Creating a Water Quality Geodatabase for the West Hawaiʻi Island Region

by

Donald Borer

A Thesis Presented to the
Faculty of the USC Graduate School
University of Southern California
In Partial Fulfillment of the
Requirements for the Degree
Master of Science
(Geographic Information Science and Technology)

December 2018

Table of Contents

List of Figures

List of Tables

Acknowledgements

I thank all the water quality scientists whose data I used in this thesis. These dedicated scientists believed the results of their work would be important to show the status of the ocean waters. I treated the data with great care and often envisioned them bouncing around on a boat or swimming out into the waves with their bottles to fill the great machines back in the laboratory. I am honored to pass along their work.

I am grateful to my advisor Dr. Karen Kemp for our weekly sessions discussing how to proceed and getting me back on track when I drifted off-topic. My committee members, Dr. An-Min Wu and Dr. Laura Loyola, played key roles in the initial design and instructing me how to describe a geodatabase project. I am grateful to my writing coach Mallory Graves for teaching me to ask “so what?” as she taught me how to turn my programmer reference manual writing into a readable style.

My water quality mentors from Kona, Dr. Richard Bennett and Dr. Dennis Mihalka, are the reason this document exists. I thank Dr. Bennett for his unwavering patience and dedication to teach me West Hawai‘i ocean water topics. I learned so much and had a wonderful and fun time with lively discussions in weekly calls. I thank Dr. Mihalka for teaching me about his water quality projects and encouraging calls and emails.

I thank Keith Olson for teaching me about water science at NELHA and reviewing my initial designs. I thank Dr. Tracy Wiegner for reviewing an early draft and helping me understand the scope of some projects. I thank Roxane Barraza at Esri for allowing me to have my own academic license for this project.

List of Abbreviations

CWB	State of Hawai‘i Department of Health, Clean Water Branch
EIS/R	Environmental Impact Statement/Report
EPA	US Environmental Protection Agency
GIS	Geographic Information System
GDB	Geodatabase (Esri terminology for native database)
HAR	Hawai‘i Administrative Rules
HDOH	State of Hawai‘i Department of Health
HOT	Hawai‘i Ocean Time series
NELHA	Natural Energy Laboratory of Hawai‘i Authority
PacIOOS	Pacific Islands Oceans Observing System
QAPP	Quality Assurance Project Plan
SGD	Submarine Ground Discharge
SQL	Structured Query Language

Abstract

An integrated ocean water quality geodatabase for the West Hawai‘i Island region is of interest to local scientists who want to assess near-shore ocean waters because the unique properties of this environment allow the data to speak for the environment as a whole. With guidance from local environmental scientists, disparate near-shore water quality projects from different organizations were combined and integrated into a single geodatabase using reproducible methods. This study used SQL query methods to extract data from regulatory monitoring and scholarly research documents into a professional spatio-temporal water science database for use in the Esri ArcGIS Pro software program. The final geodatabase contains 100,000 analyte results from 15,000 samples at 300 stations. The geodatabase, data, methods, images and data and documentation sources used to produce these results were published at GitHub free for use. This geodatabase provides a GIS foundation to support the future development of online web maps and story maps that can be used to inform the public about ocean water quality changes.

Chapter 1 Introduction

Ocean water quality monitoring is essential in efforts to characterize, evaluate, and ensure protective standards and best practices for long-term management of precious water resources. On the west side of Hawaiʻi Island, water scientists have expressed the need to develop a comprehensive database which incorporates the work of various scientists over multiple years, and at different locales along the near-shore ocean waters. This thesis describes the development of an integrated spatio-temporal database of water quality data that can be used for data-driven solutions in this region.

The scientists involved in this study identified a critical disconnect in how water quality data are utilized and understood in academic, regulatory, and public contexts. This disconnect is due largely to a lack of standardized methods of water quality data distribution and aggregation as well as the need for new sharing mechanisms on behalf of the entities (scientific, academic, and regulatory) involved in water quality management in the West Hawaiʻi Island region. Additionally, the lack of efficient ways to integrate disparate data along with overall challenges in the various methodologies used for water quality monitoring have prohibited scientists from creating an authoritative knowledge baseline for influencing water quality management standards and regulations. A robust, spatio-temporal database which factors in years of water quality data would enable people to understand and visualize ocean water quality trends and issues, as well as allow scientists to perform meaningful analyses and studies that represent a much larger source of sample data.

This research offers a two-fold solution to support the needs expressed by the water scientists and other entities involved in this study: 1) provide an ocean water quality geodatabase developed in a GIS as a baseline for quantitative studies and analyses, and 2) provide a foundation for the development of an interactive mapping narrative tool (a web GIS) for water quality education, drawing from a wealth of resources which help to tell the complex story of ocean water quality trends via maps, photos, and other multimedia.

1.1. Motivators of this Project

There is a community of concerned residents and scientists on the Island of Hawaiʻi who would like to tell the story about changes in the water quality along the shoreline of West Hawaiʻi. I worked remotely with local environmental scientists Dr. Richard H. Bennett, Dr. Dennis J. Mihalka, and Keith Olson to understand water quality monitoring, choose projects, and design the geodatabase.

Dr. Bennett, a Hawaiʻi Island environmental scientist living in Kailua-Kona, believes the integration and sharing of this diverse collection of data will contribute to public awareness and provide baseline data for future analysis. Dr. Mihalka, in this document identified as the *kayak* scientist, shared his data from a forty-week long sampling series in Keauhou Bay. Since he wants to inform the public and plan future research projects, he needs a way to publish his results in a publicly accessible database with a known design to share with other scientists. Mr. Olson, the Chief Science Officer of Natural Energy Laboratories Hawaiʻi Authority, and I discussed water quality at Keāhole Point and the region in general. These scientists shared their ideas and insight on how to move forward towards the creation of a water quality management geographic information system.

1.2. Project Overview

The local environmental scientists who motivated this project had the following requirements: 1) combine water quality files from different sources, locations and time ranges into a geodatabase; 2) create methods for the implementation and maintenance of a map-based water quality monitoring data system; and 3) provide a technical foundation from which scientists will be able to create new kinds of water quality visualizations for their ongoing efforts to engage the public.

Scientists collect near-shore ocean water samples on which they perform numerous chemical and microbiological measurements in order to understand the environment. The values of these analyte measurements are used to identify nutrients introduced by land discharge from human and animal waste, agriculture operations, golf course fertilizers, and aquaculture waste disposal. In the West Hawaiʻi region, the characteristics of these analyte values are important to the public because they are used by the State of Hawaiʻi to address the long-term effects of shoreline property development as well as immediate needs like publishing beach pollution advisories.

The State of Hawaiʻi Department of Health (HDOH) provides a set of best management practices to prevent or reduce pollution in the near-shore water for this region (see Chapter 2). Scientists need to follow these practices, methods and measures to make statements regarding exceedances in their results and conclusions. In consultation with water quality scientists, this project identified the required analyte values collected across multiple surveys and research projects and combined them into a single, comprehensive geodatabase. These efforts will provide the basis for a visualization tool to support the scientists need for online tools, and will help them follow and examine the effectiveness of the management practices.

1.3. The Study Area

The general study area is the west side of Hawaiʻi Island, the largest and easternmost of the main chain of islands (see Figure 1). Hawaiʻi Island is also called The Big Island. The well-known city on the west side is Kailua-Kona and on the east side, Hilo.

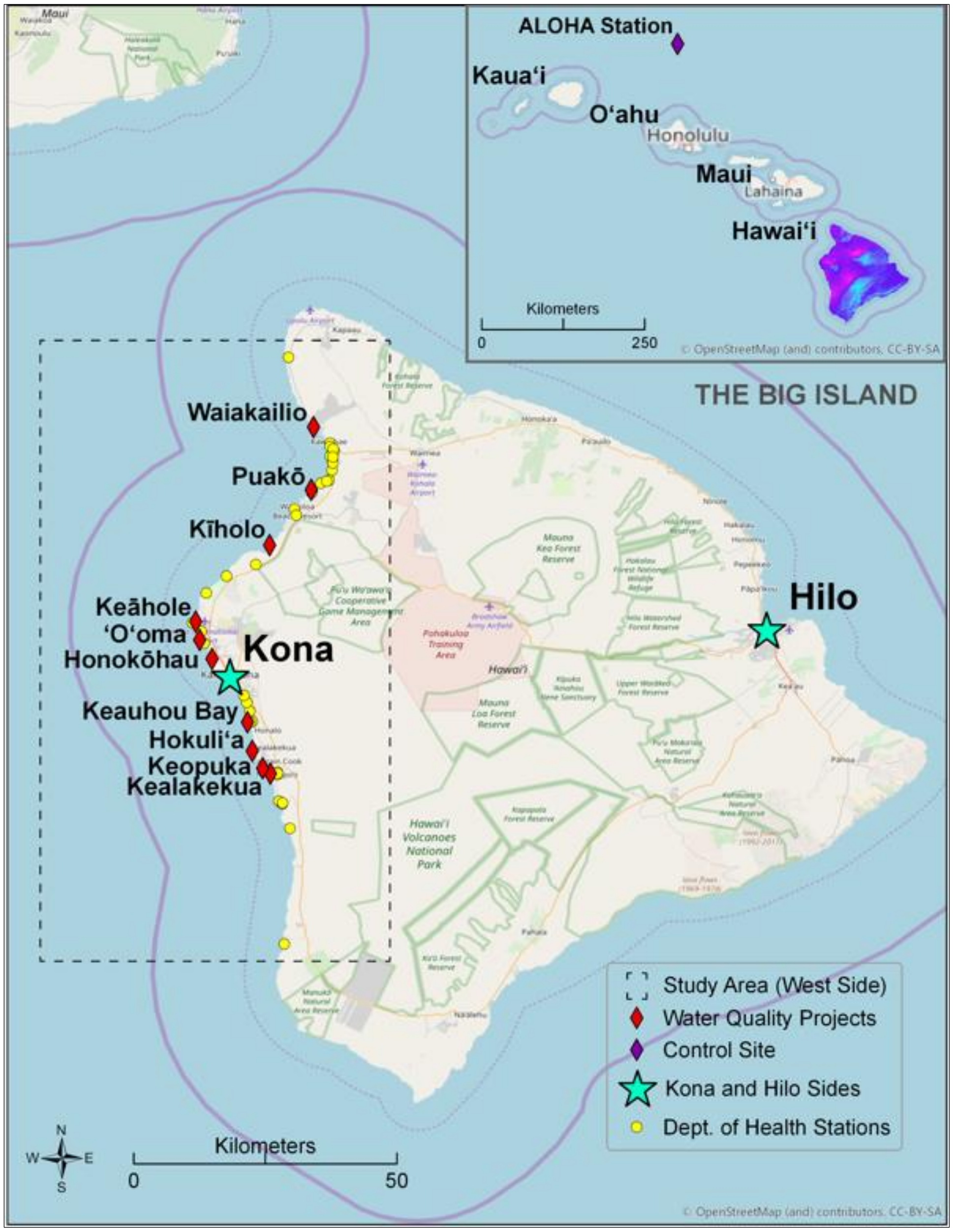


Figure 1. The study area is the west side of Hawai'i Island. Points indicate water quality project sites included in this study.

To understand the mechanisms affecting ocean water quality in West Hawai'i requires understanding the geography, primarily the location of the volcanos and the prevailing weather pattern. Figure 2 shows the five volcanoes making up the island. The two largest volcanoes, Mauna Kea and Mauna Loa reach to high elevations of 13,796ft (4,205m) and 13,679ft (4,169m), respectively (Lau and Mink 2006).

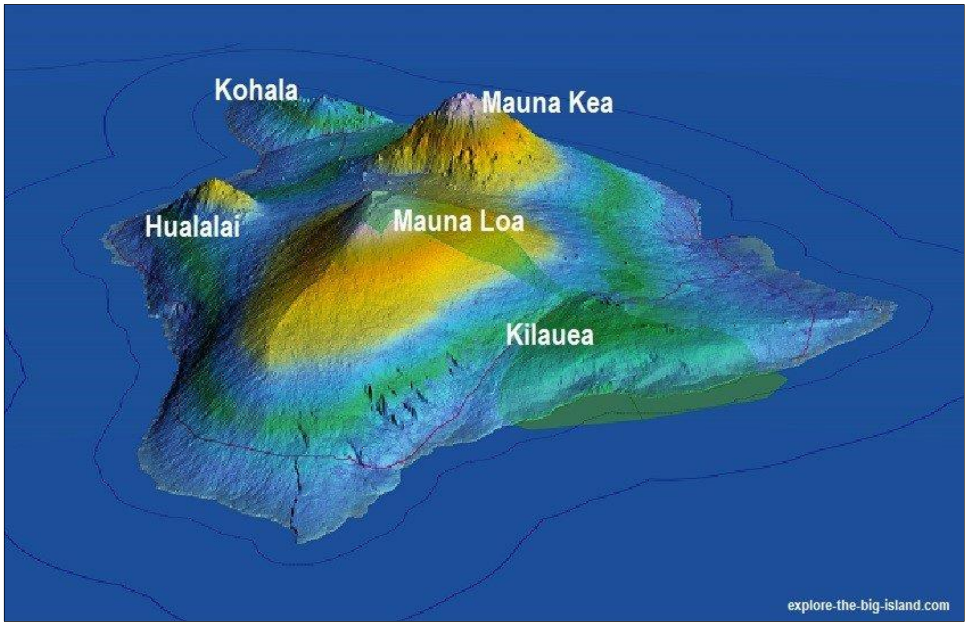


Figure 2. Hawai‘i Island Volcanoes. Source: https://explore-the-big-island.com, accessed August 7, 2018.

In the Hawaiian Islands, the trade winds from the northeast are the main source of rain (Lau and Mink 2006). It rains often on the windward Hilo side, with an annual precipitation around 300 in, or 7.6 m, because this is where the trade winds blow against the tall volcanos, draining the clouds. In contrast, it does not rain often on the leeward, arid west side, with annual precipitation around 8 in, or 20.3 cm. Figure 3 shows the annual rainfall (in inches) on the island.

