
"Our organization (staff and/or volunteers) is comfortable entering our own data without external help from consultants or others."	1.64
"We could use some help getting our data into EPA WQX."	1.67
"If we had more time, we would do more with our data."	1.83
"We would like to see our data on a map with other data from around the Sound."	1.83
"If a centralized database existed for Long Island Sound, our organization would consider using it to store and/or share our data."	1.87
"Our data are all backed up and recoverable if something were to happen to our primary data files."	1.90

Table 3. Statements for which respondents most strongly disagreed (on average):

Statement	Mean Score (1=strongly agree, 4=strongly disagree)
"Our data are currently readily available for public download, either through our website or a third-party website."	3.13
"We understand what EPA WQX is and how to use it."	3.26
"We want to use some other method [i.e., not excel] for data entry."	3.36
"Most or all of our data are currently uploaded to EPA's WQX system."	3.62

Statements for which all respondents had similar opinions (lowest standard error):

- "If a tool existed to help prepare our data for entry into EPA WQX, we would consider using it." (Mean = 1.37 ± 0.0895)
    - Note: Least variation (greatest agreement) among respondents.
- "We want to share our monitoring data with the public." (Mean = 1.41 ± 0.1165)
    - Note: Second least variation (second greatest agreement) among respondents.
- "Most or all of our data are currently uploaded to EPA's WQX system." (Mean = 3.62 ± 0.1195)
    - Third least variation (third greatest agreement) among respondents.

2.3 Data Generator In-Depth Interviews (n = 7)

GOAL: Interviews were conducted with a subset of respondents from the first survey who were actively engaged in monitoring. The main goal was to explore opinions about data management and visualization and to get a deeper sense of how various types of groups (sizes, corporate structure, etc.) handled their data and what they needed in a data management and visualization platform.

METHODS: In these interviews (n = 7), we solicited in-depth discussions on a variety of data-related topics and sought to capture both their opinions (qualitative) and to translate their thoughts on certain questions into a quantitative format that could be compared among interviewees. Below, the results are summarized and then on the following pages each table shows the discussion question or theme at the top (in bold text) as well as the responses (presented anonymously) to a series of more specific sub-questions. Representatives from The Maritime Aquarium at Norwalk, Coalition to Save Hempstead Harbor, University of Rhode Island Watershed Watch, Western Connecticut State University, Farmington River Watershed Association, The Last Green Valley, and Interstate Environmental Commission were included in these interviews.

2.3.1 Results of Data Generator Interviews (n = 7)

While each interview was a back and forth discussion, a specific list of questions were asked to stimulate discussion, presented in the following subsections. Highlights from the responses are included.

2.3.1.1 Data Sharing and Communication (Data Generator Interviews, n = 7)

- Question asked: What are your goals for data sharing and communication?
  - See Table 4 for responses.
- All respondents indicated that creating easy-to-share maps or visualizations was extremely important, for both outreach and in-house communications.
- Automated formatting of data for upload to EPA's WQX was also highly ranked, along with reducing the number of times they need to reformat data for different outlets and uses.

Table 4. Respondents ranked their goals for data sharing and communication. In the "respondents" columns, darker green colors indicate greater importance to the respondents and blank spots indicate the question is not relevant to the organization. In the column showing the average of all respondent scores, the lengths of the blue bars correspond with the value of the averages.

What are your goals for data sharing and communication? (1 = not important, 5 = most important)	Average	Respondents						
		A	B	C	D	E	F	G
Create easy-to-share maps or visualizations that demonstrate the health of our waters to better communicate with lay audiences	5.00	5	5	5	5	5	5	5
Create easy-to-share maps or visualizations as an in-house communication tool to communicate on our work and the health of our water resources	4.57	5	5	5	5	5	5	2
Automated formatting of data for uploading to EPA's WQX	4.57	5	5	3	5	5	4	5
Reduce the number of times data must be re-formatted to share with partners or uploaded to data portals	4.21	2	5	3	5	5	5	5
Streamline the data collection and management process to reduce staff time.	4.14	4	5	4	5	5	2	4
Streamline the data collection and management process to improve quality of data.	4.14	3	5	4	5	4	5	3
Use data to fundraise for our organization and our watershed	3.67	5	4	4	3		4	2
Maintain ownership of data	3.57	3	4	4	1	5	5	3

2.3.1.2 Data Sharing (Data Generator Interviews, n = 7)

- Question Asked: Where or with whom would you like to share your data?
  - See Table 5 for responses.
- Most organizations are already sharing their data though annual reports and about half share their data through their organization's website and with local municipalities and State agencies.
- Many respondents would like to share their data to EPA WQX and would like to begin or improve their sharing of data with local municipalities.

Table 5. Respondents indicated where they currently share their data (check mark) and where they would like to share their data (raised hand). Presence of both symbols indicates the group currently shares to some extent, but would like to share more. A blank space indicates they do no currently share their data with that outlet, nor is it a priority. For the Unified Water Study, a check mark indicates their organization is currently part of the UWS, where all participants are required to share their data; a raised hand would indicate a desire to join the UWS. Data are sorted based on the number of raised hands, highlighting areas needing development and support.

Where have you committed to sharing your data and where would you like to share your data?								
(✓ = committed, ✋ = would like to)	**Total (#✓- #✋)**	Respondents						
		A	B	C	D	E	F	G
Federal databases (e.g. WQX)	3 - 4	✋		✋	✋	✓	✓✋	✓
Local municipalities	4 - 4	✋	✓	✓✋		✓✋	✋	✓
State agencies (e.g. CT DEEP, NYS DEC)	all - 1	✓	✓	✓✋	✓	✓	✓	✓
Organization's annual report	5 - 1	✓	✓	✋	✓	✓	✓	
Organization's website	4 - 1	✓		✓	✋		✓	✓
Other local, regional, or national, non-federal databases (e.g. URI Watershed Watch)	4 - 0	✓	✓	✓	✓			
Unified Water Study (UWS)	3 - 0	✓	✓					✓

2.3.1.3 Measuring Success (Data Generator Interviews, n = 7)

- Question asked: How do you measure the "success" of your water quality monitoring?
  - See Table 6 for results.
- Almost all groups already track water quality trends as an indicator or plan to soon.
- Four of the seven groups track social media indicators and membership rates.
- Four groups would be interested in an option or tool to track the number of times their data are downloaded.
- Two groups currently track visits to their website, one group is planning to implement this, and two groups would be interested in this tool if it became available.
- Most groups use the LIS Report Card in outreach efforts and plan to use the UWS Report Card once it is available.

Table 6. Respondents reported on the indicators they use to measure success as currently using (check mark), planned (raised hand), possible metric they would add if available (raised hand), and not interested in this metric ("x"). A blank space indicates no response. Data are sorted based on popularity of a metric for development (# of raised hands). For the question on the UWS, darker green colors indicate greater likelihood. Data are sorted based on the number of responses indicating use of a metric is planned (empty checkbox) plus indication if interest in adding a metric (asterisk), highlighting areas needing development and support.

What indicators do you use to measure 'success' for your water quality monitoring program?								
(✓ = currently using, ✋ = would like to, ✗ = not interested)	**Total (#✋ - #✓)**	Respondents A	B	C	D	E	F	G
Number of times your data are downloaded	4 - 0	✋	✋	✋		✗	✗	✋
Visits to your website to view monitoring map	3 - 2	✓	✗	✋	✓	✋	✗	✋
Improved water quality trends annually	2 - 5	✋	✓	✓	✓	✋	✓	✓
Number of media hits reporting on, sharing, or discussing your monitoring efforts	1 - 4	✓	✓	✋	✓		✓	✗
Increased organization membership rates	0 - 4	✓	✓		✓	✓		
Visits to Social Media	0 - 1					✓		
Do you use the Long Island Sound Report Card when communicating to your constituents?	0 - 5	✓	✓		✓	✗	✓	✓
Will you use UWS report card? (1-5, with 5 "definitely")		5	5	5	2	5	3	4

2.3.1.4 Data Collection (Data Generator Interviews, n = 7)

- Question asked: How do you collect your data in the field?
    - See Table 7 for results.
- All groups use paper data sheets (or field books) and will continue to do this into the future because of concerns about the potential for losing data collected solely in digital form.
- Six of the seven groups were interested in a digital data collection form.

2.3.1.5 Management Tools Used (Data Generator Interviews, n = 7)

- Question asked: What tools do you use to manage your data?
    - See Table 7 for results.
- Of the organizations interviewed, some employ citizen scientists or interns to collect data. Participants collect the data in the field, and then pass the data sheet to someone in the organization who is responsible for data entry.
- Data are entered into an online form (5 out of 7 respondents) or into a form that is stored in the cloud (4 out of 7 respondents).

2.3.1.6 Visualization, Analysis, and Sharing (Data Generator Interviews, n = 7)

- Question asked: What tools do you use to visualize, analyze, and share your data?
    - See Table 7 for results.
- Excel is used by everyone surveyed.
- ArcGIS is used by 5 of the 7 groups with the remaining 2 wanting to add this functionality, if a tool were made available.

Table 7. Respondents indicated how they currently manage and analyze data (check mark) and applications they would like to use if the process was made easier (raised hand). One digital data collection form used by respondents and noted as being very good was Terrapin Trackers; others that were considered "OK" were iNaturalist and Frog Watch. Other tools mentioned included Google Sheets linked to Infogram by Prezi, CT DEEP's Microsoft Excel data entry template for bacteria, Anecdata, Microsoft Access, and R. All groups would continue to maintain paper copies but most would like to improve the ease of data entry, management, upload to databases, analysis, and visualizations.

	Total (#✓- #✋)	**Respondents**						
		A	**B**	**C**	**D**	**E**	**F**	**G**
How do you document and collect your water quality monitoring data in the field?								
Paper Sheet	all	✓	✓	✓	✓	✓	✓	✓
Digital Data Collection Forms	2 - 4	✓		✋	✋	✋	✓	✋
What tools/workflows do you use to consolidate and manage your water quality monitoring data?								
Participants hand deliver, mail in, or scan & email paper sheets (do not enter data).	6 - 0	✗	✓	✓	✓	✓	✓	✓
Participants enter data into an online data form (not stored on local computer).	5 - 0	✓		✓	✓		✓	✓
Data manager enters data into an online data form (not stored on local computer).	4 - 0	✓		✓	✓		✓	
Participants enter data into a data sheet stored on a local computer, then transfer.	3 - 0	✗	✓	✓		✓		
Data manager enters data into a data sheet or database stored on a local computer.	2 - 0	✗				✓	✓	
What tools do you use to visualize, analyze, and share your data?								
Excel and Excel graphs	all	✓	✓	✓	✓	✓	✓	✓
ArcGIS desktop	5 - 2	✓	✋	✓	✋	✓	✓	✓
ArcGIS web-application	2 - 2	✓		✓	✋			✋
custom built database	3 - 1	✓		✋			✓	✓
custom built visualization/ web-application	1 - 0	✓					✗	

2.3.1.7 Content Delivery (Data Generator Interviews, n = 7)

- Question asked: How do you deliver content to your stakeholders?
  - See Table 8 for responses.
- Email, website, and in-person were methods used by all groups.
- Social media was used by 5 out of 7 groups, with direct mailing only used by 2 groups.
- In general, groups use oral presentations, newsletter stories, social media posts, poster presentations, and annual reports as the most prevalent style of communication.

Table 8. Respondents indicated how they deliver content and the form of the communication product. A check mark indicates current use; a raised hand indicates they would like to do this in the future. Data are sorted based on the total number of affirmative responses.

	Total	Respondents						
		A	B	C	D	E	F	G
How do you deliver your content?								
Email	all	✓	✓	✓	✓	✓	✓	✓
Website	all	✓	✓	✓	✓	✓	✓	✓
In-person	all	✓	✓	✓	✓	✓	✓	✓
Social Media	5		✓	✓	✓	✓	✓	
Direct Mailings	2		✓		✓			
What style or type of communications do you use?								
Oral presentation	all	✓	✓	✓	✓	✓	✓	✓
Newsletter story	6	✓	✓		✓	✓	✓	✓
Social-Media posts	5		✓	✓	✓	✓	✓	
Poster presentation	5	✓	✓	✓	✓		✓	
Detailed Annual Reports	5	✓	✓		✓		✓	✓
Highlights of an Annual Report	5	✋	✓		✓	✓		✓
Flyer or brochure	4	✓	✓		✓		✓	
Website blog or post	4	✓	✓	✓	✓			
Online data explorer	3	✓		✓			✓	
Website story map	1			✓				

2.3.1.8 Desired Stakeholder Action (Data Generator Interviews, n = 7)

- Question asked: What do you hope your stakeholders will do with your data?
  - See Table 9 for responses.
- Becoming a vocal supporter of the environment to politicians, at local meetings, and to friends and neighbors was a priority for most groups, with all three target audiences being roughly equivalent in importance.
- Stimulating people to join the organization as a volunteer was also highly ranked.
- Stimulating financial contributions was important for the groups who relied on donations to supplement their budget; some groups do not take donations.

Table 9. Respondents ranked their objectives for stimulating action. In the "respondents" columns, darker green colors indicate greater importance to the respondents and blank spots indicate the question is not relevant to the organization. In the column showing the average of all respondent scores, the lengths of the blue bars correspond with the value of the averages.

If you use your data to support a call to action, what action do you wish your audience would take? (1=not important, 5=important)	Average	Respondents						
		A	B	C	D	E	F	G
Become a vocal supporter to politicians	4.67	5	5	4		5	5	4
Join as a volunteer	4.50	5	5	5	3	5	4	
Become a vocal supporter at local meetings	4.43	5	5	5	5	5	5	1
Become a vocal supporter to neighbors and friends	4.43	5	5	3	5	5	5	3
Financially contribute to my program	4.25	5	5	4	3			

2.3.1.9 Program Needs (Data Generator Interviews, n = 7)

- Question asked: What needs exist in your monitoring program?
  - See Table 10 for responses.
- Almost all groups indicated they have *insufficient staff time* to manage and analyze data.
- Most groups could use greater support in analyzing data as they feel they lack the *in-house expertise* to do more advanced analyses of the data.

Table 10. Respondents ranked their perceived gaps. In the "respondents" columns, darker green colors indicate greater importance to the respondents and blank spots indicate the question is not relevant to the organization. In the column showing the average of all respondent scores, the lengths of the blue bars correspond with the value of the averages.

What perceived gaps exist with regards to operating your monitoring program? (1=not important, 5=important)	Average	Respondents						
		A	B	C	D	E	F	G
Insufficient staff time available to manage data	4.57	4	5	5	5	5	4	4
Insufficient staff time available to analyze data	4.29	4	5	2	5	5	5	4
Insufficient in-house expertise to manage data	4.29	3	5	2	5	5	5	5
Insufficient funds to support monitoring activities in the field	4.00	4	5	5	5		4	1
Insufficient in-house expertise to analyze data	4.00	2	5	1	5	5	5	5
Insufficient pool of participants to collect data	2.86	3	1	3	2	5	3	3
Data coming in from multiple locations slows down or complicates the management process	1.67	1		2				2

2.3.1.10 Question asked: What audiences do you seek to reach with your data, and what messages are you sharing? (Data Generator Interviews, n = 7)

- *Key Audiences:*
  - Teachers, students, both directly and indirectly through curriculum development
  - Government (from local to Federal)
  - Some groups expressed that having multiple audiences can create challenges for them
  - The general public
  - Other non-profits and partner groups
- *Key Messages:*
  - Health of the Sound or particular areas within the Sound
  - Which processes are natural and which are not (i.e., climate change)
  - How individuals can play a role in protecting the Sound
  - Public health concerns related to pollution

2.3.1.11 Question asked: What would you like to see in a new data tool or platform to support your program? (Data Generator Interviews, n = 7)

- Interactive maps (emphasized by many interviewees)
- Real-time sharing of observations and data
- Multivariate visualizations
- Ability to format data for WQX entry
- Integration with ArcGIS and other software
- Ability to retrieve data from a given group's website
- Ease of use is key (data entry and retrieval); needs to be streamlined and enable new users within and across organizations to be able to pick it up with little training
- Designated point of contact for questions and support
- Ability to make comparisons across space and time
- Real-time sharing of observations and data
- Tools to better integrate data from other sources (e.g. temperature, tidal stage, precipitation, etc.) with a group's own data
- Interested in having ability to enter data but restrict publishing to the public until organization is ready to do so (e.g., completion of QA/QC processes, sensitive data, or any other reasons)
- Strong metadata required
- Needs to have good alignment with other efforts (e.g., Blue Plan, LIQWIDS, Unified Water Study)

3 End Users of Long Island Sound Data: Needs and Priorities

GOAL: The present study sought to assess the needs of the management, research, and advocacy communities around Long Island Sound.

METHODS: Interviews of representative end users of data were conducted by phone in the fall and early winter of 2019-2020. The individuals contacted for these interviews represented government entities (State and Federal), non-profit organizations, and consulting firms. One of the non-profit representatives interviewed was from a national organization that had addressed similar issues in the past, but who does not work in the Long Island Sound region, and so their comments are broken out separately. These interviews were designed to explore the needs and priorities of end users of Long Island Sound data. Interviews (n = 7) were conducted with representatives from the State of New York, the State of Connecticut, River Network, Save the Sound, the Environmental Protection Agency, Tetra Tech, and NERACOOS.

The interviewers (the authors) used the following as a framework to facilitate discussion: "We are looking for ways to build a community among all of the various groups collecting data around Long Island Sound and assess what opportunities there are to make Long Island Sound data more broadly available for multiple uses. We'd like to ask you to think about data, both the data you collect at your department or organization and the data collected by others that you use for analysis, synthesis, policy, or management:

- What types of data do you use from around Long Island Sound in your work? Who is collecting it? How do you use it?
- How are those data stored and shared by the data generators? Once you get the data from the group collecting it, how are the data stored and shared by your organization?
- What criteria do you use when determining which data are going to be included in your analyses? Do you use any data collected by non-academic and non-governmental groups (like non-profits, municipalities, volunteer groups)? Would you?
- Do you have your own database? If yes, who can upload data? Who can download data? Any plans to build one or make any changes?
- How easy is it to find the data you want about the Sound? Where do you look?
- Do you think it would be valuable to have a shared, centralized database for LIS data? What would it need to offer for you to use it?
- Would you use data visualizations provided by a database interface, or are you more likely to create your own data visualizations? How do you currently visualize data?"

3.1 Data Currently Utilized by End Users (End User Interviews, n = 7)

Many types of data are being collected around Long Island Sound. These data are utilized by different end user groups based on their specific needs. Water quality parameters for both non-marine (lake/stream assessments) and marine systems included ambient parameters (e.g., temperature and salinity), and bacteria (e.g., *Enterococci*, *E. coli,* and fecal coliform). Additional data types mentioned during interviews included discharge, habitat sea floor mapping, biological data, physical and biogeochemical data, ocean forecast, macroinvertebrates, and

macroalgae. Some of the local data are rolled up into larger efforts, such as the National Coastal Assessment data, with data collected by the Connecticut Department of Energy and Environmental Protection (CT DEEP) and sent to the EPA for inclusion in the program.

The data being used are collected by a broad pool of sources, including CT DEEP, the Interstate Environmental Commission (IEC), NEIWPCC, New York Department of Environmental Conservation (DEC), local government partners (e.g., NY DEC uses a Suffolk County QAPP), as well as a range of locally or regionally-focused groups including non-profits (e.g., Save the Sound, Harbor Watch), volunteers, citizen scientists, and academic groups along the Sound in New York and Connecticut, from the East River to the Race in the Eastern Sound.

The interviewees shared a variety of end uses for those data, such as informing managers about additional monitoring needs or creating regional models of biotic or abiotic systems. A publicly visible end use that is critical to supporting a healthy environment is informing other organizations and enabling them to do analyses and syntheses of data; this in turn supports the establishment of environmental criteria and for the integration of these criteria and status reports on environmental quality into published official reports (e.g., “priority water body list”) and fact sheets. The end users interviewed stressed the need for high quality data for these uses, such as those from citable reports, collected following a strict quality assurance plan, and from “certified” (or “approved”) laboratories. They also emphasized the need for available raw data, not just reports and graphs generated by others.

3.2 Accessing Data and Storage of Data (End User Interviews, n = 7)

While processes varied among interviewees, there seemed to be a lack of an easy way to locate and obtain relevant data. Data were typically gathered informally by contacting individuals, who then shared data in various formats such as PDFs, database links, or (most commonly) Excel workbooks. Some data were posted on the individual websites of data generating groups, but most of this data transfer seems to be through informal relationships and direct requests to individuals from people in agencies and organizations who have personal knowledge of who is monitoring what and where.

For the end users of data, without a central repository, most data acquisition is through informal pathways; finding data requires considerable effort with follow-up and can involve high costs. Some end users noted differences in how readily available data tend to be among different types of organizations, or between organizations of similar type, such as Federal agencies, State agencies, academic researchers, citizen scientists, non-profits, or others. Several end users expressed the challenges faced when trying to find specific data, citing issues such as State beach data being unavailable publicly and municipal stormwater reporting data being unavailable in usable formats. Academic data were perceived as highly elusive.

Since many of the end users both aggregate data from multiple sources and collect their own data, many of these end users also have to address sharing their own data as well. One of the government groups interviewed currently shares data only through Freedom of Information Act (FOIA) requests and they have to manually pull those data when requested. The

Environmental Research Division's Data Access Program (ERDDAP) system was mentioned as one important tool, which is a National Oceanic and Atmospheric Administration (NOAA) server combining regional and national data.

Several challenges were shared related to data storage and sharing. Determining how to manage groups' data in a systematic way is a challenge, especially when data are used by an agency or organization but then not kept in a long term repository. Excel spreadsheets were mentioned by many of these end user groups as their primary tool for aggregating and storing data for other uses. End users shared that they do not retain nor maintain those data after their personal use for assessments or other end goals are completed.

Data are stored and shared independently by different groups using different formats and systems; some users reported hiring staff support to better organize and manage their database. Likewise, it can be challenging to find the best methods to manage and respond to different requests for data made by outside groups. Interviewees also mentioned a need to find ways to avoid data storage and sharing solutions from quickly becoming outdated.

3.3 Criteria for Data Included in Analysis (End User Interviews, n = 7)

We sought to understand what criteria the end users employed when determining which data sources could be included in their work. End users expressed preferring and "trusting" data with features/standards such as:

- ERDDAP compatibility
- An approved EPA-approved Quality Assurance Project Plan (QAPP), required for any work funded by the EPA
- Sample analysis conducted in Environmental Laboratory Accreditation Program (ELAP) approved laboratories or from State-certified laboratories (other than ELAP)
- Compatible with regulatory agencies' water quality standards
- Clear and complete metadata
- Groups with a proven track record of generating quality data (based on criteria listed above)
- Association with a State or City laboratory

The degree of openness to the use of "citizen science", "community-led science" or crowd-sourced data varied, but most users were open to data from a variety of sources provided that they met these (or similar) standards. One interviewee indicated they only use data from CT DEEP and academic data collectors because of concerns about the reliability of other data sources. However, others expressed that as long as groups are meeting their EPA QAPP requirements or other recognized standards, then data should be included and data collector type shouldn't matter.

One other issue raised concerned the consistency and knowledge in clearly identifying the methods, analytics, or units; it can be challenging to aggregate data or assess its potential utility to a given project with different groups using different naming conventions.

3.4 Current Databases and Tools (End User Interviews, n = 7)

The end users interviewed mentioned some existing tools and additional tools in development. There was significant discussion with all of the interviewees around WQX, and in particular the new tool being developed for Long Island called the Long Island Water Quality Integrated Data System (LIQWIDS). There was hope that a re-designed WQX will soon show data trends, data gaps, sampling location gaps, and more, but there was also acknowledgement that getting data into and out of WQX/WQP is really challenging. LIQWIDS is an interface for WQX that will help extract data from WQP in different ways. However, LIQWIDS cannot be used with other data sources without going through WQX. Through the development of LIQWIDS, a consultant will provide training to groups who need help getting the data into WQX, but a long term plan for support is still in the decision-making stage.

Most of the end users interviewed had their own discrete database or system in place, including custom databases (or several), Excel files complemented by R code, Access databases, or others. CT DEEP is currently building a database in-house; it is still undecided or unclear as to whether it will be available to community users. NYS DEC has dispersed databases in place, but these systems aren't integrated across data types. Many interviewees mentioned significant costs associated with data management, both in terms of dollars spent directly on database services (e.g., developers) as well as staff time.