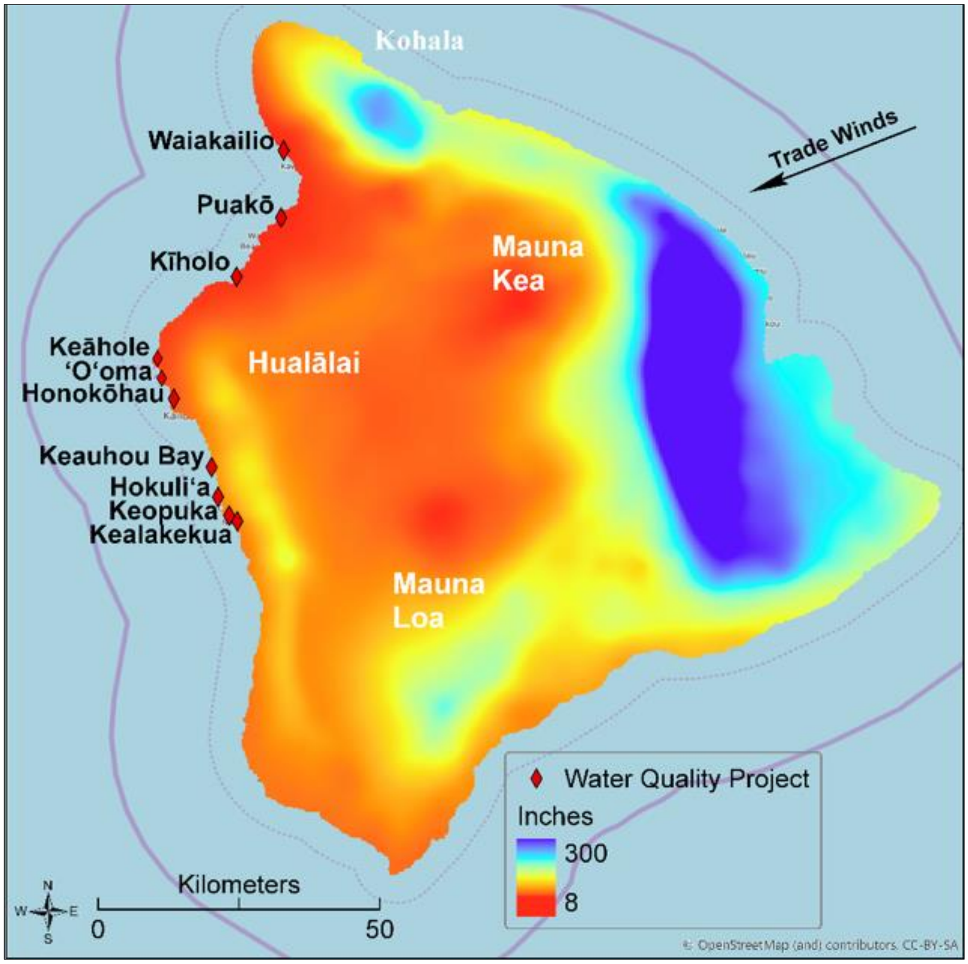


Figure 3. Annual Rainfall on Hawaiʻi Island. Data source: Giambelluca et al. 2013.

The hundreds of inches of rain per year on the windward side permeate through the discontinuous and inhomogeneous layers of volcanic rocks to eventually discharge into the ocean on both sides of the island as submarine ground discharge (SGD) (Laws 1993). SGD occurs from fractures, faults, lava tubes or diffuse seepage at the shoreline (Knee et al. 2010). Knee et al. estimate these flows are as high as five million gallons per mile every day at some locations. Since the discharge quantities are so large, a water quality scientist's overall objective is to measure the additional nutrients introduced by anthropogenic sources (Dollar and Atkinson 1992). As the water flows underground from the volcanoes to the ocean, nutrients from sources such as silicates from the basalt rocks, nitrogen and phosphorus from natural and anthropogenic sources, pathogens from soils or human/animal waste, and other nutrients are dissolved or suspended in the water.

1.4. Water Quality Geodatabase Development Overview

To begin the design and development of the proposed geodatabase, Dr. Bennett and I discussed the available data sources and chose ten projects as the initial baseline data for a GIS water quality monitoring system (Figure 4). The contents of the geodatabase include: four data tables from private consultant scientists, two long time series (decades long sampling regime) provided by State of Hawaiʻi scientists, three data tables from professors and students at the University of Hawaiʻi and a nearly yearlong series from a citizen scientist (Dr. Mihalka) who collected samples while paddling a kayak.

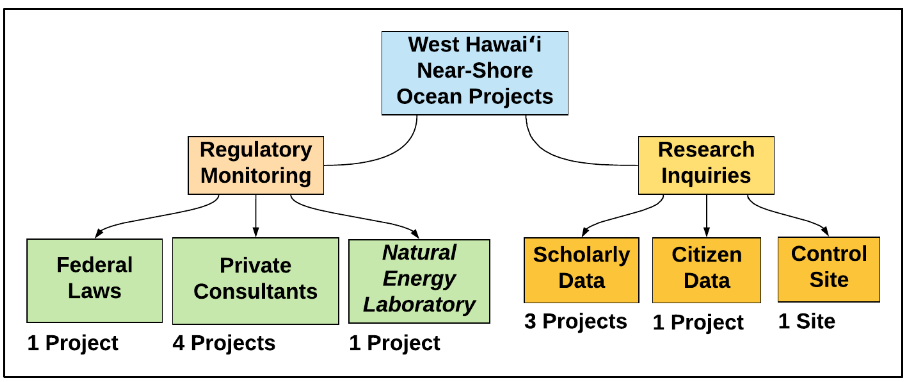


Figure 4. Introduction to West Hawaiʻi water quality data survey projects.

Esri ArcGIS software programs rely on a set of objects collectively called a geodatabase as the common data structure necessary to access spatial and temporal geoprocessing tools. The geodatabase architecture is the native ArcGIS structure for managing data input, output from geoprocessing methods, and distributed data access, and for sharing data and results with the public (Zeiler and Murphy 2010).

Integrating numerous research projects into a water quality geodatabase deliverable was achieved in five major steps, shown in Figure 5. First, data tables from numerous documents were extracted from various document types. Analytes such as nitrogen, phosphorus, oxygen, chlorophyll a, turbidity, temperature, salinity, *Enterococci* spp. and *C. Perfringens* were uniquely identified for inclusion in the database. Then, sampling station locations identified in each data source were geocoded using MyMaps GIS by Google. Next, a professional water quality database schema was modified into ArcGIS geodatabase objects and Microsoft-SQL query scripts were developed as the methods to process different arrangements of input data before moving the data into the geodatabase objects in the fourth step. The final step produced the key deliverable of this project: the integration of matched analytes, collected over many years of disparate projects, into a single geodatabase shared openly on GitHub.

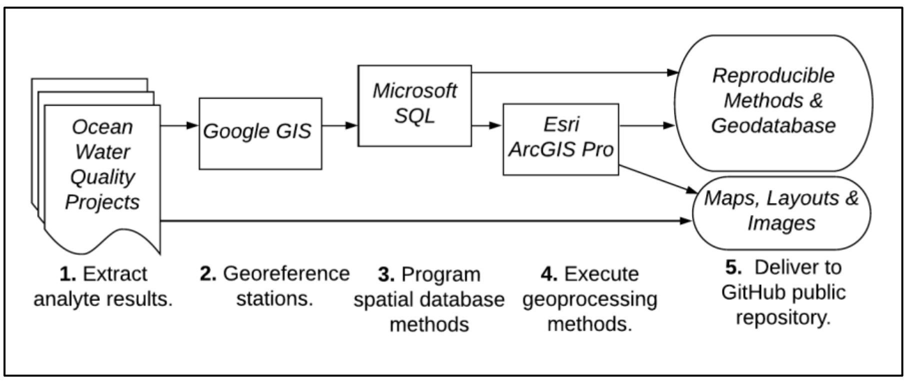


Figure 5. Development workflow for this study.

1.5. Thesis Organization

The general narrative structure of this document reflects the way in which the project was conducted. More specifically, the nature of this study required an understanding of water quality in West Hawaiʻi gleaned from a number of perspectives; chief among these was the group of environmental scientists who guided the objectives and requirements of this project. Each of the scientists has unique methods of data collection, data standards and goals, and various platforms for analyzing water quality data. Also, it is imperative to understand and describe the environmental and geophysical context of the study region, site-specific information is critical to the validity of the results and conclusions.

The structure of this paper integrates and reinforces how each scientist contributed to the development of the geodatabase. It includes a large number of exhibits that provide additional narrative content (many of which are included in the Appendices of this paper to reduce the length of the thesis' main chapters). This rich narrative content is included to support a key goal of this project—to lay the foundation for a future (and natural) extension of this research, the creation of an interactive web map (i.e. an Esri Story Map). By providing this additional material that was collected in order to develop the geodatabase, content is readily available to inspire those who wish to continue telling their stories about West Hawaiʻi water quality.

This thesis contains five additional chapters. Following this chapter, Chapter 2 provides background information regarding unique West Hawaiʻi water quality topics and describes how water quality data is visualized in a geographic information system. Chapter 3 explains and presents contextual maps for each of the ten projects including a temporal analysis relating similar project types using timelines. Chapter 4 covers the design and implementation of the geodatabase. Chapter 5 details the GitHub public repository where the resultant geodatabase, methods and associated content can be downloaded by anyone. Finally, Chapter 6 wraps up this project with a list of limitations and future development proposals.

Chapter 2 Protecting and Assessing Near-shore Water Quality

This chapter provides a framework to understand the basis on which water quality data is collected and utilized in West Hawai‘i. There are several different kinds of documents that identify and protect the quality of the near-shore ocean water in the West Hawai’i region.

This chapter first describes the federal legislation and state guidelines that guide water quality monitoring in this region. Responding to these legislative and administrative guidelines, consultant scientists produce water quality assessments within Environmental Impact Statements (EIS) and scholarly researchers carry out research to understand the basis for water quality management in general. This chapter establishes the documentation background to support why I extracted data from EIS and research publications to include in a common geodatabase.

Once the water quality documentation framework that determined the nature of this geodatabase project is outlined, the data sources and analytes of interest are introduced. Next, an explanation of how water quality data is represented in GIS is provided.

2.1. Legislative Foundations

The U.S. Environmental Protection Agency (EPA) and the State of Hawai‘i Department of Health (HDOH) decree water quality requirements and regulatory activities on Hawai‘i Island. These provide the foundations for the development of the geodatabase.

2.1.1. Federal Legislation

Section §303(d) of the Clean Water Act (CWA, full title: Federal Water Pollution Control Amendments of 1972) (Congress H.R 1972, 105-106) requires states to establish water quality standards and submit pollution exceedance data to an Environmental Protection Agency (EPA) database (HDOH 2012b). The Clean Water Act was amended to include the Beach

Environmental Assessment and Coastal Health (BEACH) Act which requires coastal states to develop mechanisms for detecting pathogens along with requiring public notifications of any exceedances (US Congress 2000, 4-5).

The geodatabase developed in this project includes the fifty thousand analyte results that the Hawaiʻi Department of Health (HDOH) has collected since 1973 to support the Clean Water Act and BEACH Act. The exceedances are easily found in the data since each sample includes a true/false flag indicating whether a pollution event occurred and the values of the analytes in exceedance.

2.1.2. Hawaiʻi State Programs

The HDOH performs water quality surveys to support these two EPA programs. Specifically, the CWA is concerned with chemical measurements and the BEACH program is concerned with microbiological measurement results. As required by the CWA and BEACH Act, the State of Hawaiʻi has numerous documents outlining the standards for water quality surveys and the HDOH provides documents to guide all ocean water quality researchers. State programs regarding criteria specific to marine water quality are administrative rules (HDOH 2014), a description of the beach pathogen monitoring program (HDOH 2017a) and a quality assurance project plan (QAPP) (HDOH 2012a) specifying detailed scientific criteria.

2.1.2.1. Hawaiʻi Administrative Rules

While the HDOH provides detailed uses and criteria for marine water quality with specific rules and classifications for West Hawaiʻi locations (HDOH 2014, 45-55), the Hawaiʻi Administrative Rules (HAR) classify the water off West Hawaiʻi into two classes (see Figure 6):

- Class AA – water should remain as near as possible to its natural pristine state "with an absolute minimum of pollution or alteration of water quality from any human-caused source or actions" (HDOH 2014, 15-16)
- Class A – water should be protected for "recreational purposes and aesthetic enjoyment" (HDOH 2014, 16-17).

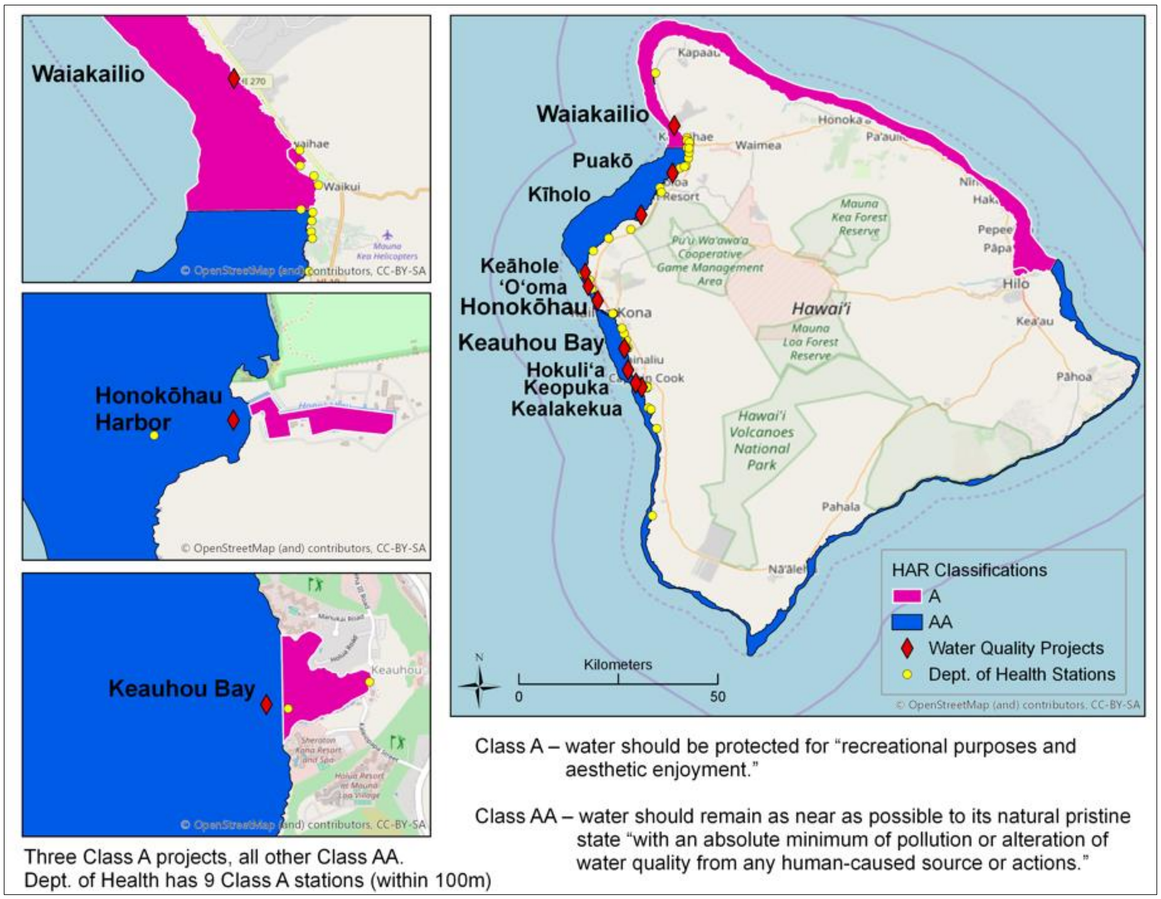


Figure 6. Water quality classes in Hawai'i. Points along the coast indicate the location of projects included in this thesis.

These A and AA classifications are important because there are different rules, separate criteria, and exclusion lists for specific areas. Regardless of the classification, these rules specify a minimum set of analytes, methods and measurement requirements. This regulation is important because scientists need to use this baseline of analytes, draw charts and make conclusions based on the relevant formulas and values.

The algorithms and constants in the publication *Rationale of the Development of Area-Specific Water Quality Criteria for the West Coast of the Island of Hawai'i and Procedures for their Use* are used by the scientists whose work is included in this study because "they are tailored to the Kona Coast and better represent achievable water quality goals for the area than do the present statewide criteria" (HDOH 1997, 15). The HDOH updated earlier HAR rules with new West Hawai'i site-specific exceedance formulas using known submarine groundwater discharge properties. This document is critical because the scientists in this study producing regulatory data use it to determine if a measurement taken at a location is in exceedance.

2.1.2.2. HDOH Beach Monitoring Program

Of interest to ocean users are beach closures due to pollution exceedances; they simply want to know if it is safe to go into the water. The scientists at the HDOH explain their beach testing protocol and notification procedures and practices in the Beach Monitoring Program (HDOH 2017a). If water quality results are in exceedance, the analyte values and the specific reason for the closing the beach or other near-shore waters are published by the HDOH. They use community partners, such as the Surfrider Foundation, universities, county lifeguards, and the visitor industry, to warn the public by posting signs on the beach and other methods such as email and social media. An on line web map provides a visual way to locate current advisories and provides more details on the extent of the pollution event and expected resolution (http://emdweb.doh.hawaii.gov/cwb/wqd/viewer/).

2.1.3. Quality Assurance Project Plan (QAPP)

As required by the CWA and BEACH Act, the QAPP document outlines detailed quality requirements a scientist must meet when analyzing and publishing water quality analyte results (HDOH 2012a). For example, various tables in that document describe how to perform field measurements and specify minimum standards for laboratory instruments. HDOH uses specific procedures and two bacterial analytes to monitor the beaches for pathogens. *Enterococcus* spp. is the indicator and *Clostridium perfringens* is the secondary tracer of sewage contamination.

2.2. Water Quality Data Sources

The purpose of this project was to assemble qualified water quality sample data from regulatory documents and scholarly researchers into one geodatabase. A key source of data are marine assessments which are a required component of Environmental Impact Statements. Another important source is reports by scholarly researchers who produce a wide range of data tables also intended to address the state of water quality in the West Hawaiʻi region.

It is important to note that all data used in this project are open and readily available to the public. The geodatabase created is also open and none of the data used are proprietary or used without permissions.

2.2.1. Environmental Impact Statements (EIS)

The Environmental Impact Statements (EIS) documents are rich and underutilized sources of water quality data and analysis (pers. comm. Richard Bennett, multiple dates). They are public information and available free of charge at a HDOH online archive.

An EIS is a wide-ranging set of environmental studies and survey information submitted to the government to assess the potential impact of a development. A private party associated with a potential land developer hires a water quality scientist experienced in the area to perform an analysis for publication within the EIS. The scientist, working as a private consultant, designs, implements and publishes a quantitative analysis of the additional nutrients the proposed project would add to the naturally occurring quantities. There are four EIS water quality assessments included in the thesis geodatabase produced by three well-known consultants.

2.2.2. Scholarly Publications

Data tables and analysis produced by scholarly researchers are an important source of water quality data tables in this thesis. These scholars typically publish tables containing only the geometric mean of their raw data, averaged at each location over the time range of their data collection period. The geometric mean (the *nth* root of the product of the *n* numbers), as opposed to the arithmetic mean (add the *n* numbers and divide by *n*), is almost universally used for recreational water quality data (Wymer 2007). Even though the raw data is generally not available, these averaged water quality data are published in tables in journals, reports, theses and dissertations. Therefore, the number of samples and the time range are key parameters to record.

For this thesis project, Dr. Tracy Wiegner (pers. comm. via email, April 24, 2018) and Leilani Abaya (pers. comm. via email, April 26, 2018), from UHH Marine Science Laboratory, and Fiona Langenberger (pers. comm. via email, August 10, 2018), from the Pacific Islands Ocean Observing System (PacIOOS) at the University of Hawaiʻi at Mānoa, agreed to publishing their raw academic data in the public geodatabase.

2.3. GIS and Ocean Water Quality Data

This section describes what water quality data looks like in a GIS and introduces the cartographic formats used in figures in this document. On a map, water quality stations are visualized as points on a line in the ocean, each point is a row in a table with latitude and longitude fields in a relational database management system (DBMS). For this water quality DBMS, there is a table of the locations (stations) where scientists collected each water sample for testing and separate related data tables for analyte results at these locations.

Typically, the stations are aligned along an imaginary line called a transect originating from the shoreline and extending out into the ocean. For example, Figure 7 shows a map of a water quality project in Keauhou Bay where the six red diamond station symbols are arranged in a line starting at the parking lot and then outward at progressively longer distances to 500 m into the mouth of the bay. Dr. Dennis Mihalka used a GPS device to navigate his kayak to these same locations every week to take water samples, perform measurements, and record his results in a spreadsheet. In all sampling location figures of this document, a red diamond symbol is used for station locations and the base map is from Open Street Map (www.openstreetmap.org).

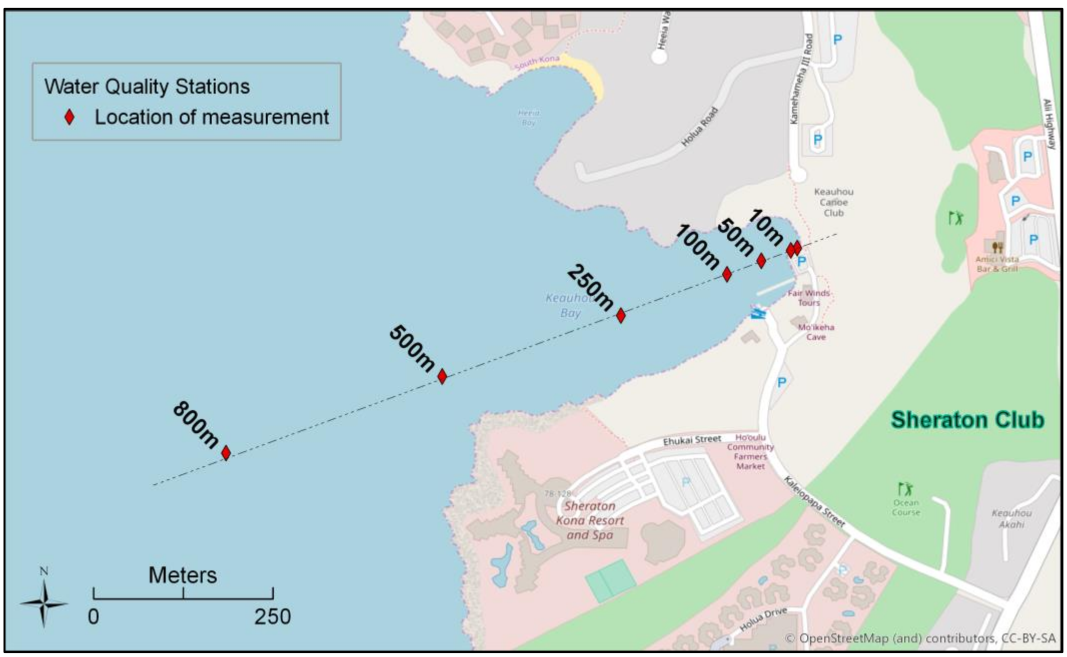


Figure 7. An example of station locations along a transect.

Using a larger time series project as an example, the scientist at Keāhole Point measures many transects designed around an aquaculture and energy research site (see Figure 8). This style of map displays different spatial contexts at various scales. At lower resolution scales (where the map covers a large area), the transects resemble single dots on the map; but at higher resolutions, the transects and individual stations become visible. At the higher resolution, the design of the transects and the stations along them are apparent (bottom two frames in Figure 8).

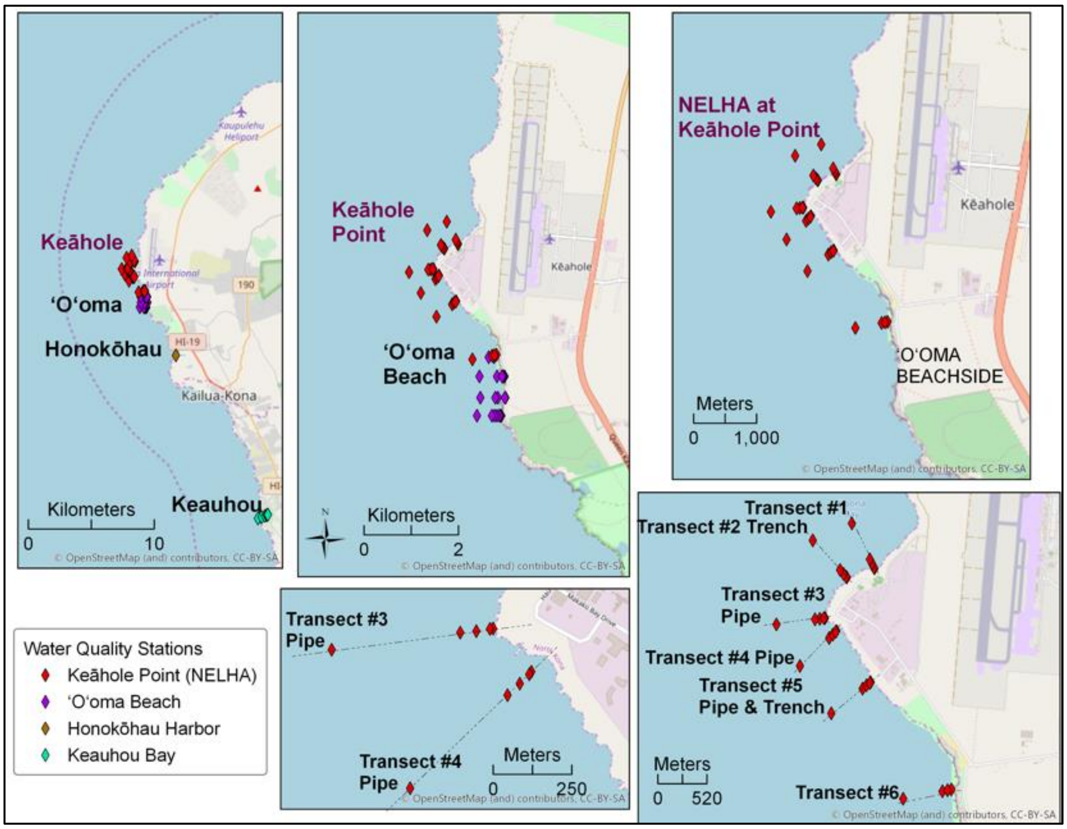


Figure 8. Water quality transects and stations at various spatial scales.

Inside the GIS, the data are organized into a location table and data tables. GIS software provides techniques to query the data associated with a location (see Figure 9). Chapter 4 describes how the geodatabase was designed using the relational links necessary to use the GIS geoprocessing tools and features for extraction of the sample results from station locations.

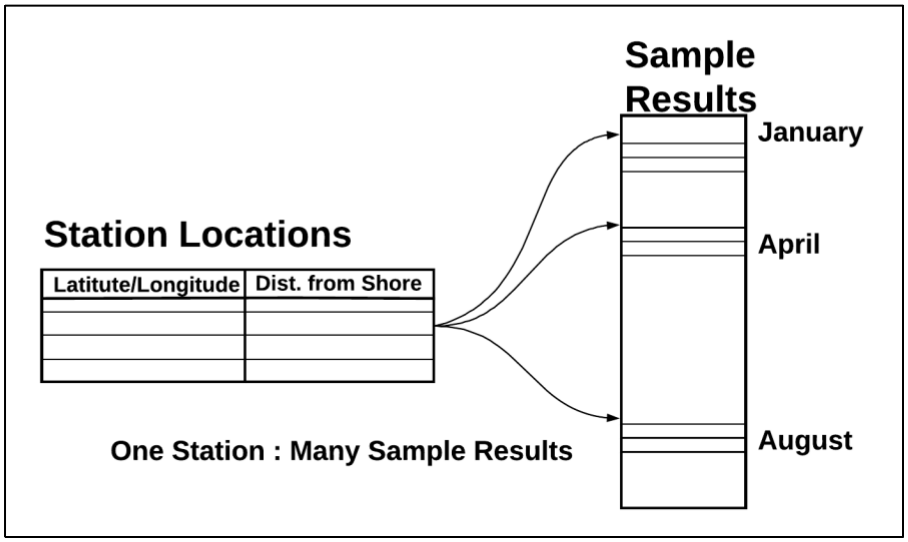


Figure 9. The one-to-many relational table design for water quality in the GIS.

Chapter 4 outlines the geodatabase schema design for the station locations and sample results tables and the methods to link them together. But first, the next chapter describes the projects, including their maps and images, that were integrated into the geodatabase.

Chapter 3 Background on the Water Quality Projects in the Geodatabase

Scientists like to study the West Hawai'i region due to its small size, consistent seasons, and arid climate (Lau and Mink 2006). The west region is favorable for water quality research because of the oligotrophic (i.e. normally low in nutrients) near-shore surface waters and a dominant source of nutrients that allows scientists to make conclusions about the larger environment. Thus, the critical factor guiding water quality studies in this region is the presence of submarine ground discharge (SGD). The scientists whose data are included in this geodatabase consider this unique subsurface hydraulic process as the dominant source of nutrients in their analyses. This chapter introduces each water quality project included in the geodatabase. Table 1 summarizes the transect and station counts for each project and notes the reason the data were collected.

Table 1. Transect and Station Counts for Water Quality Projects

ProjectID	Transects	Stations	Project Location	Notes
LTS1	6	30	Keāhole Point	Long term time series
LTS2	0	48	HDOH Sites	Long term time series
PC4	3	15	Waiakailio Bay	Private consultant project in Kohala district in Class A waters
UHH3	16	16	Puakō	Scholarly inquiry
UHH1	1	1	Kīholo Bay	Scholarly inquiry using buoys to create real-time network of data
PC3	3	15	‘O‘oma Beach	Private consultant project adjacent to Keāhole Point
CZ1	1	8	Keauhou Bay	Citizen scientist
UHH2h & UHH2k	8	4	Honokōhau Harbor & Kealakekua Bay	Scholarly inquiry project at two tourist locations
PC2	3	17	Keopuka Lands	Private consultant project in conservation land area
PC1	7	21	Hokuli‘a	Private consultant project surrounding a golf course
Control	0	1	ALOHA Station	Possible control site north of O‘ahu Island

The projects were categorized into five data source groups: State of Hawaiʻi, Hawaiʻi County, University of Hawaiʻi at Hilo, local citizens and a control site by University of Hawaiʻi at Mānoa. The purpose of the data collection, either regulatory monitoring or research inquiries, further differentiates the data sources. Regulations require scientists to produce data to monitor the quality of the ocean water in the public interest. Research inquiries support scientific objectives to better understand the environment and develop new monitoring mechanisms. The research inquires discussed in this chapter are different kinds of projects documented by scientists at the University of Hawaiʻi and a citizen scientist paddling a kayak.