The lack of cohesion among formats and compatibility of data among the generators and users of data requires ongoing effort both by individual data generators and by the managers and other end users who aggregate those data for various purposes. Concerns raised by these groups about their existing systems included the need for long-term stewardship or support for a private database, concerns that some existing systems are outdated and need to improve consistency, as well as the feeling that the Federal and State databases are neither well understood nor utilized by the public.

For visualization of data, most end users were producing outputs within their organization and not using any shared or regional tools. Some end users expressed an interest in new tools that could be used to visualize long-term trends in different parameters. At present, most interviewees were using static graphs and maps on their websites; it was noted that these are time consuming and costly to update. LIQWIDS, which is in development now, was raised as a potentially important tool for data visualization for Long Island. The Sound Health Explorer (Save the Sound) provides a regularly updated portal for visualization of bacteria and beach closure data, and pulls data from multiple public, regional sources. Most end users did not express concern about lacking in-house expertise for visualization, which was notably different from the concerns expressed by data generators on this topic.

3.5 Thoughts on a Shared, Centralized Database for LIS Data? (End User Interviews, n = 7)

Some end users, mainly those representing government agencies, suggested using WQX as the repository for everyone's data. There was acknowledgement that this suggestion would require ongoing support from one department or specific point of contact (e.g., related to version changes, technology changes) and would cause ongoing challenges as WQX changes. For those advocating for the use of WQX, they also expressed that if that were not the solution, it would be important for the solution to automate the import of data to WQX.

Most other interviewees were either open to or enthusiastic about the idea of a shared, central database and had helpful thoughts and suggestions on implementation. Some expressed that they would be looking for ease of accessibility to data collected by multiple community groups, especially groups focusing on embayments, where data available for broader analysis have historically been lacking.

Suggestions included the following:

- Well-developed search function for finding relevant data.
- Robust documentation of data (i.e., metadata) to allow other users to decide if the data can be used in a given analysis or application.
- Built-in display and analysis tools. Ability to display both regional summaries as well as drill down on specific, local, raw data.
- Interoperable with other databases.
- Make it a "win-win" and give groups something for providing their data, such as reporting and analysis tools. WQX/WQP does not currently "give anything" for the data provided. Whatever is built should be attractive to groups and make them want to submit data.
- Establish a high burden on constraints to uploading data, requiring groups to report out on methods, units, and assorted quality assurance metrics. The data intended for upload does not need an EPA-approved QAPP nor do analyses need to be conducted in an ELAP-certified lab, but it needs to be clear which standards included in these quality assurance frameworks have been met.
- End users want to be able to filter for data that have met certain thresholds or standards for quality of the data. Some specific concerns were expressed about how to filter or assess data already in WQX/WQP.
- Look at user-generated data and assess the easiest path to get all data into a central database at launch; explore whether spreadsheet templates, R code, or another method would facilitate initial entry into the system.

Concerns expressed by the end users included:

- Some end users expressed concern about maintaining the ability to present the data on their organizational website and not be exclusively redirected to a third-party website.
- Finding a design that allows others to upload data and make it an area for shared data.

- Water quality data are complex, and concern was expressed about further complexity that would be added through the inclusion of biological data such as zooplankton, macroinvertebrates, eelgrass, and other relevant data types.
- Need to consider how to address continuous versus discrete data records (e.g., every time stamp or hourly averages or some other time interval).
- Would one master database be preferable or multiple databases that can communicate easily with each other and facilitate an exchange of data?
- Ensure that the solution is adaptive and whatever gets built should be relevant for future methods and data collected at assorted spatial scales and time intervals.
- Need to determine the potential audience for both the database itself and its outputs.
- Need to determine how data will be cited or attributed to the groups collecting the data.

4 Review of Existing Tools

GOAL: Preliminary information was collected on twelve data management systems to see if any existing platforms would be well-suited to support the types of data being collected by monitoring organizations in and around Long Island Sound. This list of potential platforms is not exhaustive relative to all possible options available, but any mentioned during our survey and interview process were explored to determine whether they might be a good fit for this effort.

METHODS: A summary of each system can be found in Appendix 6.2 (page 67). Of the twelve data platforms reviewed, five stood out as potential contenders that may be appropriate for the quantity of data being collected and the number of potential platform users. These platforms are able to handle multiple users; have the ability to restrict users to certain tasks within the platform (data upload only, data reviewer, etc.); can handle different types of data such as discrete, time series, and photos; have customizable templates for bulk data uploading; can accommodate data validation and quality control metadata; and have visualization tools or capabilities to easily incorporate data into websites.

4.1 Recommended Platforms

The available options (listed alphabetically) that seemed best suited to the needs of this community were:

- Aquatics Informatics: AQUARIUS Samples and AQUARIUS Time-Series
- Chesapeake Commons: Water Reporter
- Geotech Systems: EnviroData and EnviroSpace
- Kisters: WISKI
- Locus: EIM

Each platform is summarized in the tables below, listed in alphabetical order.

Company	Aquatics Informatics
Platform	AQUARIUS Samples; AQUARIUS Time-Series
Website	https://aquaticinformatics.com/
Accessed	Cloud based (through web).
Data Supported	"Samples" supports discrete data; "Time-Series" supports continuous data. Two separate programs that can connect with each other.
Upload Type	Real-time or in batches using Excel files (rigid in format though).
Outputs/Visualizations	API - Can allow programs to talk to one another (open program, but not open source); Web developer can use API to bring in visualization tools to a website.
Pricing	Single user $6,000/year for Samples (more if adding time-series). Pricing is based on number of users.
Contact	Mike Casey
Notes	Based on review and phone meeting with company, seems like this option has a strong possibility of being compatible with the needs of many LIS monitoring groups.

Company	Chesapeake Commons
Platform	Water Reporter
Website	https://www.waterreporter.org/pricing/index.html
Accessed	Web-based.
Data Supported	Qualitative/Quantitative.
Upload Type	Photos and comments when using it more as a social network platform. Spreadsheets when using it as a database storage platform.
Outputs/Visualizations	Interactive embeddable maps; download report data.
Pricing	Free, $20/mo, or $40/mo options listed on website. Erin Hoffman said $29/mo subscription fee per organization and an optional $400 set-up fee.
Contact	John Dawes and Erin Hoffman
Notes	Potentially good option for compatibility with anticipated needs. At this time does not offer all the tools that some of the other platforms offer, but it seems like there are plans for additional tools (e.g., WQX output or all users/groups uploading to the same space).

Company	Geotech Systems
Platform	EnviroData, EnviroSpace
Website	https://geotech.com/products/envirodata
Accessed	Desktop, Client server, Cloud-based (choose one option).
Data Supported	Designed to handle many types of data (discrete, continuous, and non-numerical or categorical).
Upload Type	Fully customizable Excel imports; can make unique importers for each group based on the way they currently store data.
Outputs/Visualizations	Vbsite.net/mapdemo/asp (requires a "modest" license fee, puts data into Google Maps for anyone to access); Envirodata for the web has login but could omit so the public could access (tables of data, no graphs and limited visualization options). EnviroSpace is an ArcGIS compatible service to bring data into GIS for mapping.
Pricing	Startup costs would be > $10,000. Additional fees in the range of $1,000s. Pricing is based on the number of users.
Contact	Dave Rich
Notes	Seems like ease of use may be helpful for the different types/experience levels of users that make up LIS monitoring groups.

Company	Kisters
Platform	WISKI
Website	https://www.kisters.net/NA/products/wiski/
Accessed	Web based; have a field app for additional cost.
Data Supported	Discrete, continuous, and biological in one platform.
Upload Type	Excel templates, real-time component for time-series data.
Outputs/Visualizations	Graphs, tables, maps (multiple analytes on one graph, different years).
Pricing	$30,000-35,000 one-time fee, then annual maintenance fee about 20% of one-time fee. Non-profit discounts available. Concurrent licensing.
Contact	Steve Elgie
Notes	Currently used by Cape Cod Commission (see discussion in main text).

Company	Locus EIM
Website	https://locustec.com/applications/environmental-information-management/
Accessed	Web-based.
Data Supported	Primarily discrete, but possibly customizable for other types.
Upload Type	Excel files (using templates), manual upload, real-time in the field with access to internet through the web.
Outputs/Visualizations	Can add a public-facing system so anyone can visualize the data, not just the users.
Pricing	Estimated at $20,000 to set up and then an additional $15,000 per year. Pricing varies based on number of users and quantity of data.
Contact	Kate Stevenson
Notes	The platform is currently more focused on data related to compliance, but they want to be able to work with LIS monitoring groups and customization is possible. Pricing possibly negotiable.

4.2 Unsuitable Platforms

The other platforms that were reviewed and which seem less likely to meet the needs of the community included:

- Aquatic Informatics: WaterTrax
- Gulfbase
- Ocean Data View
- United Nations Environmental Data Explorer
- LIQWIDS (Long Island Water Quality Integrated Data System)
- The Water Quality Exchange and Water Quality Portal

A summary of each system can be found in Appendix 6.2 (page 67).

Aquatic Informatics: WaterTrax, Gulfbase, Ocean Data View, and United Nations Environmental Data Explorer were not designed for the types of data storage, sharing, and visualization identified as being needed by the community.

LIQWIDS (Long Island Water Quality Integrated Data System) is a new system in development in partnership between NYSDEC and USGS to consolidate public and private nitrogen data collected on Long Island Sound related to the Long Island Nitrogen Action Plan into one data sharing system. While it will be a useful tool for stakeholders, it is limited in geography to Long Island and by only using data already entered into WQX it will not meet some of the broader needs identified through this study. The system will retrieve data from the EPA Water Quality Portal, which will then be shared on a public facing website to show trends, maps, modeling

that predicts water quality, and water quality projects that are completed or in progress on Long Island. While this system is not compatible at present with all of the data being collected across Long Island Sound monitoring groups, the code will be available on GitHub and may be used as a model on which to build a data sharing system for all water quality data being collected in Long Island Sound. Opportunities for partnership and coordination with this effort may be possible in the future.

The Water Quality Exchange and Water Quality Portal currently do not seem to meet as many needs as other platforms reviewed. The process to gain access to the system and prepare data for uploading is technically complex and can be overwhelming for some potential users. Since entering data into WQX is a priority for Federal and State partners, during the review of the other platforms it was asked if the platform under review produced WQX-compatible outputs to make uploading data to WQX more efficient; all of those reviewed and listed as those "best suited" above had that capability or were working on creating it.

4.3 Interviews With Others Who Have Adopted Platforms

In addition to the reviews of each existing database, the Cape Cod Commission (CCC) was contacted in order to discuss their process as they worked to implement one database system for four different monitoring groups collecting data on Cape Cod. After researching many of the same platforms listed above, and determining it would be difficult to compel the groups to use WQX, the CCC began working with Kisters WISKI and reported being very satisfied with their choice. There are staff at the CCC who are database and computer "savvy" so they built their own data importing modules, as opposed to having Kisters design the modules. It took approximately six months to upload all the historical data from four partner organizations dating back to the 1990s. Each group required a customized import process due to the different data formats of the historical data, but once the import modules were made, the CCC indicated that the uploads were easy. The biggest hurdle they faced was rectifying sampling locations and names because some groups were sampling at the same location but using different naming conventions. The CCC reported being very happy with WISKI's ability to handle and track quality assurance and quality control procedures. It should be noted that the CCC has each monitoring group send them their quality assured data annually for the CCC to upload as opposed to each group uploading their own data. The CCC plans to design their own visual tools and not use the ones that are built into WISKI.

During our interview process, it was also suggested we reach out to River Network, which is a non-profit that assisted with a similar database development effort for monitoring groups in North Carolina. In the initiation of their effort to streamline data sharing and storage across groups, they framed the effort as a community-driven collaboration working with all stakeholders from the "boots on the ground" data collectors all the way up through the data-chain to the end users. This effort is called the "Water Data Collaborative." They engaged State partners at the project outset, to ensure the needs of key stakeholders would be met in whatever product was developed. As a result, the State now reportedly encourages data generators to use the collaborative system, which also translates the data for import to WQX.

This system fosters groups throughout the data sharing process so that redundant effort is minimized and groups are supported through the process of importing to WQX.