Figure 10 shows the taxonomy of projects within the geodatabase. The bottom layer of the hierarchy (shown in the green boxes) are the specific kinds of projects presented in this chapter. The map in Figure 11 shows the kinds of projects and their locations (except the control site which is not on Hawaiʻi Island).

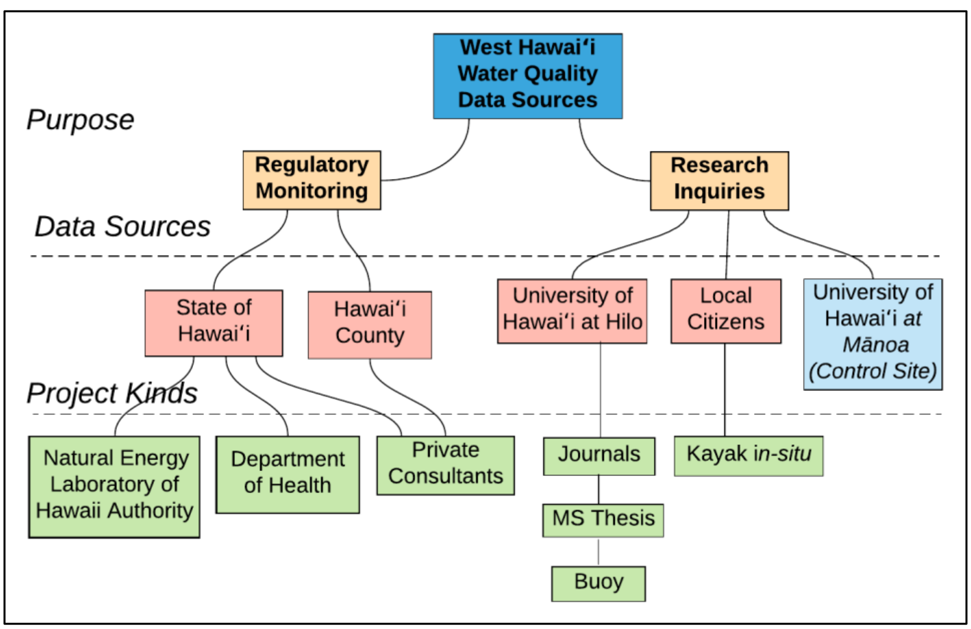


Figure 10. Water Quality Project Taxonomy

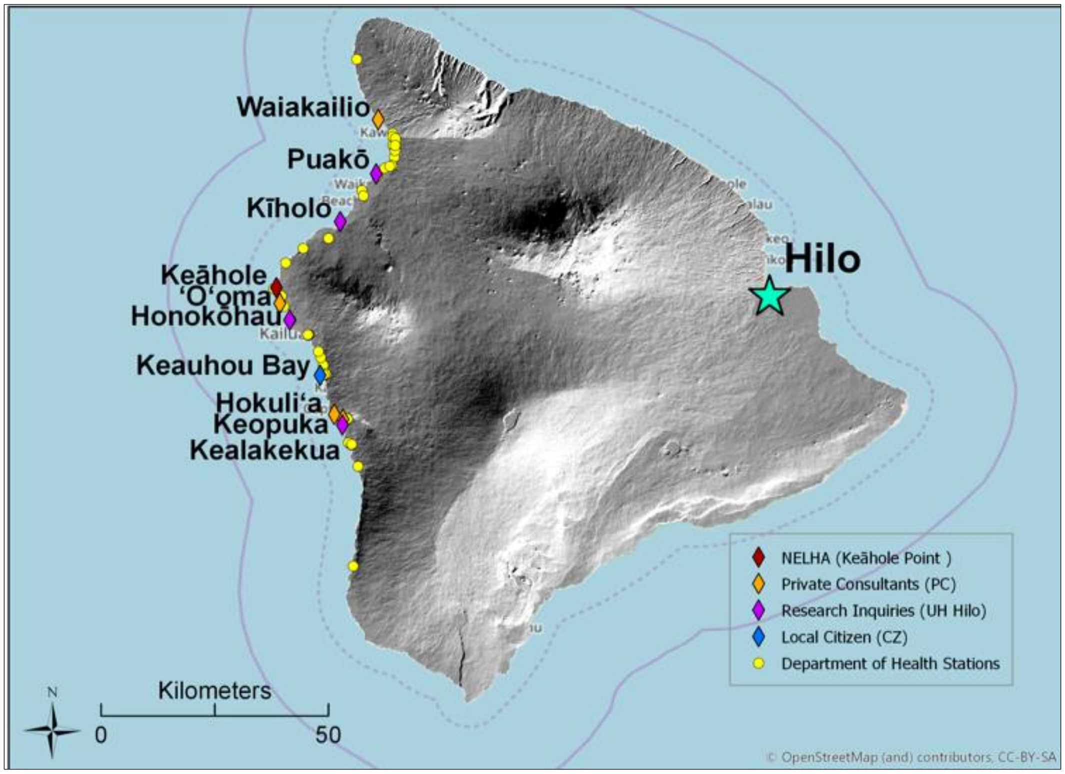


Figure 11. Project locations and types. The hillshade raster (contributed by Robert Whittier) highlights the volcanoes.

The primary contributions to the geodatabase include two regulatory long time series datasets: Natural Energy Laboratory of Hawaiʻi Authority (NELHA) and the Hawaiʻi Department of Health (HDOH). The data set produced by NELHA covers twenty-five years and the set produced by the HDOH covers over forty-five years. NELHA data set is not only very large but it is consistent in both analyte results and sample interval which makes the data important when researching changes over the twenty-five period at Keāhole Point. The HDOH data set is important for two reasons. this data is the source of the exceedances reported to the EPA and provides a record of all beaches that were closed because of pollution events.

The common analytes recorded by the scientists who produced the data for the geodatabase include nitrogen, phosphorus, chlorophyll a, oxygen, turbidity, temperature, silica, salinity, *Enterococci* spp. and *C. Perfringens*. These are summarized in Table 2. As noted earlier, the *Enterococci* spp. bacteria indicator has special importance in this area because the State of Hawaiʻi Department of Health (HDOH) bases beach closures and re-openings on the values of this analyte, as directed in the BEACH Act.

Table 2. The most common analytes. Source: USGS 2018 (except as noted).

Name	West Hawaiʻi Concerns
Salinity, silicates, temperature	Indicates submarine ground discharge properties (Lau and Mink 2006).
Enterococci spp., *C. Perfringens*	Bacteria are common single-celled organisms and are a natural component of lakes, rivers, and streams. Most of these bacteria are harmless to humans; however, certain bacteria, some of which normally inhabit the intestinal tract of warm-blooded animals, have the potential to cause sickness and disease in humans.
Total nitrogen, ammonium nitrogen, nitrate+nitrite nitrogen	Nitrogen, in the forms of nitrate, nitrite, or ammonium, is a nutrient needed for plant growth.
Total phosphorus	A common constituent of agricultural fertilizers, manure, and organic wastes in sewage and industrial effluent.
Chlorophyll a	Indication of primary biological productivity (Lau and Mink 2006).
Turbidity	Turbidity is the measure of relative clarity of a liquid.
Dissolved oxygen	Measure of how much oxygen is dissolved in the water.

Each section below presents a summary and visual representation of each project. Ten projects and a control site were chosen for the initial effort to create a water quality monitoring system and geodatabase for the West Hawaiʻi region. Dr. Bennett chose most of the sites and we discussed the importance of each site in weekly conference calls. I created a contextual map of each project and a matching appendix for each project provides additional narrative content and figures intended to contribute to the eventual development of a multi-media map presentation to share water quality information and science with the public. It is important to note that most maps included in this chapter and the appendices were generated from the geodatabase whose construction is described in the next chapter.

3.1. Natural Energy Laboratory and Department of Health Projects

The geodatabase contains two regulatory time series projects from NELHA and HDOH (see Figure 12). A comprehensive dataset from Keāhole Point was contributed by Keith Olson from NELHA. The consistent sampling interval, analyte types that match state law, and the spatial distribution of transects surrounding the project area make this robust data well-suited for the geodatabase (GDB) for this data integration project. It provides a foundation for comparing analyte values against all the other projects since there is always data available from NELHA for any time range.

The HDOH data is important because of the extended duration and bacteria analytes that determine whether to close polluted beaches and transmit advisories out to the public. The scientists supporting this study expressed a need for data to understand the sampling pattern before and after beach closure events. While the HDOH data set is not as consistent in sample interval and sampling distribution as NELHA, they represent different agendas. The HDOH dataset is reconciling two EPA programs for multiple locations along the coast.

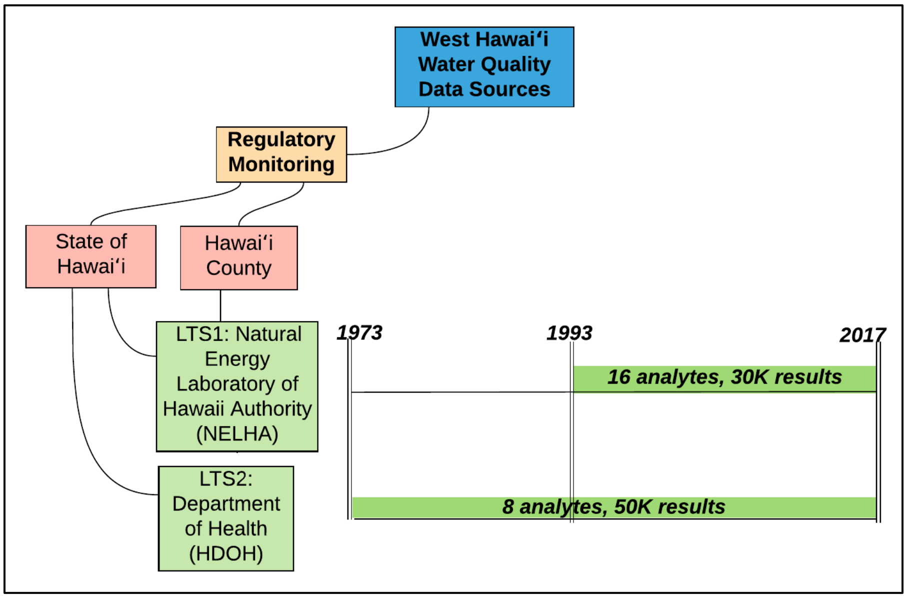


Figure 12. The long time series regulatory projects by the Natural Energy Laboratory (LTS1) and Department of Health (LTS2)

3.1.1. NELHA sampling at Keāhole Point (LTS1)

NELHA is the largest seawater operation in the world. Pipes draw cold, nutrient rich water from deep in the ocean to support numerous aquaculture and commercial operations. The waste water is then discharged via trenches back to the ocean. As part of their water quality monitoring protocol, NELHA scientists collect oceanic water samples at thirty stations every ninety days (pers. comm. with Keith Olson, December 2017) (Figure 13 and Appendix A). Mr. Olson designed the transects around the intake pipes and discharge trenches and records analyte results matching the state regulatory requirements.

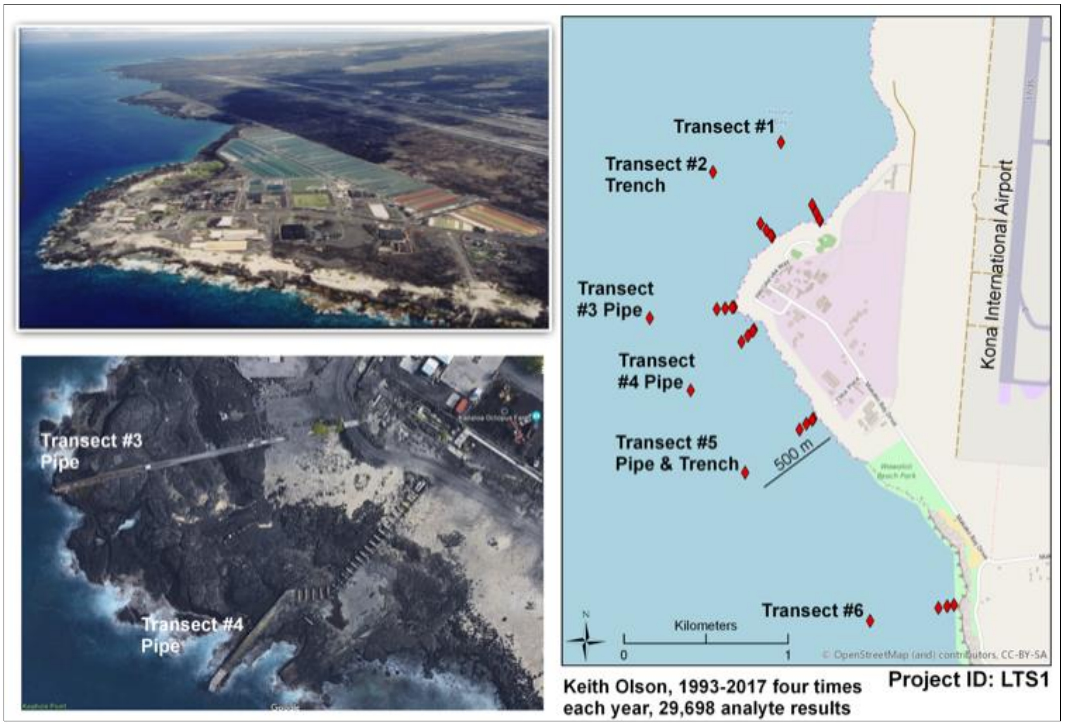


Figure 13. Keāhole Point long time series project (LTS1). Data Source: pers. comm. Keith Olson.

There are six transects each with five stations arranged around Keāhole Point, located near trench and pipe locations. The analytes recorded are: ammonia nitrogen, chlorophyll a, nitrate+nitrite nitrogen, dissolved oxygen, phosphorus, pH, salinity, silicates, total nitrogen, total dissolved phosphorus, temperature, and turbidity.

3.1.2. Department of Health sampling (LTS2)

The HDOH Clean Water Branch takes surface ocean samples up and down the coast with most locations right off a beach or pier (see Figure 14 and Appendix B). There are data recorded for a total of 48 stations in this series spread all along the coast. HDOH employees wade into the ocean with a bottle on a stick to collect a water sample about a meter below the surface. The sample bottle is then sent to a laboratory for processing and the analyte results are published on a public data portal. I developed a SQL method to extract, import and manipulate this data to a set of tables for incorporation into the geodatabase.

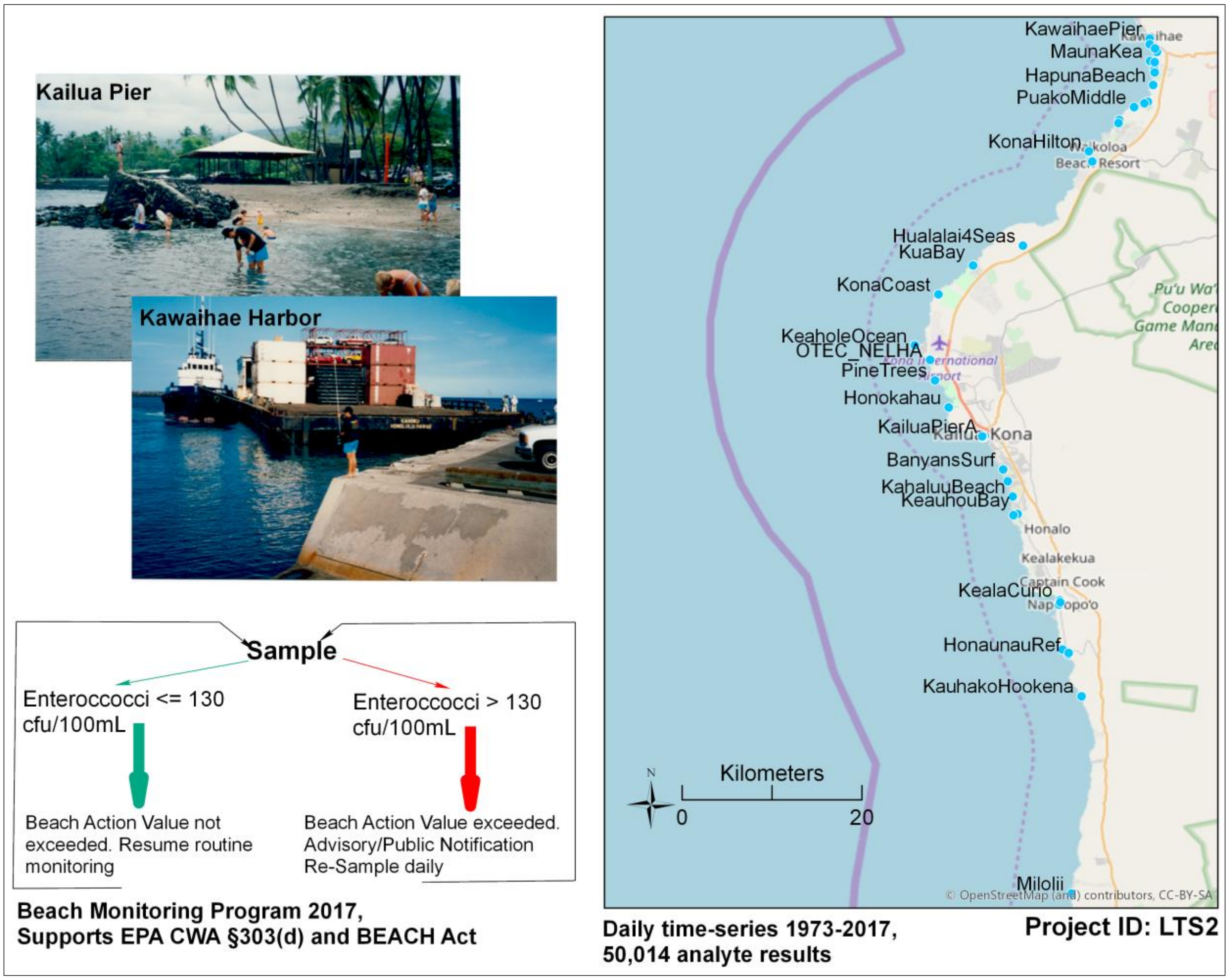


Figure 14. HDOH long time series project (LTS2). Station data source: PacIOOS, www.pacioos.hawaii.edu/metadata/cwb_water_quality.html.

The analytes recorded are: *Enterococcus* spp., *C. Perfringens*, temperature, salinity, turbidity, pH and oxygen along with a true/false 'Alert' flag on each sample. The Alert flag is true if the analyte results are in exceedance as determined by criteria in the Beach Monitoring Program document.

This project is the backbone of the thesis. It supports the requirement to be able to determine changes in quality over time and provides the background data on the specific values for analytes regarding beach closures to the public. The decision to close a beach for pollution events is directly related to this data.

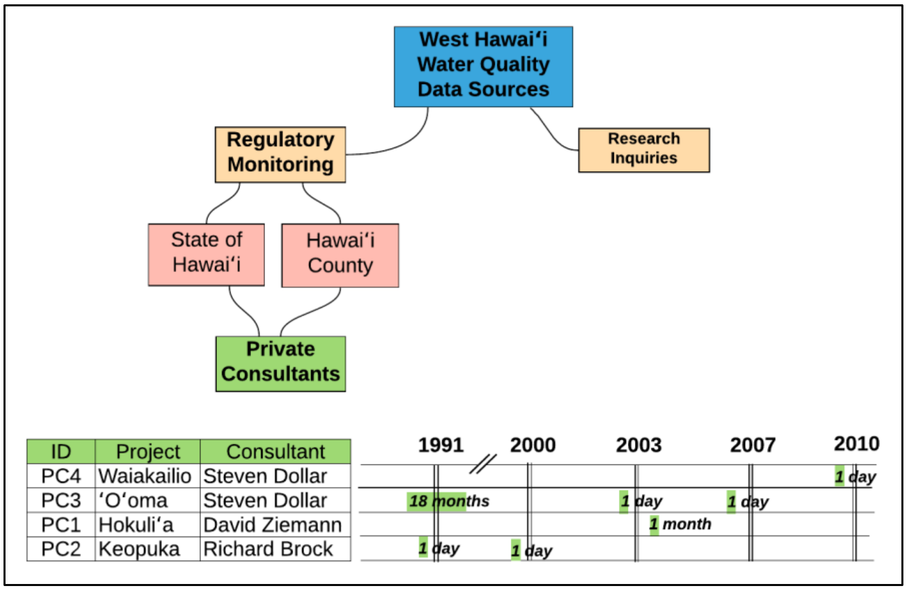
The next section also describes time series data although the sampling times are not as repetitive (or long term) since private consultants' projects are undertaken for a specific proposed development.

3.2. Private Consultant Projects

Private consultants use multiple transects and choose station lengths to cover the extent of the SGD gradients to record analyte results conforming to regulatory requirements. The purpose of these projects is to determine if a proposed shoreline development will have significant effect on the environment, ending with a conclusion recommending whether or not the government should approve the project.

The private consultant projects are rich sources of both data and area specific analysis. Each project is unique in analytical richness and the determination of whether the development should proceed. I attempted to identify the specific characteristics of project designs using the maps in the individual projects in the following sections. I wrote methods to extract the data tables from PDFs and integrated these into the geodatabase for this thesis.

Figure 15 shows the timeline of each project with some common dates in 1991-1992 and 2003-2004 ranges. Both Dr. Steven Dollar and Dr. Richard Brock included in the referenced reports multiple data tables from previous surveys to provide a time series baseline for future researchers.



ID	Project	Consultant
PC4	Waiakailio	Steven Dollar
PC3	ʻOʻoma	Steven Dollar
PC1	Hokuliʻa	David Ziemann
PC2	Keopuka	Richard Brock

Figure 15. Private consultant projects

3.2.1. Waiakailio Bay (PC4)

Dr. Steven Dollar created a baseline data set for an EIS proposal for a shoreline development project (Dollar 2015, 100) (Figure 16 and Appendix C). I decided to include this site because Dr. Dollar provided extended narrative written in a direct understandable style and included discussion and analysis about an important gulch close by. He stated that the Honokoa Gulch is important to water quality scientists and he moved his transect #3 farther south to provide a baseline for future scientists interested in studying this gulch.

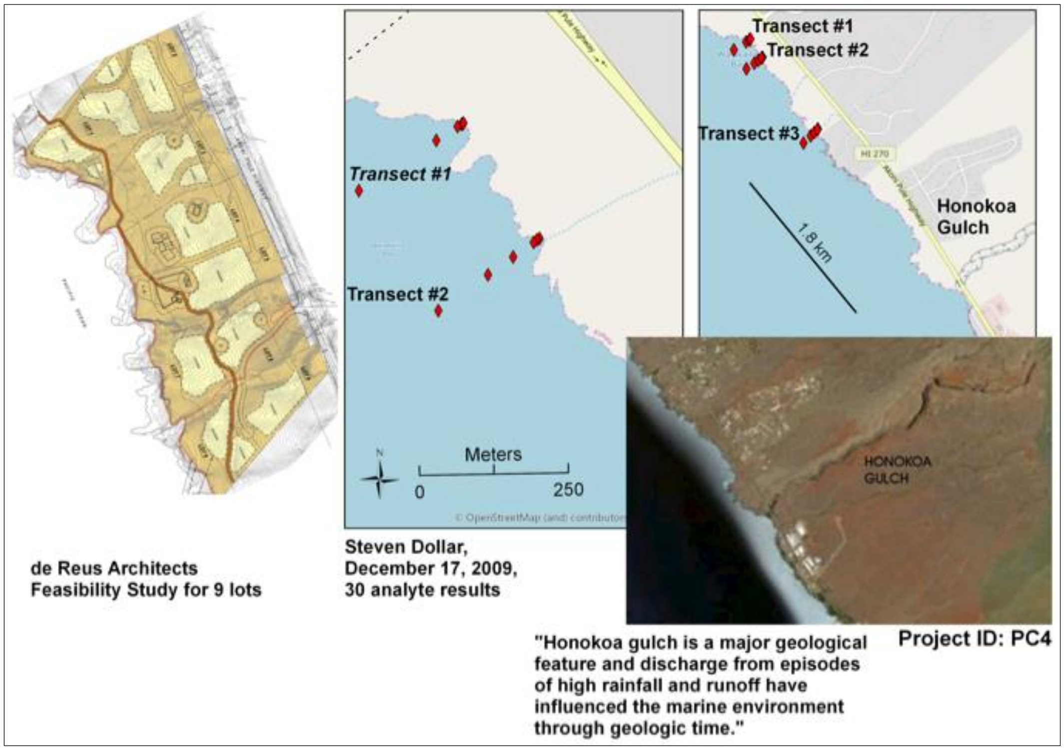


Figure 16. Waiakailio Bay, private consultant project (PC4). Station data source: Dollar 2010.

There are two transects located at the subject property and another about 1 km south. The station distances on the transects are shorter than other projects because the SGD seeps out right near the shoreline. The analytes included are: ammonia nitrogen, chlorophyll a, dissolved organic nitrogen, NO_3, dissolved oxygen, O2 saturation, 0-phosphate, pH, salinity, silicates, total nitrogen, temperature, total phosphorus, and turbidity.

3.2.2. ʻOʻoma Beach (PC3)

Dr. Steven Dollar also created a long time series baseline data set for a proposed development project at ʻOʻoma Beach (Dollar 2008) (see Figure 17 and Appendix D). He included three tables of analyte results from the current survey along with previous surveys covering a 16-year range. The proximity of this site is significant because it is adjacent to NELHA and the normal current flow is from the North to the South (as expressed to me by the collaborating scientists).

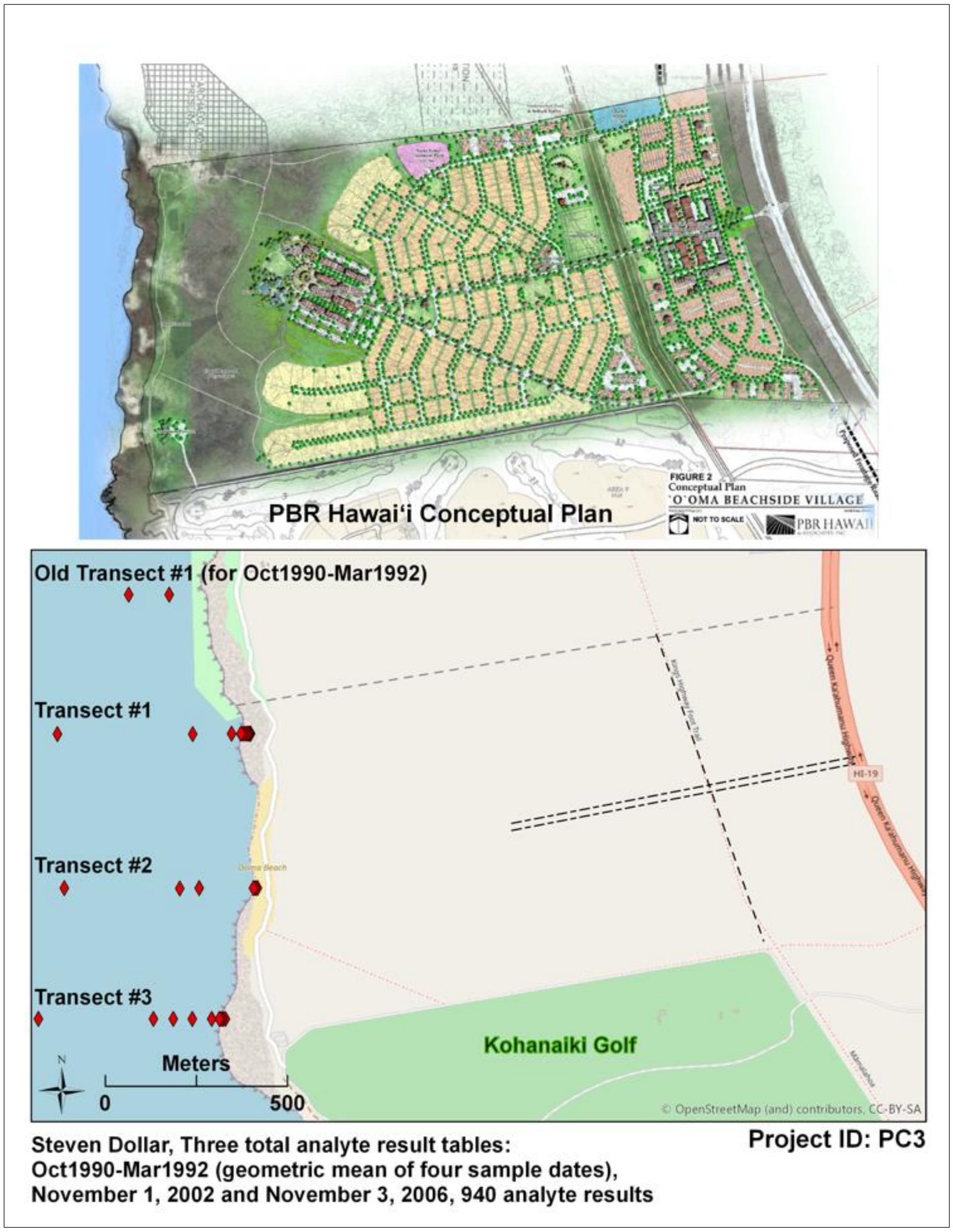


Figure 17. ’O’oma Beach, private consultant project (PC3). Station data source: Dollar 2008, Conceptual plan: PBR Hawai‘i 2009.