The interviewee from River Network had many helpful suggestions, which are paraphrased here:

- Don't develop in a vacuum. Assess what everyone is collecting and the protocols they are already using so we can meet the data generators where they live as opposed to having them fit into a system we've created without them.
- Allow groups to both "own" their own data, and be able to share it. In the River Network product, they opted for a distributed database rather than a cloud-based system so that groups could own and maintain local control over their data.
- The user experience was important, including the ability to manage volunteers alongside data.
- WQX translation was a very important aspect for all stakeholders.
- Visualizations are valuable to data generators as well as end users and motivate groups to participate. River Network outputs include WQX templates or tidy formats that are machine readable in Google Sheets, ArcOnline, or Google Data Studio. River Network can create dashboards which can be customizable for different users.

5 Conclusions and Recommendations for Next Steps

The work presented here shows that there is a need for the proposed effort, as well as a willingness to engage with a centralized database in the monitoring and water quality research community. Through surveys of local groups, we found that 75% of data generators want to share their data with the public, and yet only 22% said that their data are readily available for public download at present time (n = 39). We found that 49% of groups are not happy with their current process for entering, storing, and sharing their data (n = 39). However, since we asked groups about those processes as part of the same question, we can't discern from that whether it is the entering, storing, or sharing *specifically* that groups are not happy with. 68% of data generators would consider using a centralized database for Long Island Sound data if one existed (n = 39), indicating a high likelihood of engagement with this effort going forward.

We also found that most data end users are currently getting data through informal, person-to-person networks. Managers and other stakeholders need confidence in data quality to integrate new data sources. The database effort assessed here would enable these and other data end users to access new streams of data from sources around the Sound, increasing the amount of data available for management and decision making.

Recommendations to Consider in the Project's Next Steps:

- Groups are happy with data entry in Excel; groups would like to use this familiar format.
- Guidelines on best practices for groups on data management and quality control are needed.
- Support and tools are needed to increase engagement with the EPA Water Quality Exchange (WQX).
- Public data sharing is a priority for groups and not currently easy to do.
- Interest is high in visualization tools (mapping and graphing of individual groups' data), including real-time display options, and social media and ArcGIS compatibility. In the close-out webinar, 46 data-generators and end users were polled and out of the people responding to the question and who had an opinion, 90% indicated access to graphical output would incentivize them to use the database and 10% said it might incentivize them and 0% said no, it would not incentivize them (n = 30) (Appendix 6.3, page 73).
- Groups need a clear protocol for metadata development and a list of what information is required for inclusion of their data in the database, in order to meet the needs of different end users.
- Groups expressed concern over resource limitations that are unrelated to technology (*e.g.,* self-reported insufficiencies in staff time and expertise); thus, keeping required effort low will be key for group engagement.
- Need to define the geographic area to be served by this effort.
- All parameters should be supported; the database should allow groups to store all types of data.
- In the close-out webinar, 47 attendees were polled and out of the people responding to the question and who had an opinion, 67% of groups (n=27) indicated that the ability to

retrieve group-specific data directly from a group's own website is highly desired (Appendix 6.3, page 73).

- The database should be easy to search and data should be easily downloadable.
- The database program needs a designated point of contact for questions and support.

6 Appendices

6.1 Appendix - Data Generator Survey #1 (Fall 2019): Respondents

Table 11. Summary of Respondents to Survey #1

State	County or NYC	water body	Salt or Fresh Water	UWS? (Tier 1 or 2)	monitor water quality	monitor bacteria	Name of Organization
CT	many	harbors from Greenwich to Bridgeport, some eastern Long Island Sound area	salt		Y		The Maritime Aquarium at Norwalk
NY	unspecified	unspecified	salt		Y		Sierra Club, Long Island Group
CT, NY	New London	Fisher's Island Sound	salt		Y		Project Oceanology
NY	Suffolk	Fisher's Island Sound watershed - Hay Harbor	salt		Y		Fishers Island Seagrass Management Coalition
NY	Suffolk	Fisher's Island Sound watershed - West Harbor	salt		Y		Fishers Island Seagrass Management Coalition
CT, NY		Long Island Sound	salt			Y	Save the Sound
CT, NY	New London	Long Island Sound	salt		Y		Project Oceanology
CT	Fairfield	Long Island Sound, off Shippan Point	salt		Y		SoundWaters
CT	Fairfield	Long Island Sound, western	salt		Y		SoundWaters
NY	Westchester	Long Island Sound	salt		Y	Y	The Ursuline School
NY	Westchester	Long Island Sound	salt		Y		Purchase College (SUNY)

State	County or NYC	water body	Salt or Fresh Water	UWS? (Tier 1 or 2)	monitor water quality	monitor bacteria	Name of Organization
NY	Suffolk	Long Island Sound	salt		Y	Y	Cornell Cooperative Extension of Suffolk County
NY	Nassau	Long Island Sound, near Oyster Bay	salt		Y		The WaterFront Center
NY	Nassau	Long Island Sound, western	salt		Y	Y	SUNY Maritime College
NY	Nassau	Long Island Sound, western - narrows	salt		Y		Interstate Environmental Commission
CT	New London	Alewife Cove	salt	Tier 1	Y		New England Science & Sailing Foundation (NESS)
CT	New London	Mystic Harbor	salt		Y	Y	Clean Up Sound and Harbor (CUSH) & University of Rhode Island Watershed Watch (URIWW)
CT	New London	Mystic Harbor	salt	Tier 1	Y		Clean Up Sound & Harbors (CUSH)
CT	New London	Mystic Harbor watershed - Mystic River	salt		Y	Y	Clean Up Sound and Harbor (CUSH) & University of Rhode Island Watershed Watch (URIWW)

State	County or NYC	water body	Salt or Fresh Water	UWS? (Tier 1 or 2)	monitor water quality	monitor bacteria	Name of Organization
CT	New London	Mystic Harbor watershed - Mystic River	salt	Tier 1	Y		Clean Up Sound & Harbors (CUSH)
CT	New London	Mystic Harbor watershed - Pequotsepos Cove	salt		Y	Y	Clean Up Sound and Harbor (CUSH) & University of Rhode Island Watershed Watch (URIWW)
CT	New London	Niantic River estuary	salt	Tier 2	Y		Save the River - Save the Hills
CT	New London	Niantic River estuary	salt			Y	Save the River - Save the Hills
CT	New London	Niantic River watershed - Cranberry Meadow Brook	fresh		Y		Niantic River Watershed Committee
CT	New London	Niantic River watershed - Latimer Brook	fresh		Y		Niantic River Watershed Committee
CT	New London	Niantic River watershed - Oil Mill Brook	fresh		Y		Niantic River Watershed Committee
CT	New London	Niantic River watershed - Stony Brook	fresh		Y		Niantic River Watershed Committee
CT	New London	Stonington Harbor	salt		Y	Y	Clean Up Sound and Harbor (CUSH) & University of Rhode Island Watershed Watch (URIWW)

State	County or NYC	water body	Salt or Fresh Water	UWS? (Tier 1 or 2)	monitor water quality	monitor bacteria	Name of Organization
CT	New London	Stonington Harbor	salt	Tier 1	Y		New England Science & Sailing Foundation (NESS)
CT	New London	Thames River watershed, rivers, lakes and streams upstream of Thames River	fresh		Y	Y	The Last Green Valley
CT	New London	Wequetequock Cove	salt		Y	Y	Clean Up Sound and Harbor (CUSH) & University of Rhode Island Watershed Watch (URIWW)
CT	New London	Wequetequock Cove	salt	Tier 1	Y		Clean Up Sound & Harbors (CUSH)
CT	Middlesex	Connecticut River watershed - Eightmile River and associated tributaries	fresh		Y	Y	Eightmile River Wild & Scenic Watershed
CT	Middlesex	Connecticut River watershed - Otter Creek	fresh			Y	Save the Sound
CT	Hartford, Middlesex	Connecticut River	salt, fresh	Tier 1	Y		Connecticut River Conservancy
CT	Hartford, Middlesex	Connecticut River watershed - Connecticut River	fresh		Y	Y	Connecticut River Conservancy

State	County or NYC	water body	Salt or Fresh Water	UWS? (Tier 1 or 2)	monitor water quality	monitor bacteria	Name of Organization
CT	Hartford	Connecticut River watershed - Farmington River	fresh		Y	Y	Northwestern Connecticut Community College
CT	Hartford	Connecticut River watershed - Farmington River and multiple stream and river tributaries	fresh		Y	Y	Farmington River Watershed Association
CT	New Haven	Farm River	salt	Tier 1	Y		Friends of the Farm River Estuary
CT	New Haven	Milford Harbor watershed - Race Brook	fresh			Y	Town of Orange Health Department
CT	New Haven	Milford Harbor watershed - Silver Brook	fresh			Y	Town of Orange Health Department
CT	New Haven	Milford Harbor watershed - Wepawaug River	fresh			Y	Town of Orange Health Department
CT	Litchfield, Fairfield, New Haven	Housatonic River watershed - Lake Lillinonah	fresh		Y		Citizen Led Environmental Observatory (CLEO)
CT	Litchfield, Fairfield, New Haven	Housatonic River watershed - Lake Lillinonah	fresh		Y		Friends of the Lake

State	County or NYC	water body	Salt or Fresh Water	UWS? (Tier 1 or 2)	monitor water quality	monitor bacteria	Name of Organization
CT	Litchfield	Housatonic River watershed - Bantam Lake and watershed	fresh		Y		The White Memorial Conservation Center, Inc. and Bantam Lake Protective Association
CT	Fairfield	Ash Creek	salt		Y		Bridgeport Aquaculture School
CT	Fairfield	Ash Creek	salt		Y		Mill River Wetland Committee
CT	Fairfield	Ash Creek watershed - Rooster River	fresh		Y	Y	Town of Fairfield
CT	Fairfield	Black Rock Harbor	salt			Y	Save the Sound
CT	Fairfield	Black Rock Harbor	salt	Tier 1	Y		Ash Creek Conservation Association, Inc.
CT	Fairfield	Black Rock Harbor	salt		Y		Bridgeport Aquaculture School
CT	Fairfield	Black Rock Harbor	salt		Y		Bridgeport Aquaculture School
CT	Fairfield	Byram River	salt, fresh			Y	Save the Sound
CT	Fairfield	Byram River	salt, fresh		Y		Town of Greenwich
CT	Fairfield	Byram River watershed - Pemberwick Creek	fresh			Y	Save the Sound
CT	Fairfield	Cove Harbor	salt	Tier 1	Y		Town of Darien
CT	Fairfield	Darien Harbor	salt	Tier 1	Y		Town of Darien

State	County or NYC	water body	Salt or Fresh Water	UWS? (Tier 1 or 2)	monitor water quality	monitor bacteria	Name of Organization
CT	Fairfield	Fairfield County rivers (many)	salt, fresh		Y	Y	Harbor Watch
CT	Fairfield	Greenwich Harbor	salt			Y	Save the Sound
CT	Fairfield	Greenwich Harbor watershed - Horseneck Brook	fresh			Y	Save the Sound
CT	Fairfield	Holly Pond	salt		Y		SoundWaters
CT	Fairfield	Holly Pond	salt	Tier 1	Y		SoundWaters
CT	Fairfield	Holly Pond	salt		Y		SoundWaters
CT	Fairfield	Housatonic River	salt	Tier 1	Y		Town of Stratford
CT	Fairfield	Housatonic River watershed - Candlewood Lake	fresh		Y		Brookfield WPCA
CT	Fairfield	Housatonic River watershed - Candlewood Lake	fresh		Y	Y	Western Connecticut State University
CT	Fairfield	Housatonic River watershed - Lake Housatonic and all tributaries	fresh		Y	Y	Western Connecticut State University
CT	Fairfield	Housatonic River watershed - Still River	fresh		Y		Brookfield WPCA

State	County or NYC	water body	Salt or Fresh Water	UWS? (Tier 1 or 2)	monitor water quality	monitor bacteria	Name of Organization
CT	Fairfield	Housatonic River watershed - Still River	fresh		Y	Y	Northwestern Connecticut Community College
CT	Fairfield	Housatonic River watershed - Still River	fresh		Y	Y	Western Connecticut State University
CT	Fairfield	Mianus River	salt, fresh			Y	Save the Sound
CT	Fairfield	Mianus River	salt, fresh		Y		Town of Greenwich
CT	Fairfield	Norwalk Harbor	salt		Y		The Maritime Aquarium at Norwalk
CT	Fairfield	Norwalk Harbor	salt		Y	Y	Harbor Watch
CT	Fairfield	Norwalk Harbor	salt		Y	Y	Western Connecticut State University
CT	Fairfield	Norwalk Harbor watershed - Norwalk River	fresh		Y	Y	Western Connecticut State University
CT	Fairfield	Norwalk Harbor, inner	salt	Tier 2	Y		Harbor Watch
CT	Fairfield	Norwalk Harbor, outer	salt	Tier 2	Y		The Maritime Aquarium at Norwalk
CT	Fairfield	Sasco Brook	salt, fresh		Y	Y	Town of Fairfield
CT	Fairfield	Saugatuck River estuary	salt		Y	Y	Harbor Watch
CT	Fairfield	Scott Cove	salt	Tier 1	Y		Town of Darien
CT	Fairfield	Southport Harbor / Mill River	salt	Tier 1	Y		Town of Fairfield

State	County or NYC	water body	Salt or Fresh Water	UWS? (Tier 1 or 2)	monitor water quality	monitor bacteria	Name of Organization
CT	Fairfield	Southport Harbor / Mill River	salt		Y	Y	Town of Fairfield
CT	Fairfield	Southport Harbor / Mill River	salt		Y		Mill River Wetland Committee
CT	Fairfield	Stamford Harbor	salt	Tier 1	Y		SoundWaters
CT	Fairfield	Stamford Harbor	salt		Y		SoundWaters
NY	Westchester	Larchmont Harbor	salt			Y	Save the Sound
NY	Westchester	Mamaroneck Harbor	salt	Tier 2	Y		Derecktor Shipyards
NY	Westchester	Mamaroneck Harbor	salt			Y	Save the Sound
NY	Westchester	Mamaroneck Harbor watershed - Beaver Swamp Brook	fresh			Y	Save the Sound
NY	Westchester	Mamaroneck Harbor watershed - Guion Creek	fresh			Y	Save the Sound
NY	Westchester	Mamaroneck Harbor watershed - Mamaroneck River	fresh			Y	Save the Sound
NY	Westchester	Mamaroneck Harbor watershed - Sheldrake Lake	fresh			Y	Save the Sound
NY	Westchester	Mamaroneck Harbor watershed - Sheldrake River	fresh			Y	Save the Sound
NY	Westchester	Milton Harbor	salt			Y	Save the Sound

State	County or NYC	water body	Salt or Fresh Water	UWS? (Tier 1 or 2)	monitor water quality	monitor bacteria	Name of Organization
NY	Westchester	Milton Harbor watershed - Blind Brook	fresh			Y	Save the Sound
NY	Westchester	New Rochelle Harbor	salt			Y	Save the Sound
NY	Westchester	New Rochelle Harbor	salt	Tier 1	Y		Save the Sound
NY	Westchester	Playland Lake	fresh			Y	Save the Sound
NY	Westchester	Premium Mill Pond watershed - Premium River	salt			Y	Save the Sound
NY	NYC	East River	salt		Y	Y	SUNY Maritime College
NY	NYC	East River	salt			Y	NYC Water Trail Association
NY	NYC	East River - Bronx River	salt	Tier 1	Y		Bronx River Alliance
NY	NYC	East River - Bronx River	salt		Y	Y	Bronx River Alliance
NY	NYC	East River - Flushing Bay	salt		Y	Y	York College of CUNY
NY	NYC	Eastchester Bay	salt			Y	Save the Sound
NY	NYC	Eastchester Bay	salt	Tier 2	Y		Save the Sound
NY	NYC	Eastchester Bay watershed - Hutchinson River	fresh			Y	Save the Sound
NY	NYC	Hudson River	salt			Y	NYC Water Trail Association
NY	NYC	Hunter Island Bay	salt			Y	Save the Sound
NY	NYC	Hunter Island Bay	salt	Tier 1	Y		Save the Sound
NY	NYC	Jamaica Bay	salt			Y	NYC Water Trail Association
NY	NYC	Jamaica Bay	salt		Y	Y	York College of CUNY

State	County or NYC	water body	Salt or Fresh Water	UWS? (Tier 1 or 2)	monitor water quality	monitor bacteria	Name of Organization
NY	NYC	New York Harbor	salt			Y	NYC water trail association
NY	NYC	New York Harbor tributaries and outfalls	salt			Y	NYC Water Trail Association
NY	Suffolk	Flax Pond	salt		Y		The Friends of Flax Pond
NY	Suffolk	Goldsmith Inlet	salt	Tier 1	Y		Group for the East End
NY	Suffolk	Huntington-Northport Bay Complex	salt	Tier 2	Y		Cornell Cooperative Extension of Suffolk County
NY	Suffolk	Huntington-Northport Bay Complex	salt		Y	Y	Cornell Cooperative Extension of Suffolk County
NY	Suffolk	Mattituck Creek	salt	Tier 1	Y		Group for the East End
NY	Suffolk	Nissequogue River	salt	Tier 1	Y		SWAN - Salonga Wetland Advocates Network
NY	Suffolk	Nissequogue River mouth - Sunken Meadow Creek	salt		Y		Sachem North HS
NY	Suffolk	Nissequogue River Mouth, Kings Park Bluff	salt		Y		A Day in the Life of a River
NY	Suffolk	Port Jefferson Harbor	salt	Tier 1	Y		Setauket Harbor Task Force
NY	Suffolk	Stony Brook Harbor	salt		Y		The Friends of Flax Pond

State	County or NYC	water body	Salt or Fresh Water	UWS? (Tier 1 or 2)	monitor water quality	monitor bacteria	Name of Organization
NY	Suffolk	Stony Brook Harbor mouth - West Meadow Creek	salt		Y		The Friends of Flax Pond
NY	Suffolk	Suffolk County waters	salt		Y	Y	Cornell Cooperative Extension of Suffolk County
NY	Nassau	Cold Spring Harbor	salt		Y		The WaterFront Center
NY	Nassau	Hempstead Harbor	salt		Y	Y	Coalition to Save Hempstead Harbor
NY	Nassau	Hempstead Harbor	salt	Tier 1	Y		Coalition to Save Hempstead Harbor
NY	Nassau	Hempstead Harbor	salt		Y		Interstate Environmental Commission
NY	Nassau	Little Neck Bay	salt			Y	Save the Sound
NY	Nassau	Little Neck Bay	salt				Douglas Manor Environmental Association
NY	Nassau	Little Neck Bay	salt		Y		Interstate Environmental Commission
NY	Nassau	Little Neck Bay	salt	Tier 2	Y		Interstate Environmental Commission
NY	Nassau	Little Neck Bay watershed - Gabblers Creek	fresh			Y	Save the Sound
NY	Nassau	Little Neck Bay watershed - Udalls Mill Pond	fresh			Y	Save the Sound

State	County or NYC	water body	Salt or Fresh Water	UWS? (Tier 1 or 2)	monitor water quality	monitor bacteria	Name of Organization
NY	Nassau	Little Neck Bay watershed - Alley Creek	fresh			Y	Save the Sound
NY	Nassau	Manhasset Bay	salt		Y		Interstate Environmental Commission
NY	Nassau	Manhasset Bay	salt	Tier 1	Y		Interstate Environmental Commission
NY	Nassau	Manhasset Bay	salt		Y	Y	Manhasset Bay Protection Committee
NY	Nassau	Oyster Bay	salt		Y		The WaterFront Center
NY	Nassau	Oyster Bay complex	salt	Tier 1	Y		Friends of the Bay
NY	Nassau	West Harbor	salt		Y		The WaterFront Center

6.2 Appendix - Existing Data Platform Review: Summaries

Company	Aquatics Informatics
Platform	AQUARIUS Samples; AQUARIUS Time-Series
Website	https://aquaticinformatics.com/
Accessed	Cloud based (through web).
Data Supported	"Samples" supports discrete data; "Time-Series" supports continuous data. Two separate programs that can connect with each other.
Upload Type	Real-time or in batches using Excel files (rigid in format though).
Outputs/Visualizations	API - Can allow programs to talk to one another (open program, but not open source); Web developer can use API to bring in visualization tools to a website.
Pricing	Single user $6,000/year for Samples (more if adding time-series). Pricing is based on number of users.
Contact	Mike Casey
Notes	Based on review and phone meeting with company, seems like this option has a strong possibility of being compatible with the needs of many LIS monitoring groups.

Company	Aquatics Informatics
Platform	WaterTrax
Website	https://watertrax.com/
Accessed	Not applicable.
Data Supported	Not applicable.
Upload Type	Not applicable.
Outputs/Visualizations	Not applicable.
Pricing	Not applicable.
Contact	Customer service representative (name unknown).
Notes	Mainly compliance based and for municipality use. Agreed that AQUARIUS Samples would be a better fit.

Company	Chesapeake Commons
Platform	Water Reporter
Website	https://www.waterreporter.org/pricing/index.html
Accessed	Web-based.
Data Supported	Qualitative/Quantitative.
Upload Type	Photos and comments when using it more as a social network platform. Spreadsheets when using it as a database storage platform.
Outputs/Visualizations	Interactive embeddable maps; download report data.
Pricing	Free, $20/mo, or $40/mo options listed on website. Erin Hoffman said $29/mo subscription fee per organization and an optional $400 set-up fee.
Contact	John Dawes and Erin Hoffman
Notes	Potentially good option for compatibility with anticipated needs. At this time does not offer all the tools that some of the other platforms offer, but it seems like there are plans for additional tools (e.g., WQX output or all users/groups uploading to the same space).

Organization	Environmental Protection Agency
Platform	Water Quality Exchange (WQX); Water Quality Portal (WQP)
Website	https://www.epa.gov/sites/production/files/2015-11/documents/wqx_factsheet.pdf
Accessed	Web-based
Data Supported	Physical, chemical, biological, habitat, index, and metrics.
Upload Type	Excel files based on a very specific format.
Outputs/Visualizations	Can download into a CSV, tab-separated, or excel file. No visualization tools available.
Pricing	Free
Contact	Not applicable.
Notes	Cumbersome process for setting up an account. Two or three step registration. First the organization needs to have an ID. If no ID is already available, then contact WQX helpdesk to set up the Organization ID. Then the user needs to register with CDX and WQX following the steps here: https://www.epa.gov/waterdata/wqx-web-account-registration. WQX gives you a user name. CDX must verify you. Then you have to call to get a key to finish registering. Emails from system sent with moderate delays; overall, took three days to just complete registration. Appears you have to login to WQX to upload through CDX. Webinar videos do not match up to the templates that are available now, which may prove to be difficult to interpret when trying to upload data.

Company	Geotech Systems
Platform	EnviroData, EnviroSpace
Website	https://geotech.com/products/envirodata
Accessed	Desktop, Client server, Cloud-based (choose one option).
Data Supported	Designed to handle many types of data (discrete, continuous, and non-numerical or categorical).
Upload Type	Fully customizable Excel imports; can make unique importers for each group based on the way they currently store data.
Outputs/Visualizations	Vbsite.net/mapdemo/asp (requires a "modest" license fee, puts data into Google Maps for anyone to access); Envirodata for the web has login but could omit so the public could access (tables of data, no graphs and limited visualization options). EnviroSpace is an ArcGIS compatible service to bring data into GIS for mapping.
Pricing	Startup costs would be > $10,000. Additional fees in the range of $1,000s. Pricing is based on the number of users.
Contact	Dave Rich
Notes	Seems like ease of use may be helpful for the different types/experience levels of users that make up LIS monitoring groups.

Company	Gulfbase
Website	Gulfbase.org
Accessed	Not applicable.
Data Supported	Not applicable.
Upload Type	Not applicable.
Outputs/Visualizations	Not applicable.
Pricing	Not applicable.
Contact	Not applicable.
Notes	Gulfbase itself is a free resource database for Gulf of Mexico-related information but does not specifically house scientific data. GRIIDC does that aspect, but at this time citizen science is not incorporated. At this time, Gulfbase does not seem to be a relevant resource for LIS monitoring groups to use as a model.

Company	Kisters
Platform	WISKI
Website	https://www.kisters.net/NA/products/wiski/
Accessed	Web based; have a field app for additional cost.
Data Supported	Discrete, continuous, and biological in one platform.
Upload Type	Excel templates, real-time component for time-series data.
Outputs/Visualizations	Graphs, tables, maps (multiple analytes on one graph, different years).
Pricing	$30,000-35,000 one-time fee, then annual maintenance fee about 20% of one-time fee. Non-profit discounts available. Concurrent licensing.
Contact	Steve Elgie
Notes	Currently used by Cape Cod Commission (see discussion in main text).