3.2.3. Hokuli‘a (PC1)

During August 2003, Dr. David A. Ziemann (Ziemann and Klein 2003, 1-26) took ocean samples from a boat at seven transects adjacent and north of the Hokuli‘a golf course (see Figure 18 and Appendix E). Their reason for having so many transects was to search for differences in nutrient discharges by two nearby golf courses (Sheraton and Hokuli‘a Golf Clubs). The main purpose (by Court order) was to measure the contribution of additional nutrients sourced from the operation of the Hokuli‘a Golf Club.

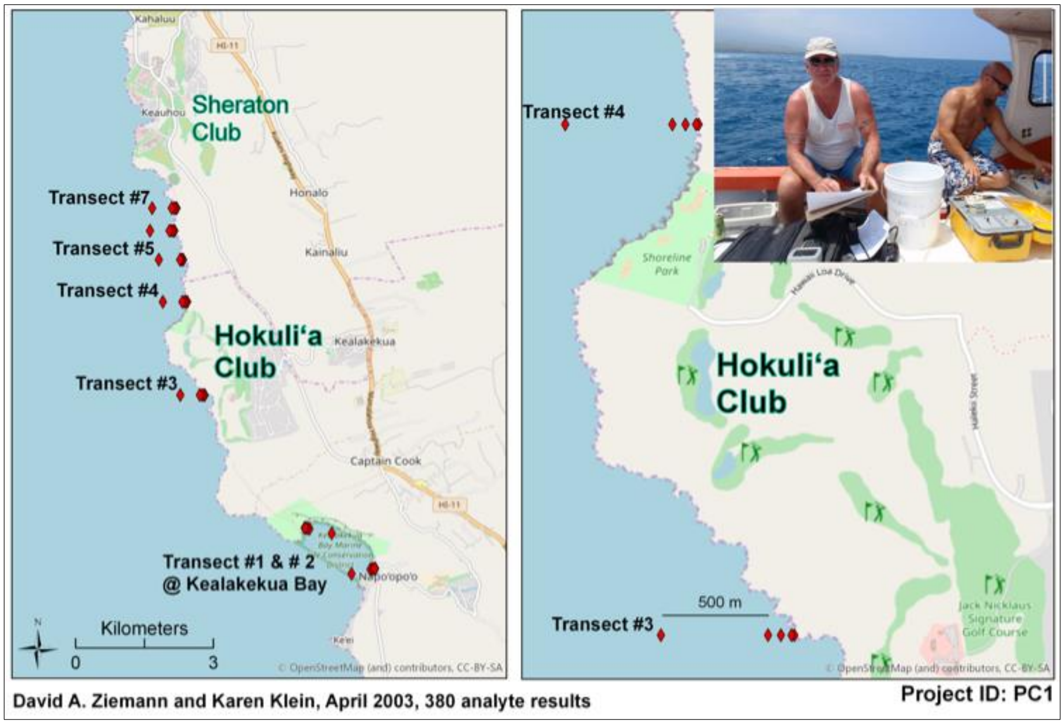


Figure 18. Hokuli‘a, private consultant project (PC1). Station data source: Ziemann and Klein 2003, Picture in right frame: Richard Bennett.

The analytes included are: total nitrogen, ammonia nitrogen, nitrate+nitrite nitrogen, total phosphorus, chlorophyll a, turbidity, dissolved oxygen, temperature, salinity, silicates, pH, NO_3 and 0-phosphate.

3.2.4. Keopuka Lands (PC2)

Dr. Richard Brock recorded water quality results at Keopuka Lands for a proposed nine lot development (Brock 2000, 35-47) (see Figure 19 and Appendix F). He shared these two surveys to establish baseline conditions at Keopuka. Dr. Bennett stated this project could be considered a control site for this study area because the property is still natural (undeveloped).

Another reason for consideration as a control site is this location is the proximity to Kealakekua Bay and Hokuli'a which are covered by two other projects in this thesis.

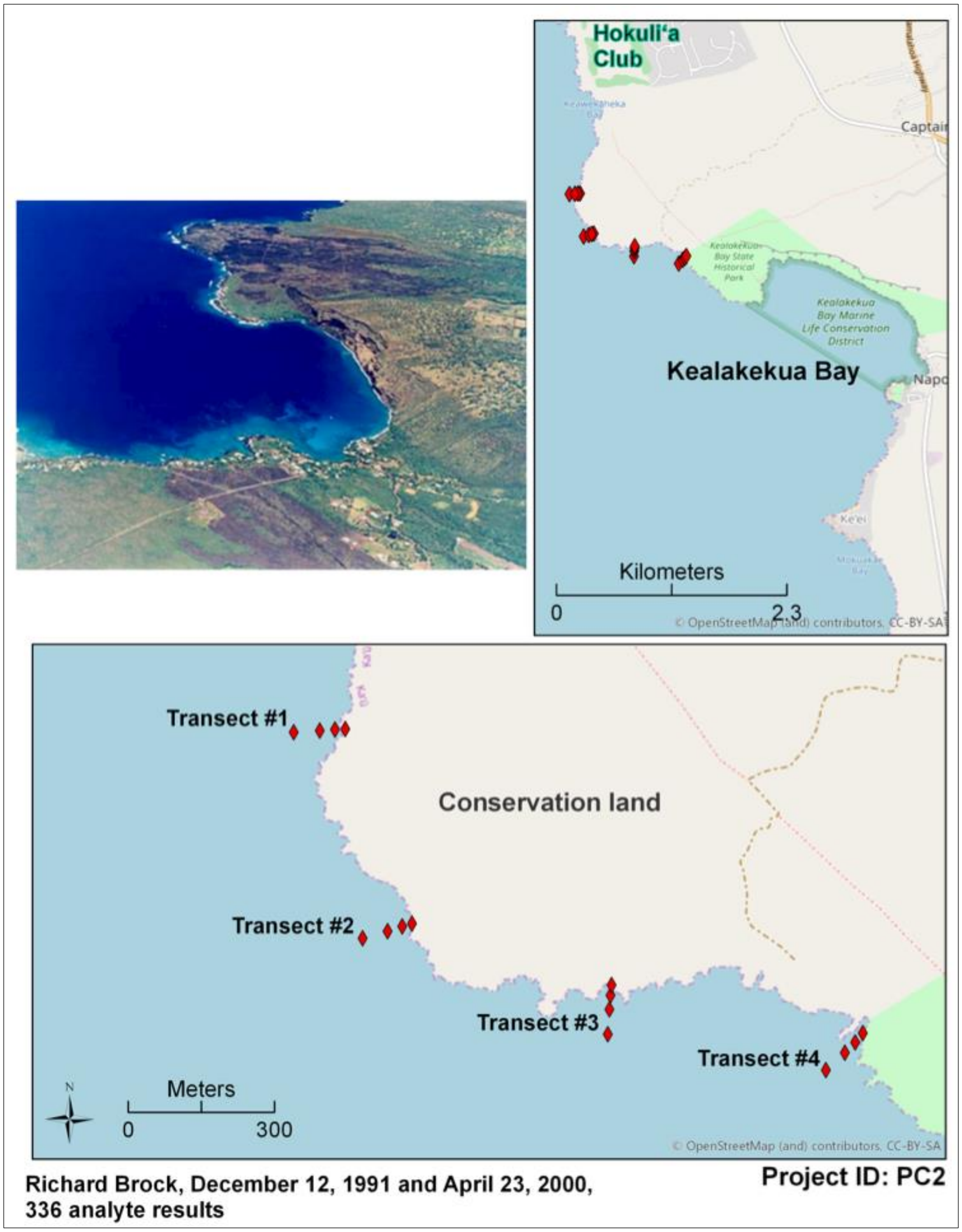


Figure 19. Keopuka, private consultant project (PC2). Station data source: Brock 2000, Aerial Picture: Star Bulletin (http://archives.starbulletin.com/2000/09/15/news/story4.html).

The analytes included are: ammonia nitrogen, chlorophyll a, dissolved organic nitrogen, NO_3, dissolved oxygen, O_2 saturation, 0-phosphate, pH, salinity, silicates, total nitrogen, temperature, total phosphorus, and turbidity.

This completes the overview of the regulatory projects. The next section covers the scholarly researchers who have moved beyond the requirements of the laws to measure additional analytes and test new techniques to track water quality changes.

3.3. Scholarly Research Projects

This thesis includes research inquiries from scientists at the University of Hawai‘i at Hilo (UHH) and a citizen scientist (see Figure 20 and Appendix G-J). The UHH data sets include data collected and summarized over 1 to 15 month periods. Since Dr. Mihalka’s project, while also a short time series conducted for research purposes, is categorized separately as citizen science, it is discussed separately in the next section.

The data for scholarly projects is different than regulatory ones in both the spatial distribution, temporal scale and analyte results. Scholarly researchers choose the transect and station locations based on a pre-survey of the study area to identify the SGD locations often using salinity, temperature and pH. Each of the research projects included here have published only one table of data to represent the data collected over the entire date range of the project, giving the geometric mean of all the individual samples at each sampling location. The scientists also introduce additional analytes such as $\delta^{15}N$ and pheophytin to support their research objectives.

As noted above, scholarly scientists provide a geometric mean of the data instead of the raw results for each sample day. Therefore, the temporal properties of the data extend for the entire duration of the project and the geodatabase attribute *StartDate* and *StopDate* store these values (see Chapter 4 for schema details). For example, in Figure 20, it can be seen that the Puakō project covers November 2015 to July 2016 as specified in Leilani Abaya's thesis tables (see the green timeline bar in Figure 20).

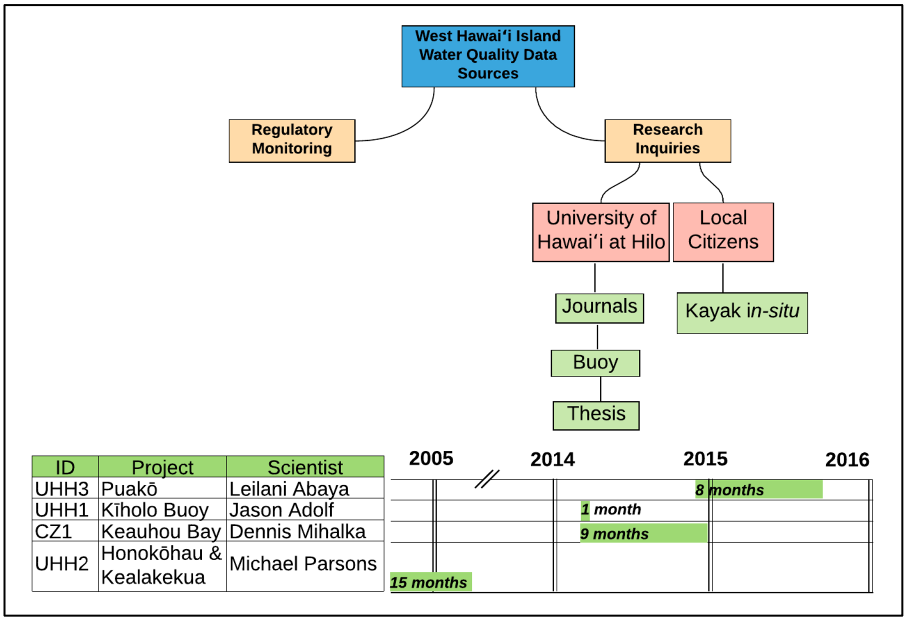


Figure 20. Research inquiry projects

3.3.1. Kīholo Bay Buoy (UHH1)

This project includes a network of buoys relaying water quality analyte results to a website intended to lay the foundation for a system of near real-time monitoring in the future (see Figure 21 and Appendix G). Dr. Jason Adolf deployed a buoy in Kīholo Bay for a month and the Pacific Island Ocean Observing System (PacIOOS) data server published the results via their W3C portal (www.pacioos.hawaii.edu). I developed a SQL query method to export, convert and integrate this data into the geodatabase.

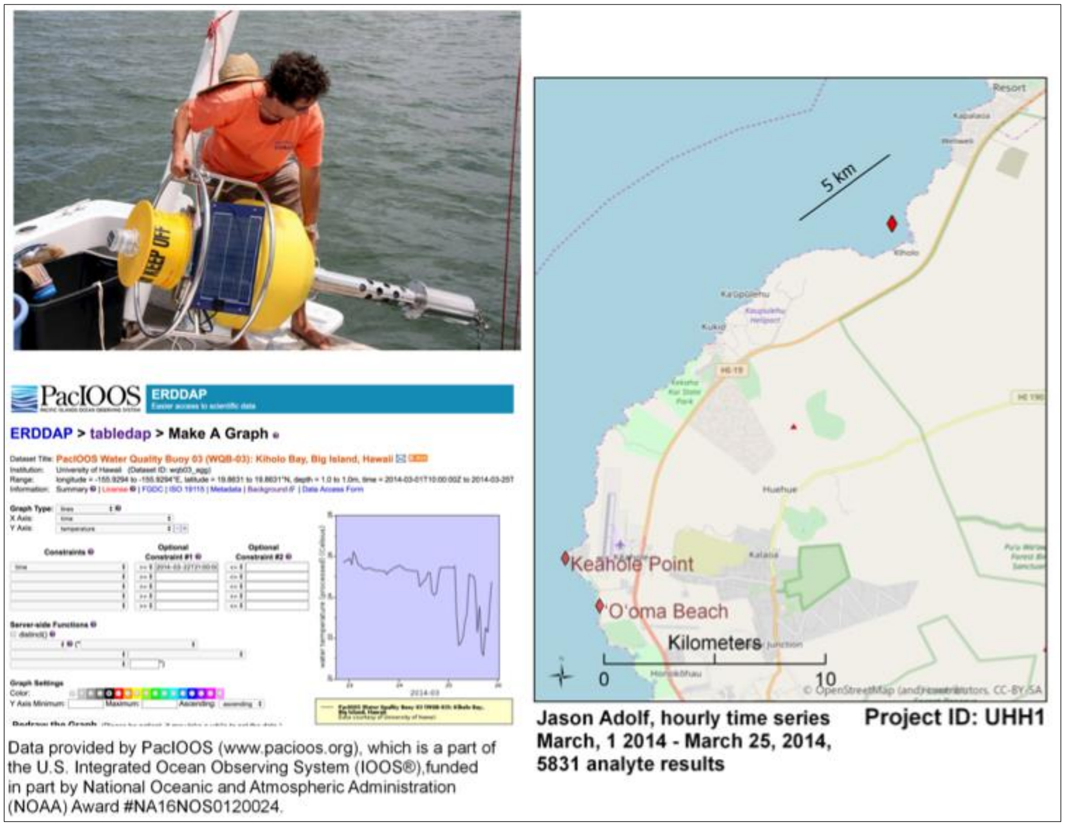


Figure 21. Kīholo Buoy, research inquiry project (UHH1). Station data source: PacIOOS, www.pacioos.hawaii.edu/metadata/WQB03agg.html, Picture: https://hilo.hawaii.edu/keaohou/2013/10/15/adolf-marine-science/, accessed July 17, 2018.