Company	Locus EIM
Website	https://locustec.com/applications/environmental-information-management/
Accessed	Web-based.
Data Supported	Primarily discrete, but possibly customizable for other types.
Upload Type	Excel files (using templates), manual upload, real-time in the field with access to internet through the web.
Outputs/Visualizations	Can add a public-facing system so anyone can visualize the data, not just the users.
Pricing	Estimated at $20,000 to set up and then an additional $15,000 per year. Pricing varies based on number of users and quantity of data.
Contact	Kate Stevenson
Notes	The platform is currently more focused on data related to compliance, but they want to be able to work with LIS monitoring groups and customization is possible. Pricing possibly negotiable.

Company	Ocean Data View
Website	https://odv.awi.de/
Accessed	Not applicable.
Data supported	Not applicable.
Upload type	Not applicable.
Outputs/Visualizations	Not applicable.
Pricing	Not applicable.
Contact	Not applicable.
Notes	Does not meet the anticipated needs of this effort.

Company	United Nations Environmental Data Explorer
Website	https://geodata.grid.unep.ch/
Accessed	Web browser.
Data Supported	National, regional, and sub-regional statistics and geospatial data.
Upload Type	Does not appear that small groups can upload, only large organizations like WHO.
Outputs/Visualizations	Maps, graphs, tables, metadata, downloadable data sets. Simple visuals. Easy to manipulate.
Pricing	Not applicable.
Contact	Not applicable.
Notes	Not applicable to this effort. Resource for the public to use to track trends on large scale (regionally, sub-regionally, or nationally). "Data is to be used for informing environmental decision-making at a variety of governmental levels to facilitate an interaction between science and policy."

Company	3Crowns Tech
Platform	Envirocoms
Website	https://www.envirocoms.com/
Accessed	Not applicable.
Data Supported	Not applicable.
Upload Type	Not applicable.
Outputs/Visualizations	Not applicable.
Pricing	Not applicable.
Contact	Not applicable.
Notes	Could not connect with a contact person there after multiple attempts to discuss options.

6.3 Appendix - Presentation and Discussion Summary from Final Webinars

Two webinars were conducted in September 2020 for the purpose of sharing results and soliciting additional comments. The invitation to participate in the close-out webinar was sent to 368 people, including all people invited to take the original survey plus key people in the State governmental agencies. Overall 49 people attended the webinar.

- On September 28, 2020, 32 people participated (not including the presenters).
- One September 30, 2020, 17 people participated (not including the presenters).

Slides from the presentation are included at the end of this Appendix. During the presentation, four poll questions were posed to the attendees, to see who was in attendance and to assess the agreement of a larger group with opinions expressed during the smaller subset of interviews. People were entering and leaving during the meeting, hence the difference in numbers responding to each question. Question 1 was a "choose all that apply", so the percentages will not total to 100%.

Question 1:
Choose all that apply. I am a...

A. End user of data (uses data for reports, newsletters, management decisions, scientific research, etc.). day 1: 14/32 (44%); day 2: 9/17 (53%); TOTAL 23/49 (47%)
B. Data generator (conducting field work or representing an organization that conducts field work). day 1: 12/32 (38%); day 2: 6/17 (35%); TOTAL 18/49 (37%)
C. Different type of participant, neither a data generator nor an end user. day 1: 3/32 (9%); day 2: 2/17 (12%); TOTAL 5/49 (10%)

No Answer. day 1: 10/32 (31%); day 2: 8/17 (0%); TOTAL 18/49 (37%)

Question 2:
"Ability to retrieve group-specific data directly from a group's own website is highly desired." Do you…

A. strongly agree	day 1: 7/32 (22%)	day 2: 3/15 (20%)	TOTAL: 10/47 (21%); 10/27 (37%)
B. agree	day 1: 2/32 (6%)	day 2: 6/15 (40%)	TOTAL: 8/47 (17%); 8/27 (30%)
C. neither agree nor disagree	day 1: 3/32 (9%)	day 2: 3/15 (20%)	TOTAL: 6/47 (13%); 6/27 (22%)
D. disagree	day 1: 3/32 (9%)	day 2: 0/15 (0%)	TOTAL: 3/47 (6%); 3/27 (11%)
E. strongly disagree	day 1: 0/32 (0%)	day 2: 0/15 (0%)	TOTAL: 0/47 (0%); 0/27 (0%)
F. no opinion	day 1: 2/32 (6%)	day 2: 0/15 (0%)	TOTAL: 2/47 (4%)
No Answer	day 1: 15/32 (47%)	day 2: 3/15 (20%)	TOTAL: 18/47 (39%)

Question 3:
Would access to built-in display and analysis tools (like mapping and graphing) at both the local and LIS-regional level incentivize you (both data generators and end-users) to use the database?

A. yes	day 1: 18/31 (58%)	day 2: 9/15 (60%)	TOTAL: 27/46 (59%); 27/30 (90%)
B. maybe	day 1: 2/31 (6%)	day 2: 1/15 (7%)	TOTAL: 3/46 (7%); 3/30 (10%)
C. no	day 1: 0/31 (0%)	day 2: 0/15 (0%)	TOTAL: 0/46 (0%); 0/30 (0%)
D. no opinion	day 1: 0/31 (0%)	day 2: 0/15 (0%)	TOTAL: 0/46 (0%); 0/30 (0%)
No Answer	day 1: 11/31 (35%)	day 2: 5/15 (33%)	TOTAL: 16/46 (35%)

Question 4:
"Have we missed anything? You may respond here or we can contact you later. If you want us to follow up with you, please include your name and how we should contact you."

The short-answer responses below were submitted by attendees:

- *How will this interface with other data organization efforts in the region? For example, I think that GMRI has led a considerable effort in the Gulf of Maine to coordinate data collection efforts.*
- *Thank you so much for providing this update and for including me in the process.*
- *This seems to be targeted at small data generators, like UWS groups. Will it also include larger academic and agency studies?*
- *Nothing to add, but providing hand holding on data management and uploading is most important to my volunteer organization.*
- *Biological Data are extremely important. The exclusive focus on nitrogen/nutrients does not speak to biointegrity and the consolidated pressures that impact biointegrity. Nutrients are just one of thousands of chemical, physical and biological pressures that affect LIS. Biointegrity also reflects the integration of many drivers of change, including climate, land disruption and invasive species.*
    - ***Author note:*** We think there may have been a misunderstanding with this attendee, as we did not propose to focus on nitrogen/nutrients during the webinar.
- *I think a context I am missing with regard to data quality/integrity is the breadth of groups and data that may enter data onto the site. Can you comment on that? Would, for example, school groups (high school or other) be entering data onto the site?*
- *I would like to have a call with you and Tracy to talk more in depth about what we are doing in NYSDEC. Perhaps find a way to work together and minimize any duplication of effort. Thanks.*
- *The only additional thought I have is that lower quality data can be useful in lieu of no data. But you already noted categories that would delineate those. This sounds like a needed project that we have heard on many fronts -- but implementation has hit a lot of walls. Also love the visualization, but wouldn't prioritize this. Having a one stop shop for*

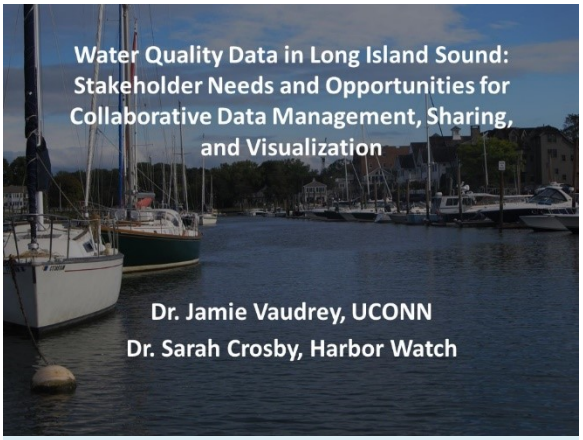
data will have such a high benefit, you will still get users. Visuals could be a phase 2 or 3 of a project -- after we see "who" users are.

- *None, but Thanks!*
- *We (NRWC) only collect freshwater data, so I am with Sarah in advocating for its inclusion.*
- *PANGAEA and other data repositories issue a unique digital identifier (doi) to data archives - for attribution/citation purposes. Not sure how feasible that would be.*
- *We already use water reporter as an organization, it'd be great if this could connect easily especially if this doesn't encompass our entire geographic region.*

Additional Verbal Comments, where respondents did not put the same comment in responses to Question 4.

- *Need to establish a minimum threshold for data entry. Need to be able to tell if data are collected using QA/QC standards.*
- *Important to be interoperable, use a standardized data format – both for importing other data and exporting data.*
- *Be sure to include larger data collectors, e.g. IEC.*
- *Open ocean community has some well-operating databases .*
- *MS4 monitoring, would be good to get their data into the database, contact them regarding whether they would use it [post discussion – could incentivize them to submit data by making one output be a "MS4 report" suitable for submitting to the State agencies: NYSDEC, CTDEEP].*

Power Point Presentation:



Thank you!

- Long Island Sound Stewardship Fund
  - Long Island Sound Funders Collaborative
  - Long Island Community Foundation
- Survey respondents, interviewees, partners at Save the Sound and Chesapeake Commons, and everyone else who has participated

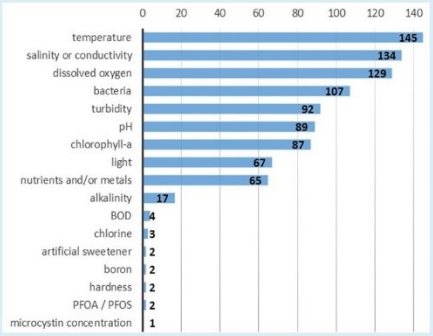
Goals

- …**Overall**: Support the management, sharing, and visualization of ***data*** from around LIS
- …**of our Study**: To assess the needs of people collecting and using monitoring data
- …**for Today**: To share our findings and *GET YOUR FEEDBACK!*

What we did

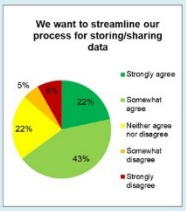
- Survey of monitoring groups
  - Who, what, where, how?
- Survey of groups' data-related needs
- Interviews with selection of stakeholders:
  - Data generators
  - Data end users
- Review of existing tools
- Report (*Coming soon!*)

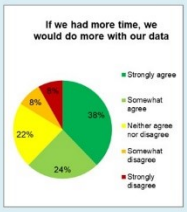
55 groups actively monitoring

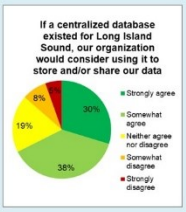


Survey Response Format

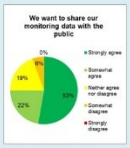
- Statement: "I like puppies!"

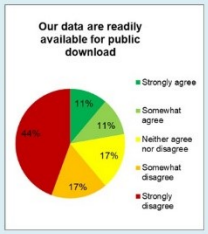


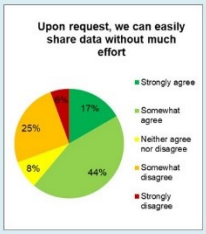




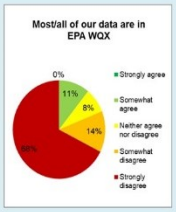
- Data management is a problem for groups
- Opportunity to build something together to solve this problem

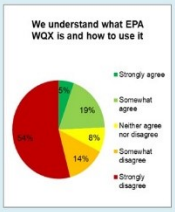


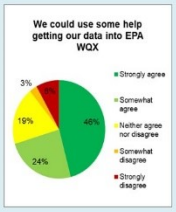


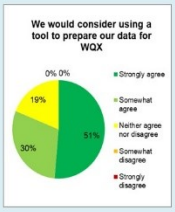


75% of respondents want to share their data with the public, and yet only 22% said their data are readily available for public download.





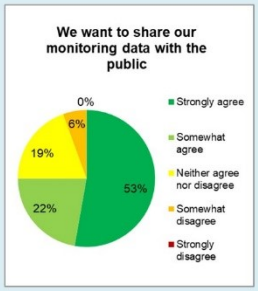


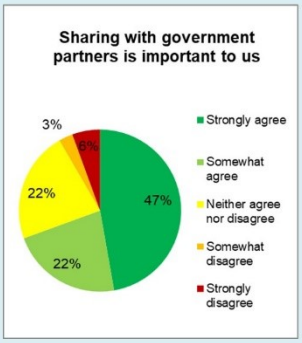


Getting data into EPA WQX is an important way to share with government partners, and is an area where groups could use support

Most "strong disagreement"

Statement	Mean Score (1=strongly agree, 4=strongly disagree)
"Our data are currently readily available for public download, either through our website or a third-party website."	3.13
"We understand what EPA WQX is and how to use it."	3.26
"We want to use some other method [i.e., not excel] for data entry."	3.36
"Most or all of our data are currently uploaded to EPA's WQX system."	3.62





Most "strong agreement"

Statement	Mean Score (1=strongly agree, 4=strongly disagree)
"If a tool existed to help prepare our data for entry into EPA WQX, we would consider using it."	1.37
"We want to share our monitoring data with the public."	1.41
"We want to use Excel for data entry."	1.5
"Collecting data under an EPA-approved QAPP (quality assurance project plan) is important to our organization."	1.52
"Sharing our data with local (towns), State (CT DEEP, NY DEC) and Federal (e.g., EPA, NOAA) governmental groups is important to our organization."	1.57

Before we move on to asking for feedback on our recommendations for next steps, any general questions or comments?

Recommendations: Do you agree?

- INTERFACE & DATA ENTRY
  - Excel is a preferred interface.
  - Clear requirements for metadata and format.
  - Allow and clearly define all water quality parameters (e.g., provide choice of units).
  - Guidelines on best practices for groups on data management and quality control are needed.

Recommendations: Do you agree?

- FEATURES
  - Search functionality.
  - Visualization tools (mapping, graphing).
  - Data must be easily downloadable (by end users AND the public).
  - Point of contact for technical support.
  - Translation for upload to EPA's WQX database.

Recommendations: Do you agree?

- CONCERNS
  - Will our staff have the time and technical skills needed?
  - Want to be able to link to the database and our data from our own website.

End Users: Do you agree?

- INTERFACE & DATA ENTRY
  - Clear requirements for *complete* metadata.
  - Include QA/QC info.
  - High burden on constrains to uploading data – must include information on QA/QC, units, methods; just uploading data without this info is not allowed.
- FEATURES
  - Well-developed search functionality, including searching on QA/QC metrics.
  - Built-in display and analysis tools, from regional summaries to station and sample level.
  - Interoperability with other databases, should follow standard data formats.

End Users: Do you agree?

- SUGGESTIONS
  - Incentivize data entry by providing visualization tools.
  - Provide support to help groups put past/old data into the database.

Concerns and Questions

- DATA OWNERSHIP
  - Include citation for data, attribute groups.
  - Most groups seemed fine with making data widely available for comparison to other areas, but wanted to be able to link directly to the data for their sites from their websites.

Concerns and Questions

- TYPES OF DATA, BEYOND POINT SAMPLING OF WATER QUALITY
  - Will biological data be included?
  - Will continuous data be included?

Concerns and Questions

- GENERAL CONCERNS
  - Adaptive and relevant for future methods and data collected at assorted spatial scales and time intervals.
  - Need to define the potential audience for both the database itself and its outputs – cannot be everything to everyone.

Next Steps: Building Connections

- Community of Practice
  - Darn, Covid! More to come!
  - Google group and listserv
- Next steps update: Save the Sound

Community Data Platform Project

Overall Project Goal
Create a standardized user-friendly data storage, visualization, and retrieval system for Long Island Sound community-generated data.

Current Initiative
Develop a scope of work and strategy for building and maintaining the data platform.

Project Partners
Save the Sound
Chesapeake Commons
Harbor Watch
Jamie Vaudrey

Community Data Platform Project

Major Activities

1) Needs Assessment using CoP surveys, discussions, focus groups.

2) Draft Features Set for data storage, visualization and retrieval based on above research activities and research of similar systems in use.

3) Revise, refine and finalize Feature Set and Project Scope based on feedback from stakeholders, user testing, and assessment of cost and maintenance requirements.

Thank you!

- ...for your feedback!
- Long Island Sound Stewardship Fund
- Save the Sound
- Chesapeake Commons

6.4 Appendix - Data Generator Survey #1 & #2: Questions and Response Format

On the following pages, the blank templates of each of the two surveys of data generators can be found. Each survey was sent to potential respondents as a web link by email. Participation was optional and participants were not compensated in any way for participation.

Introduction

Here in New York and Connecticut, dozens of groups work toward Long Island Sound's protection. We have a shared goal of a healthy and resilient estuary, and we generate of an enormous amount of data about the Sound, its embayments, and its watershed. Dr. Jamie Vaudrey (UCONN) and Dr. Sarah Crosby (Harbor Watch) are leading an effort to (1) build capacity across all of our organizations by providing the necessary information for the development of a shared, centralized database that groups can choose to use to share their data, and (2) provide a mechanism for communication among all of us as Long Island Sound data generators to promote collaboration.

This survey will be the first step in that process, and we sincerely appreciate your participation! Your answers here will be used to inform what is needed in a database that would benefit the maximum number of groups. **Even if you are not actively monitoring a water body, we would like you to proceed - this survey is also being used to gather names for people wishing to provide some feedback as an end-user of data (6 quick questions, if this is you!)**

Thank you! This work would not be possible without the generous support of the Long Island Sound Stewardship Fund (from the Long Island Sound Funders Collaborative and the Long Island Community Foundation).

Contact Information

Name of Organization

Name of water body (or multiple water bodies) where you sample. If you don't sample, please list your water body of interest.

Name of person (or people) completing this survey.

Are you or a group from your organization willing to participate in a follow-up survey or in-person interview regarding your needs in a database?

- ○ Yes.
- ○ No.

May we contact you if we have any questions about your responses on this survey?

- ○ Yes.
- ○ No.

Please provide contact information.

Name of contact person (or people).	
Email address.	
Phone number.	

Block 21

Is your monitoring program conducted under the Unified Water Study (UWS), overseen by Save the Sound?

- ○ We do not participate in the Unified Water Study.
- ○ We participate in the Unified Water study, but we also conduct additional monitoring not covered by the UWS.
- ○ All of our monitoring activities are conducted as part of the Unified Water Study.
- ○ We do not currently conduct monitoring of water quality or habitat quality in aquatic environments.

We have access to all of the information on the monitoring activities conducted by the UWS, so you do not need to provide any information on what is monitored or how monitoring is conducted when you are participating in the UWS. Please answer the following questions for your other monitoring activities.

We have access to all of the information on the monitoring activities conducted by the UWS, so you do not need to provide any information on what is monitored or how monitoring is conducted.

Do you have any additional comments or anything else you would like us to know?

Do you have any additional comments or anything else you would like us to know?

Parameters Monitored - Overview

Which parameters do you monitor?

- ☐ temperature
- ☐ salinity (salt water) or conductivity (fresh water)
- ☐ dissolved oxygen
- ☐ chlorophyll concentration or fluorescence
- ☐ turbidity (total suspended solids, NTU)
- ☐ water clarity (Secchi depth or light measurements)
- ☐ pH
- ☐ alkalinity
- ☐ nutrients and/or metals
- ☐ bacteria
- ☐ other

DD - Temperature

Digging Deeper - Temperature; How do you collect temperature data?

- ☐ digital, handheld sonde (YSI, Eureka, HydroLab, etc.)
- ☐ deployed instrument (something you leave in the water)
- ☐ thermometer
- ☐ other

You answered "other" for the previous question. Please explain.

DD - Salinity & Conductivity

Digging Deeper - Salinity & Conductivity; How do you collect salinity or conductivity data?

- ☐ digital, handheld sonde (YSI, Eureka, HydroLab, etc.)
- ☐ deployed instrument (something you leave in the water)
- ☐ refractometer
- ☐ send to an analytical lab
- ☐ other

You answered "other" for the previous question. Please explain.

DD - Dissolved Oxygen

Digging Deeper - Dissolved Oxygen; How do you collect oxygen data?

- ☐ digital, handheld sonde (YSI, Eureka, HydroLab, etc.)
- ☐ deployed instrument (something you leave in the water)
- ☐ titration kit (e.g. LaMotte, Hach, etc.)
- ☐ send to an analytical lab
- ☐ other

You answered "other" for the previous question. Please explain.

DD - Chlorophyll

Digging Deeper - Chlorophyll; How do you collect chlorophyll data?

- ☐ digital, handheld sonde for in situ fluorescence (YSI, Eureka, HydroLab, etc.)

- [ ] deployed instrument for in situ fluorescence (something you leave in the water)
- [ ] filter and send to an analytical lab, extraction and fluorescence method
- [ ] filter and send to an analytical lab, HPLC method
- [ ] other

You answered "other" for the previous question. Please explain.

DD - Turbidity

Digging Deeper - Turbidity; How do you collect turbidity data?

- [ ] digital, handheld sonde (YSI, Eureka, HydroLab, etc.)
- [ ] deployed instrument (something you leave in the water)
- [ ] our organization has a bench-top turbidity sensor, for NTU
- [ ] send to an analytical lab, for NTU
- [ ] send to an analytical lab, for total suspended solids via gravimetric technique (weigh filters)
- [ ] our organization analyzes in-house for total suspended solids via gravimetric technique (weigh filters)
- [ ] other

You answered "other" for the previous question. Please explain.

DD - Water Clarity

Digging Deeper - Water Clarity; How do assess water clarity?

- [ ] digital, handheld sonde with light sensor (YSI, Eureka, HydroLab, etc.)
- [ ] deployed instrument (something you leave in the water)
- [ ] light meter (LiCor, Biospherical, etc.)
- [ ] Secchi disk
- [ ] other

You answered "other" for the previous question. Please explain.

DD - pH

Digging Deeper - pH; How do you collect pH data?

- ☐ digital, handheld sonde (YSI, Eureka, HydroLab, etc.)
- ☐ deployed instrument (something you leave in the water)
- ☐ bench top instrument
- ☐ send to an analytical lab
- ☐ other

You answered "other" for the previous question. Please explain.

DD - Alkalinity

Digging Deeper - Alkalinity; How do you collect alkalinity data?

- ☐ digital, handheld sonde (YSI, Eureka, HydroLab, etc.)
- ☐ deployed instrument (something you leave in the water)
- ☐ bench top instrument
- ☐ send to an analytical lab
- ☐ other

You answered "other" for the previous question. Please explain.

DD - Nutrients & Metals

Digging Deeper - Nutrients and/or Metals; How do you analyze for nutrients or metals?

- ☐ send samples to an analytical lab
- ☐ use a field kit (e.g. LaMotte, Hach, etc.)
- ☐ use a probe
- ☐ other

You answered "other" for the previous question. Please explain.

Which nutrients and/or metals? Later questions will get into the forms of each nutrient.

- ☐ carbon
- ☐ nitrogen
- ☐ phosphorus
- ☐ sulfur
- ☐ mercury
- ☐ iron
- ☐ lead
- ☐ copper
- ☐ other (contaminants, human drugs, hardness, etc.)

Carbon Species

What species of carbon? **Check only those that are analyzed, not those calculated by difference**. Note that you have a choice of filtered or unfiltered samples; choose all that apply. Also note, you have a choice of analysis conducted by an analytical lab or a pre-packaged field kit (e.g. LaMotte or Hach kits). Please choose all that apply.

	Filtered Sample, Analytical Lab	Unfiltered Sample, Analytical Lab	Filtered Sample, Field Kit	Unfiltered Sample, Field Kit
carbon dioxide	☐	☐	☐	☐
inorganic carbon	☐	☐	☐	☐
organic carbon	☐	☐	☐	☐
total carbon	☐	☐	☐	☐
CDOM (colored dissolved organic matter)	☐	☐	☐	☐
carbon isotopes (del13C) in water	☐	☐	☐	☐
carbon isotopes (del13C) of particulates, via analysis of a filter	☐	☐	☐	☐
other	☐	☐	☐	☐

You answered "other" for the previous question. Please explain.

Nitrogen Species

What species of nitrogen? **Check only those that are analyzed, not those calculated by difference**. Note that you have a choice of filtered or unfiltered samples; choose all that apply. Also note, you have a choice of analysis conducted by an analytical lab or a pre-packaged chemical field kit (e.g. LaMotte or Hach kits). Please choose all that apply.