The buoy sampling mechanism records 10 analytes every hour. The analytes are temperature, conductivity, salinity, dissolved oxygen, oxygen saturation, oxygen saturation concentration, turbidity, chlorophyll a, chromophoric dissolved organic matter, and nitrate.

3.3.2. Puakō (UHH3)

A team from UHH including Leilani M. Abaya put a fluorescein dye in cesspools of shoreline residences and then traced the flow to the ocean (Abaya at al. 2008, 335-347). She did this for her 2016 Masters' thesis *Identifying Hotspots of Sewage Pollution in Coastal Areas with Coral Reefs*. This project compared water quality on land and in the ocean while measuring the effluent flow rates from the cesspools into the ocean. Leilani Abaya's study area was right next to a row of houses along a particularly porous rock shoreline (see Figure 22 and Appendix H).

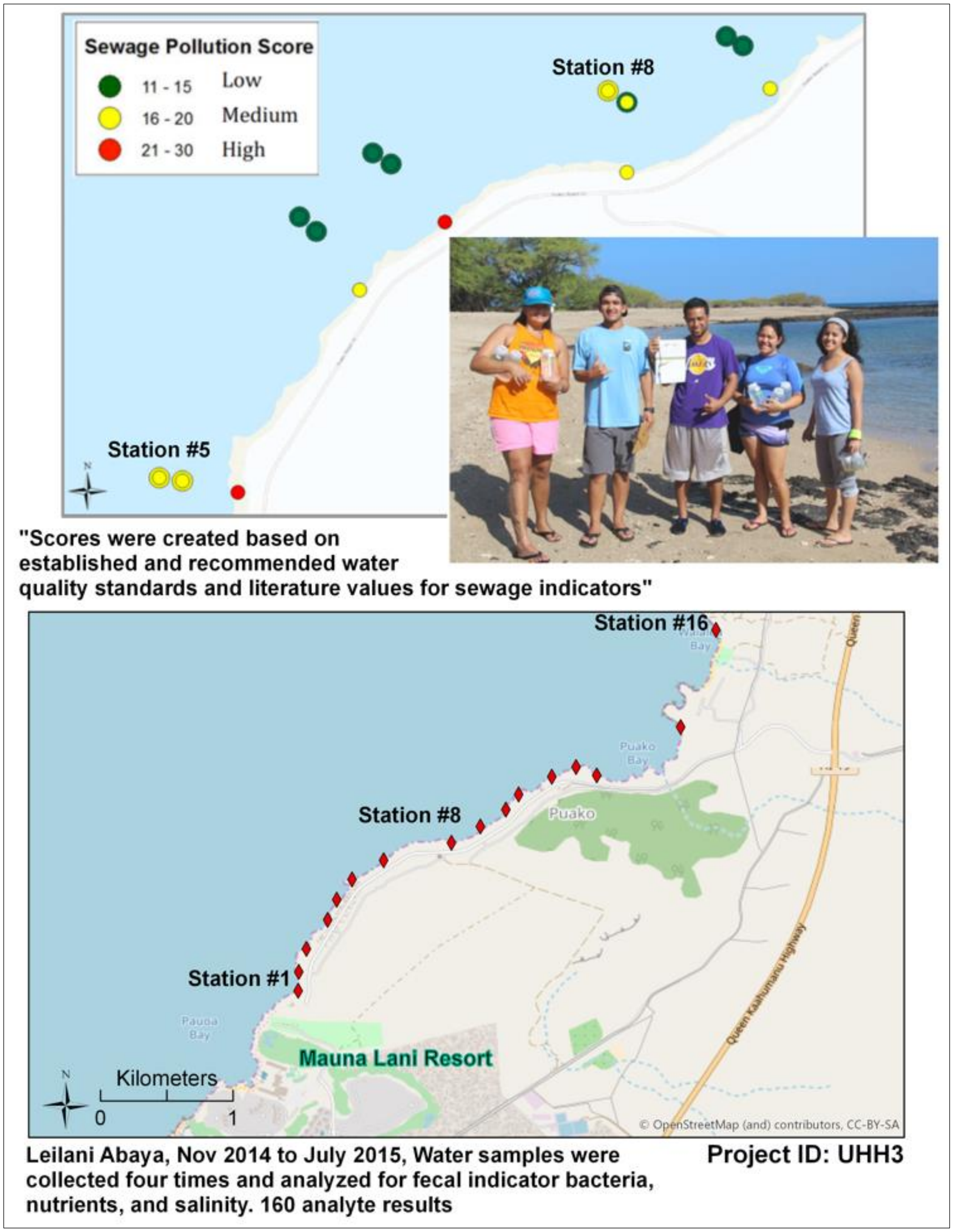


Figure 22. Puakō, scholarly research project (UHH3). Station data source: Abaya 2016, Picture in top frame: http://hilo.hawaii.edu/news/stories/2016/05/20/ leilani-abayas-master-thesis-defense/, accessed December 1, 2017.

The main objective of this research was to map sewage pollution at sixteen locations along the Puakō shoreline using three fecal indicators: *Enterococcus* spp., *C. Perfringens,* algal tissue $\delta^{15}N$. She introduces a new GIS symbology mechanism using these fecal indicators. This new combination of three indicators quantifies results into ranges as Low/Medium/High by assigning scores to polluted locations. Research about coral reefs is beyond the scope of this geodatabase, but the uniqueness of her analyte set, the introduction of a scoring algorithm to be used for map symbology, and determination of flow rates from specific cesspools adds more depth and diversity to water quality research in the study area and could provide a model for future projects.

The analytes include ammonia nitrogen, *C. Perfringens*, *Enterococci* spp, H4Sio4 silicic acid, $\delta^{15}N$, nitrate+nitrite nitrogen, phosphorus PO4, salinity, total nitrogen, and total phosphorus.

3.3.3. Honokōhau Harbor (UHH2)

Michael Parsons and colleagues from University of Hawai‘i at Hilo and the Hawai‘i Department of Land and Natural Resources collected samples over a fifteen-month period and then published the analyte results in a table in a professional journal (Parsons et al. 2008, 1138-1149). They wanted to determine if water quality posed threats to the ocean ecosystem by studying the coral reef at this popular recreational boat harbor which is close to an effluent disposal trench (see the “Disposal Pit” in middle map in Figure 23). To perform their coral reef assessment project, they arranged eight transects around the mouth of the bay near known SGD locations.

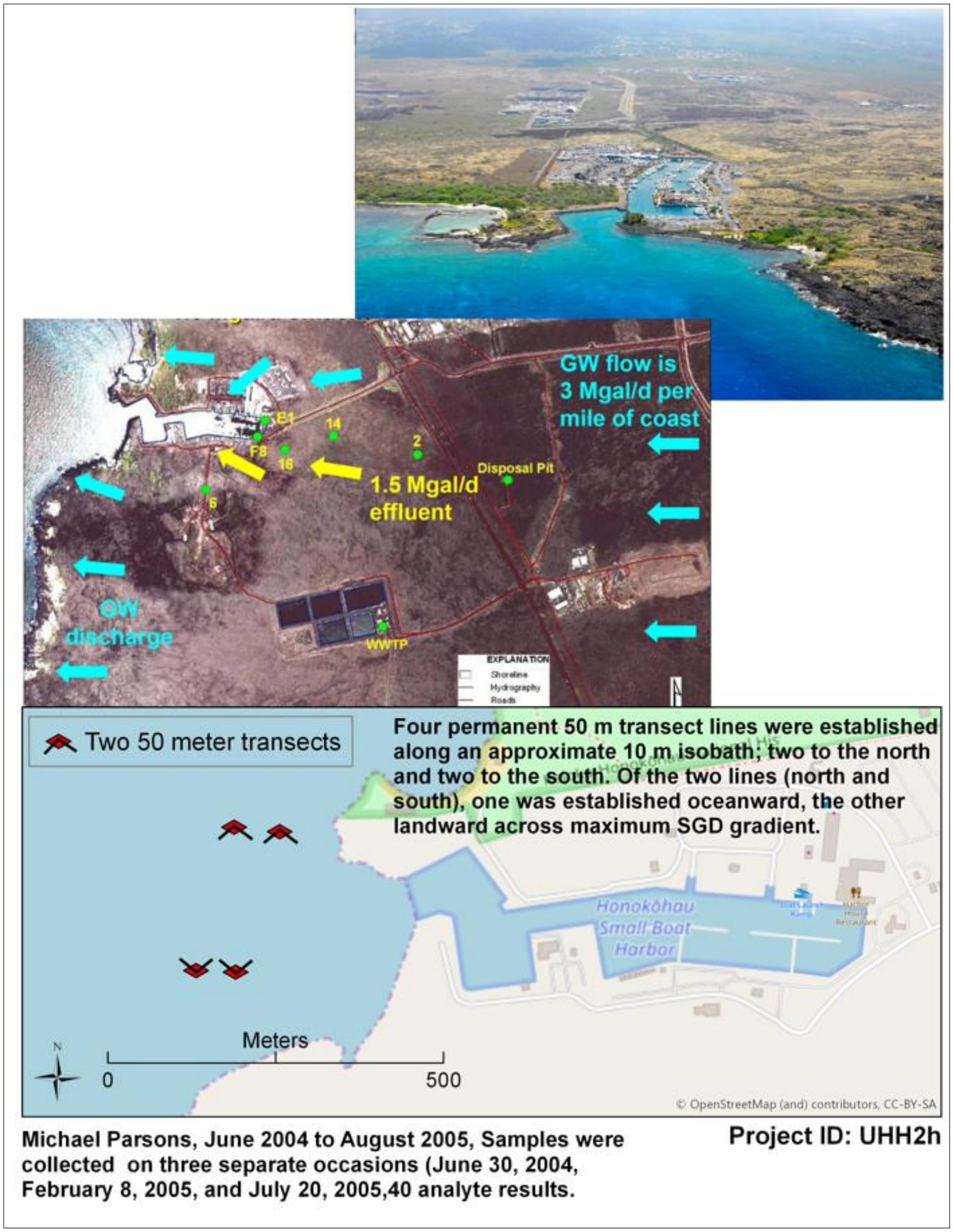


Figure 23. Honokōhau Harbor, scholarly research project (UHH2). Station data source: Parsons et al. 2008, Hunt 2018. Notation on the photograph is from the original source. Picture at top: from Marinas.com, https://marinas.com/view/marina/2gc5w6_Honokohau_Marina_and_Small_Boat_Harbor_Kailua-Kona_HI_United_States#&gid=1&pid=4, accessed August 13, 2018.

The analyte result table contains the geometric mean of the results from three sample collection days for the two transects at each of the four transect locations. The analytes include:

ammonia nitrogen, chlorophyll a, $\delta^{15}N$, nitrate+nitrite nitrogen, pheophytin, salinity, silicates, total nitrogen, temperature, and total phosphorus.

The next section discusses Kealakekua Bay which is located about 25 km south of Honokōhau Harbor. This is also a part of Parson's project where the data are recorded on the same dates using similar protocol.

3.3.4. Kealakekua Bay (UHH2k)

Parsons et al. (2008) performed the same sample regime at Kealakekua Bay as Honokōhau Harbor and the analyte results for both sites were merged into the same table of the journal article (see Appendix I for aggregated input table). Kealakekua Bay is a very popular recreational area, hosting many swimmers, kayaks and tour boats to view dolphins, coral and the Captain Cook Monument, which marks the location where the famous explorer was killed. The published table contains data collected using the same transect design and analyte sets (see bottom part of Figure 24). The bluish colors in the top map in Figure 24 shows the ocean water is colder which means there is more cool submarine ground discharge in that area. Parsons arranged the transects near and across the higher SGD locations.

The proximity of the Keopuka project adjacent to the north should be noted. A conclusion and warning from Parsons and colleagues highlights the northern transects because of their elevated nutrient concentrations affecting the coral ecosystem.

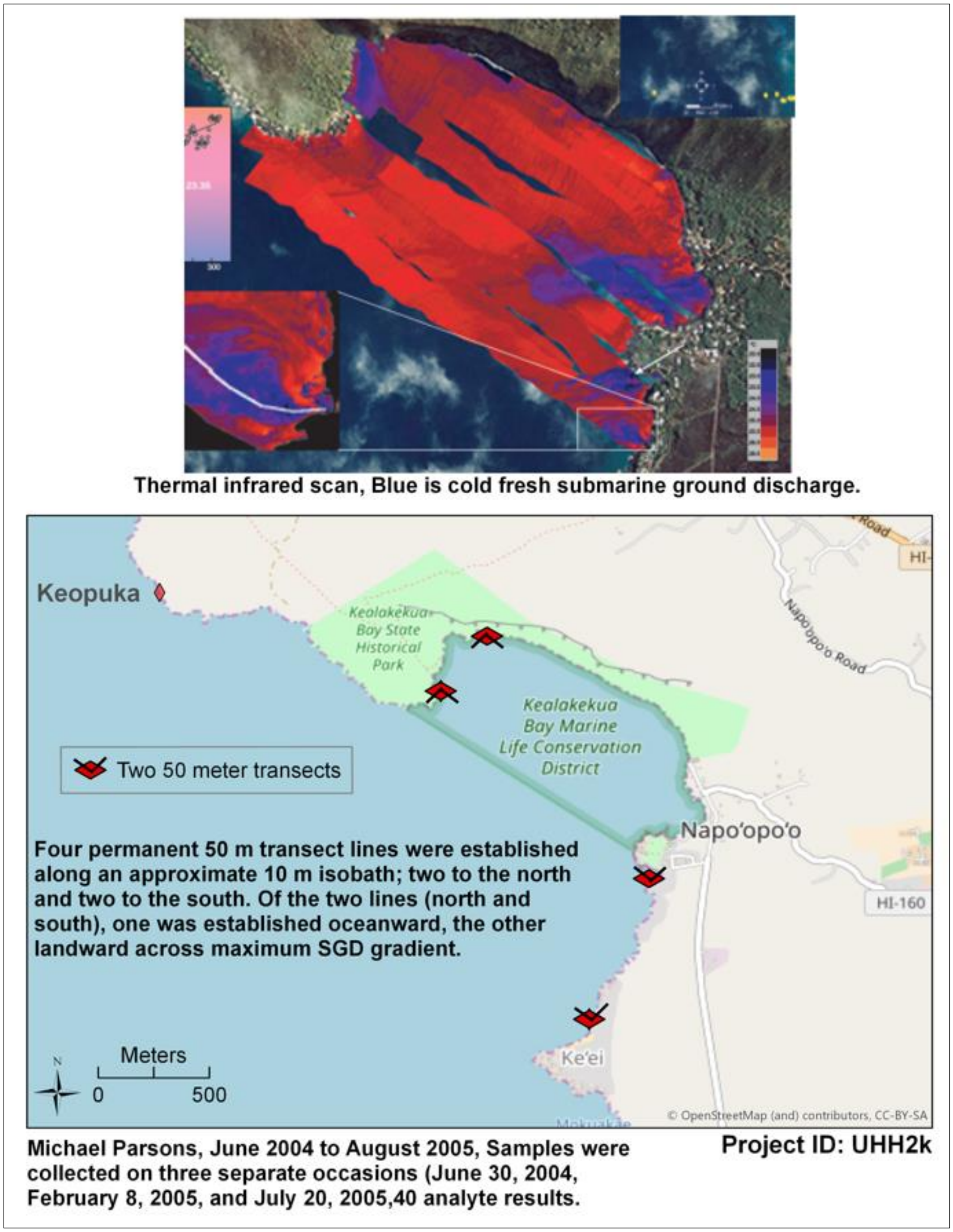


Figure 24. Kealakekua Bay, scholarly research project (UHH2k). Station data source: Parsons et al. 2008, Thermal image: Johnson et al. 2015.