	Filtered Sample, Analytical Lab	Unfiltered Sample, Analytical Lab	Filtered Sample, Field Kit	Unfiltered Sample, Field Kit
ammonium (NH4) or ammonia (NH3)	☐	☐	☐	☐
nitrate (NO3)	☐	☐	☐	☐
nitrite (NO2)	☐	☐	☐	☐
nitrate + nitrite (NO3 + NO2, or NOx)	☐	☐	☐	☐
total nitrogen	☐	☐	☐	☐
organic nitrogen	☐	☐	☐	☐
Kjeldahl nitrogen (NH3 + organic N)	☐	☐	☐	☐
particulate nitrogen (PN), via analysis of a filter	☐	☐	☐	☐
nitrogen isotopes (del15N) in water	☐	☐	☐	☐
nitrogen isotopes (del15N) of particulates, via analysis of a filter	☐	☐	☐	☐
other	☐	☐	☐	☐

You answered "other" for the previous question. Please explain.

Phosphorus Species

What species of phosphorus? **Check only those that are analyzed, not those calculated by difference**. Note that you have a choice of filtered or unfiltered samples; choose all that apply. Also note, you have a choice of analysis conducted by an analytical lab or a pre-packaged field kit (e.g. LaMotte or Hach kits). Please choose all that apply.

	Filtered Sample, Analytical Lab	Unfiltered Sample, Analytical Lab	Filtered Sample, Field Kit	Unfiltered Sample, Field Kit
ortho-phosphate / inorganic phosphate / soluble reactive phosphorus	☐	☐	☐	☐

	Filtered Sample, Analytical Lab	Unfiltered Sample, Analytical Lab	Filtered Sample, Field Kit	Unfiltered Sample, Field Kit
organic phosphorus	☐	☐	☐	☐
total phosphorus	☐	☐	☐	☐
phosphorus isotopes (del31P) in water	☐	☐	☐	☐
phosphorus isotopes (del31P) of particulates, via analysis of a filter	☐	☐	☐	☐
other	☐	☐	☐	☐

You answered "other" for the previous question. Please explain.

Sulfur Species

What species of sulfur? **Check only those that are analyzed, not those calculated by difference**. Note that you have a choice of filtered or unfiltered samples; choose all that apply (sediment samples are "unfiltered"). Also note, you have a choice of analysis conducted by an analytical lab or a field probe (voltammetric probe). Please choose all that apply.

	Filtered Sample, Analytical Lab	Unfiltered Sample, Analytical Lab	Filtered Sample, Field Probe	Unfiltered Sample, Field Probe
total sulfur	☐	☐	☐	☐
sulfate (SO4)	☐	☐	☐	☐
hydrogen sulfide (HS)	☐	☐	☐	☐
dimethyl sulfide (DMS)	☐	☐	☐	☐
methanethiol (MSH)	☐	☐	☐	☐
3-methiolpropionate (MMPA)	☐	☐	☐	☐
3-mercaptopropionate (MPA)	☐	☐	☐	☐
carbonyl sulfide (COS)	☐	☐	☐	☐
carbon disulfide (CS2)	☐	☐	☐	☐
dimethyl disulfide (DMDS), H2S, S0	☐	☐	☐	☐
thiosulfate in sediments	☐	☐	☐	☐
elemental sulfur in sediments	☐	☐	☐	☐
pyrite in sediments	☐	☐	☐	☐
sulfur isotopes (del32S, del33S, del34S, del36S) in water	☐	☐	☐	☐
sulfur isotopes (del32S, del33S, del34S, del36S) of particulates	☐	☐	☐	☐
other	☐	☐	☐	☐

You answered "other" for the previous question. Please explain.

Mercury Species

What species of mercury? **Check only those that are analyzed, not those calculated by difference**. Note that you have a choice of filtered or unfiltered samples; choose all that apply.

	Filtered Sample, Analytical Lab	Unfiltered Sample, Analytical Lab
methylmercury	☐	☐
total mercury	☐	☐
other	☐	☐

You answered "other" for the previous question. Please explain.

Bacteria Species

Digging Deeper - Bacteria; How do you collect bacteria data?

- ☐ all bacteria, acridine orange stain with fluorescence microscope counts
- ☐ all bacteria, DAPI stain with fluorescence microscope counts
- ☐ Enterococci, via IDEXX Enteroalert system
- ☐ Enterococci, via membrane filtration
- ☐ Enterococci, via qPCR
- ☐ *E. coli*, via IDEXX Colilert system
- ☐ *E. coli*, via membrane filtration
- ☐ *E. coli*, via qPCR
- ☐ other

You answered "other" for the previous question. Please explain.

Other Water Column Parameters

Digging Deeper - Other; What other parameters do you monitor in the water column?

- ☐ 1,7-DMX
- ☐ Acetaminophen
- ☐ Atenolol
- ☐ Azithromycin
- ☐ Caffeine
- ☐ Carbamazepine
- ☐ CDOM (colored dissolved organic matter)
- ☐ Chloride
- ☐ Chlorine
- ☐ Cotinine
- ☐ Ethylene glycol
- ☐ Hardness
- ☐ Optical Brighteners
- ☐ PCB's
- ☐ Primidone
- ☐ Propylene
- ☐ Surfactants
- ☐ TPH Petroleum ID (alpha)
- ☐ Urobilin
- ☐ other

You answered "other" for the previous question. Please explain.

Benthic Monitoring

Please indicate the type of organism or benthic monitoring that you conduct. Leave blank if none apply.

> Please note, a "named, standardized protocol" refers to a set of standard operating procedures that any group could choose to use; something which is openly available to anyone searching for a method. The procedures specify things like sampling frequency, number of replicates, specific terminology, etc., and are usually distributed by a single organization (e.g. CT DEEP, Save the Sound's Unified Water Study, RI Watershed Watch, etc.).

> A protocol developed by your organization or team refers to a method or standard operating procedure that only your group is currently using, though other groups may follow a similar protocol. For example, you may collect and analyze

sediment grain size using a procedure developed by your organization.

	salt water	freshwater	we use a named, standardized protocol	we use a protocol developed by our organization or research team
sediment (grain size, organic content, other analyses)	☐	☐	☐	☐
macrophytes (plants or seaweed)	☐	☐	☐	☐
benthic organisms (organisms living in or on the sediment such as worms, shellfish, crabs, etc.)	☐	☐	☐	☐
aquatic insects	☐	☐	☐	☐
fish	☐	☐	☐	☐
birds	☐	☐	☐	☐
other	☐	☐	☐	☐

What is the named protocol(s) for sediment? Please include name of organization that provides the standard operating procedure.

What is the named protocol(s) for macrophytes? Please include name of organization that provides the standard operating procedure.

What is the named protocol(s) for benthic organims? Please include name of organization that provides the standard operating procedure.

What is the named protocol(s) for aquatic insects? Please include name of organization that provides the standard operating procedure.

What is the named protocol(s) for fish? Please include name of organization that provides the standard operating procedure.

You selected "other" in the previous question. Please explain.

What is the named protocol(s) for the "other" type of sampling you selected? Please include name of organization that provides the standard operating procedure.

Additional Comments

Do you have any additional comments or anything else you would like us to know?

Additional Comments - UWS participants

Introduction

Thank you for your response to our initial survey, asking about the monitoring your organization conducts. This follow-up survey is designed to get a broader perspective on how data generators, like you, feel on a variety of topics relevant to developing a common database for Long Island Sound and developing visualization tools useful for data analysis and community outreach.

Background on our project: Here in New York and Connecticut, dozens of groups work toward Long Island Sound's protection. We have a shared goal of a healthy and resilient estuary, and we generate an enormous amount of data about the Sound, its embayments, and its watershed. Dr. Jamie Vaudrey (UCONN) and Dr. Sarah Crosby (Harbor Watch) are leading an effort to (1) build capacity across all of our organizations by providing the necessary information for the development of a shared, centralized database that groups can choose to use to share their data, and (2) provide a mechanism for communication among all of us as Long Island Sound data generators to promote collaboration. This work is possible due to the generous support of the Long Island Sound Stewardship Fund (from the Long Island Sound Funder's Collaborative and the Long Island Community Foundation).

Contact Information

Names of the person or people completing this survey.

Name of your organization(s).

Email address of the person or people completing this survey.

Entering Data

Entering Data - Please indicate the extent to which you agree or disagree with the statements.

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
1. Our organization (staff and/or volunteers) is comfortable entering our own data without external help from consultants or others.	○	○	○	○	○
2. We want to use Excel for data entry.	○	○	○	○	○
3. We want to use some other method for data entry (please describe other method in comments box).	○	○	○	○	○
4. Using a mobile application (phones or tablets) for data entry in the field or lab would be a good addition to our data entry options.	○	○	○	○	○

Any comments on data entry?

Managing Data

Managing Data - Please indicate the extent to which you agree or disagree with the statements.

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
5. Managing our monitoring data is a burden on our time (staff and/or volunteers).	○	○	○	○	○
6. Our data are currently stored digitally in an organized format that could be readily understood by someone outside of our organization.	○	○	○	○	○
7. Our data are all backed up and recoverable if something were to happen to our primary data files.	○	○	○	○	○
8. Our organization (staff and/or volunteers) is comfortable ***digitally storing*** our own data without external help from consultants or others.	○	○	○	○	○
9. Our organization (staff and/or volunteers) is comfortable ***analyzing*** our own data without external help from consultants or others.	○	○	○	○	○

Any comments on data management?

Quality Control & EPA Requirements

Quality Control & EPA Requirements - Please indicate the extent to which you agree or disagree with the statements.

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
10. We have a quality control process for our data entry.	○	○	○	○	○
11. Collecting data under an EPA-approved QAPP (quality assurance project plan) is important to our organization.	○	○	○	○	○
12. Most or all of our data are currently uploaded to EPA's WQX system (water quality exchange).	○	○	○	○	○
13. We understand what EPA WQX is and how to use it.	○	○	○	○	○
14. We could use some help getting our data into EPA WQX.	○	○	○	○	○
15. If a tool existed to help prepare our data for entry into EPA WQX, we would consider using it.	○	○	○	○	○

Any comments on quality control or EPA's WQX?

Sharing Data

Sharing Data - Please indicate the extent to which you agree or disagree with the statements.

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
16. Our data are currently readily available for public download, either through our website or a third-party website.	○	○	○	○	○
17. If a member of the public emailed me asking for data on a certain place that we monitor in a certain year, I could easily send them back that data without much effort.	○	○	○	○	○
18. Sharing our data with local (towns), state (CT DEEP, NY DEC) and federal (e.g., EPA, NOAA) governmental groups is important to our organization.	○	○	○	○	○
19. We want to share our monitoring data with the public.	○	○	○	○	○

Any comments on sharing data?

Visualizing Data

Visualizing Data - Please indicate the extent to which you agree or disagree with the statements.

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
20. We are currently using tools to visualize our data (e.g., ArcGIS, etc.).	○	○	○	○	○
21. Our organization (staff and/or volunteers) is comfortable ***mapping*** our own data without external help from consultants or others.	○	○	○	○	○
22. Our organization (staff and/or volunteers) is comfortable ***graphing*** our own data without external help from consultants or others.	○	○	○	○	○
23. We could use some help visualizing our data (e.g., making maps, graphs).	○	○	○	○	○
24. We would like to see our data on a map with other data from around the Sound.	○	○	○	○	○

Any comments on visualizing data?

Follow Up

Based on results from in-person interviews, we want to better understand how familiar people are with two specific programs/products. Please indicate the extent to which you agree or disagree with the statements.

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
25. I am familiar with the Sound Health Explorer website.	○	○	○	○	○
26. I am familiar with the Unified Water Study.	○	○	○	○	○

Overview

Overview - Please indicate the extent to which you agree or disagree with the statements.

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree

Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
27. We are currently happy with our process for entering, storing and sharing our data.				
○	○	○	○	○
28. We would like to streamline our process for storing/sharing our data.				
○	○	○	○	○
29. If we had more time, we would do more with our data.				
○	○	○	○	○
30. If a centralized database existed for Long Island Sound, our organization would consider using it to store and/or share our data.				
○	○	○	○	○

Thank you for your time! Any final comments on the overview or any other sections?