The next section discusses the citizen inquiry project performed by a collaborating citizen scientist.

3.4. Citizen Science Research

Only one project of the citizen science research type is included in the geodatabase. Dr. Dennis Mihalka DDS is a local citizen scientist who is active in water quality issues on Hawaiʻi Island. He produces a show on community television presenting information related to water quality and has personally collected water quality data in Keauhou Bay, a popular canoeing, kayaking and paddle boarding location that is also the anchorage and launching spot for several dolphin and whale watching boats.

3.4.1. Keauhou Bay (CZ1)

For his data collection in Keauhou Bay, Dr. Mihalka has taken photos with a GoPro camera at each station along his transect in this embayment in order to share his observations with the public (see Figure 25 and Appendix J). His camera design used a fixed 12" length arm on which was attached a white board with fine lines etched on it. This provided for readings to be taken at a consistent depth that could show differences of color and turbidity as well as the background under the water beyond the white board. He arranges photographs into sequences to create visual stories to teach people about the turbidity in Keauhou Bay (see the bottom of Figure 25).

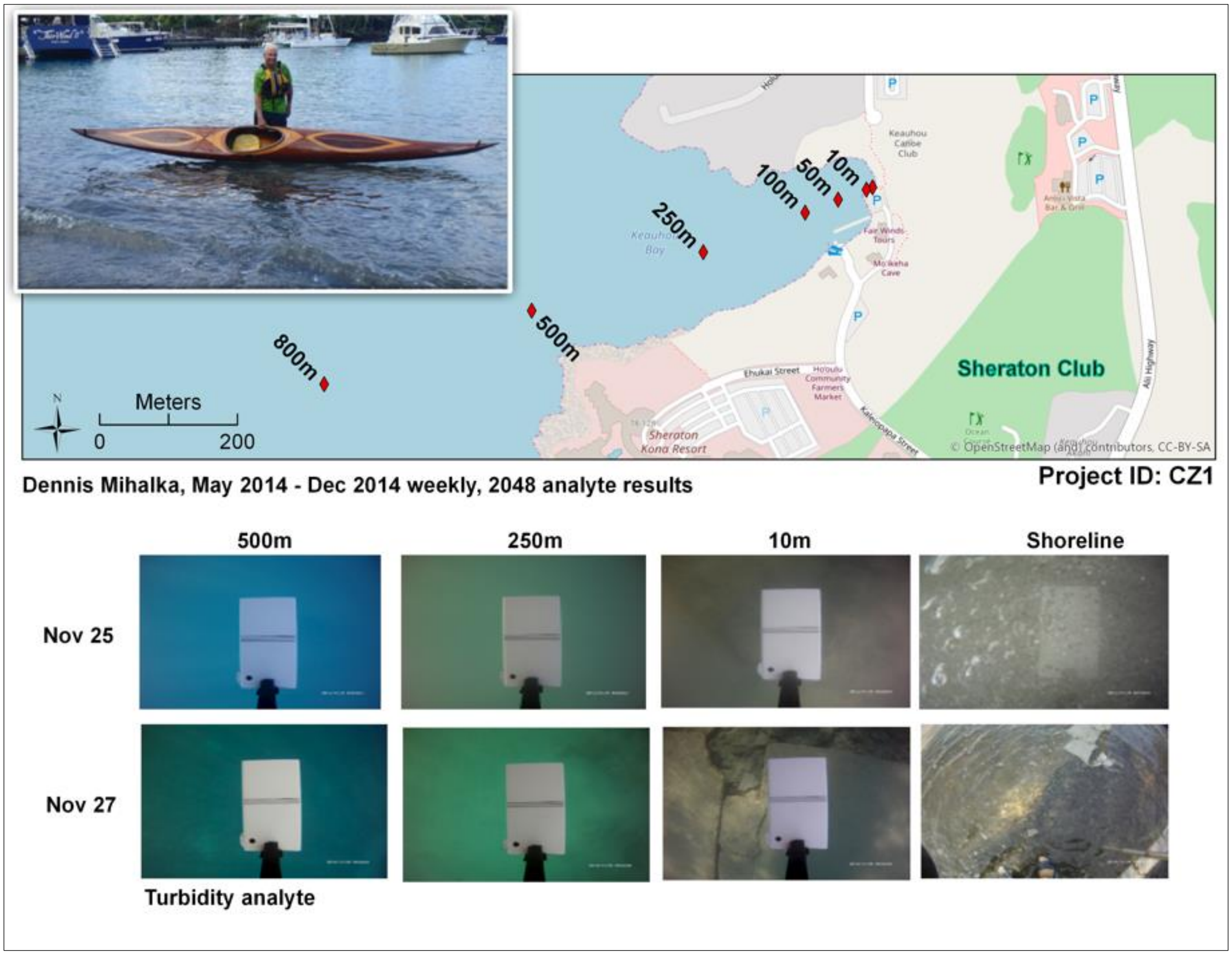


Figure 25. Keauhou Bay, local citizen research inquiry project (CZ1).
Station data source: pers. comm. Dennis Mihalka.

The analytes recorded are salinity, turbidity, temperature (water and air), wind speed and direction, wave height and direction, wind chop height and direction.

3.5. Hawai‘i Ocean Time Series (Control Site)

Dr. Bennett specified that the geodatabase needs a control site. A control is a measurement or a series of measurements that can provide a means to assess samples against the “no manipulation” or “background” condition (Montello and Sutton 2013). Dr. Bennett suggested that, for water quality analysis, the station ALOHA (A Long-Term Oligotrophic Habitat Assessment) located ~100 km north of the Island of O‘ahu can be thought of as a sampling location uninfluenced by anthropogenic manipulation of the environment (see Figure 26). The Hawai'i Ocean Time series (HOT) collected at the ALOHA site, while not specifically designed as a control for comparing with West Hawai'i data, does provide a sparse set of analyte results from the surface (defined as depth range 0-5m) that might be useful for this purpose. A selection of values aggregated by year are shown in the Appendix K.

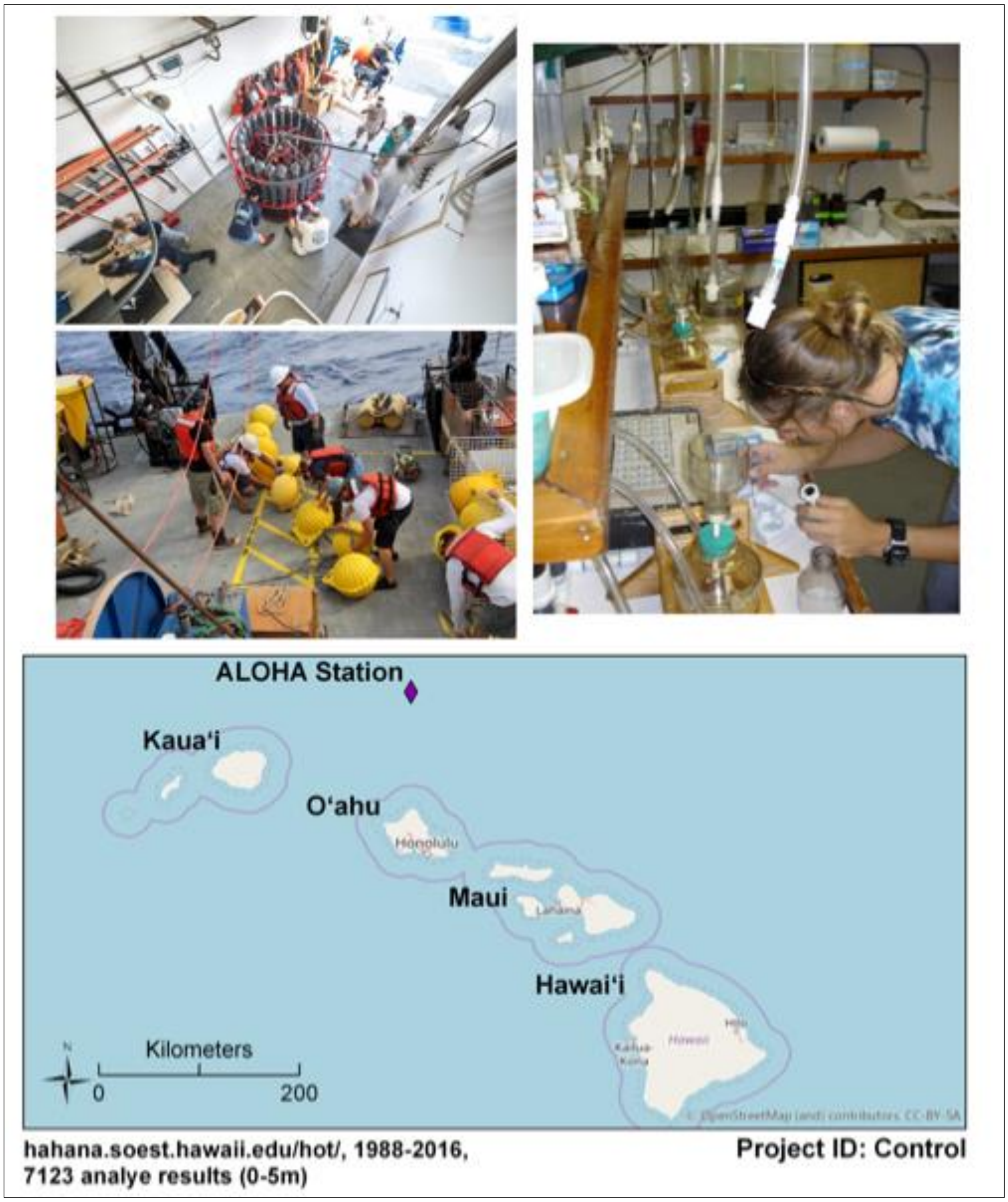


Figure 26. ALOHA Station control site (Control). Data Source: hahana.soest.hawaii.edu/hot/, Pictures: Ryan Tabata (top left), Tara Clemente (bottom left), Dan Sadler (right).

3.6. Summary

As discussed earlier, in addition to providing an integrated set of water quality data in a spatio-temporal framework, the goal of the geodatabase is to provide a GIS foundation that will support the future development of online web maps and map stories that can be used to inform the public about ocean water quality changes. Such GIS data, presented along with images, narratives and analyte results are the means by which scientists involved in this study and others may use a GIS to create new kinds of water quality visualizations. Appendices A to K contain further contextual information and maps related to each of these projects to support this future development. In addition to the final geodatabase, various descriptive information, map layouts and documents relevant to these projects are provided in an open GitHub repository which is described in Chapter 5.

The next step in developing the integrated geodatabase was to customize a professional water science database schema for this study area and populate it with data. The next chapter describes the sequence of methods performed to create the geodatabase design and to combine and integrate data files and tables from each of the projects described above.

Chapter 4 Design and Implementation of the Geodatabase

This chapter explains the design and implementation of the geodatabase in two sections. The first section describes the methods to create a set of SQL tables from the various raw data sources using Microsoft SQL Server. The second section describes how I used ArcGIS Pro to convert the SQL tables into an Esri geodatabase. The geodatabase and methods are the deliverables stored in a GitHub repository which is described in the next chapter.

4.1. Creation of SQL Tables from the Raw Source Data

The steps to make a set of SQL tables from the various raw data sources are shown in Figure 27. The first four of these steps are described in this section. The fifth step is described in the next section.

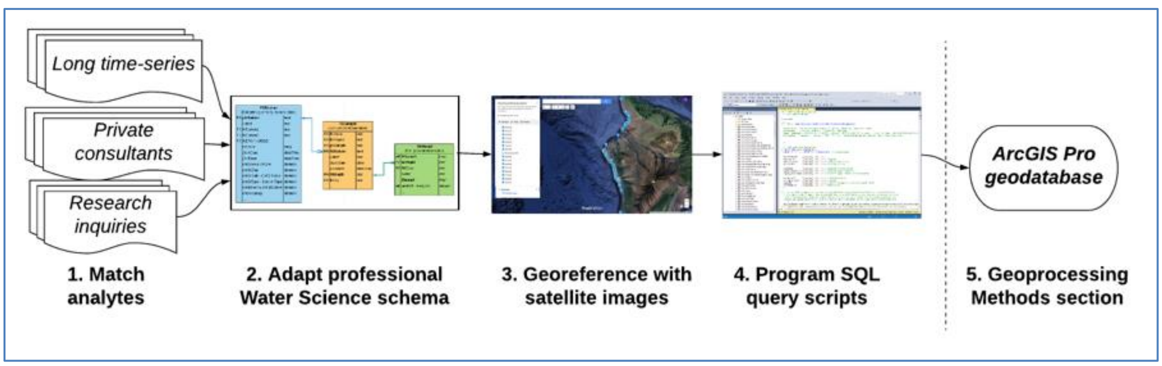


Figure 27. Steps in the design of the relational geodatabase and preparation of data

4.1.1. The Geodatabase Schema

To create the design of the geodatabase for this thesis, I adapted a professional water science database schema called FoxDB from the Illinois State Water Survey (McConkey et al. 2004 and Bartošová et al. 2005) and customized it for the West Hawaiʻi region (see Figure 28). The creators of FoxDB published their data in SQL tables, documented their schema, and provided live data with query instructions so others could easily build on their designs.

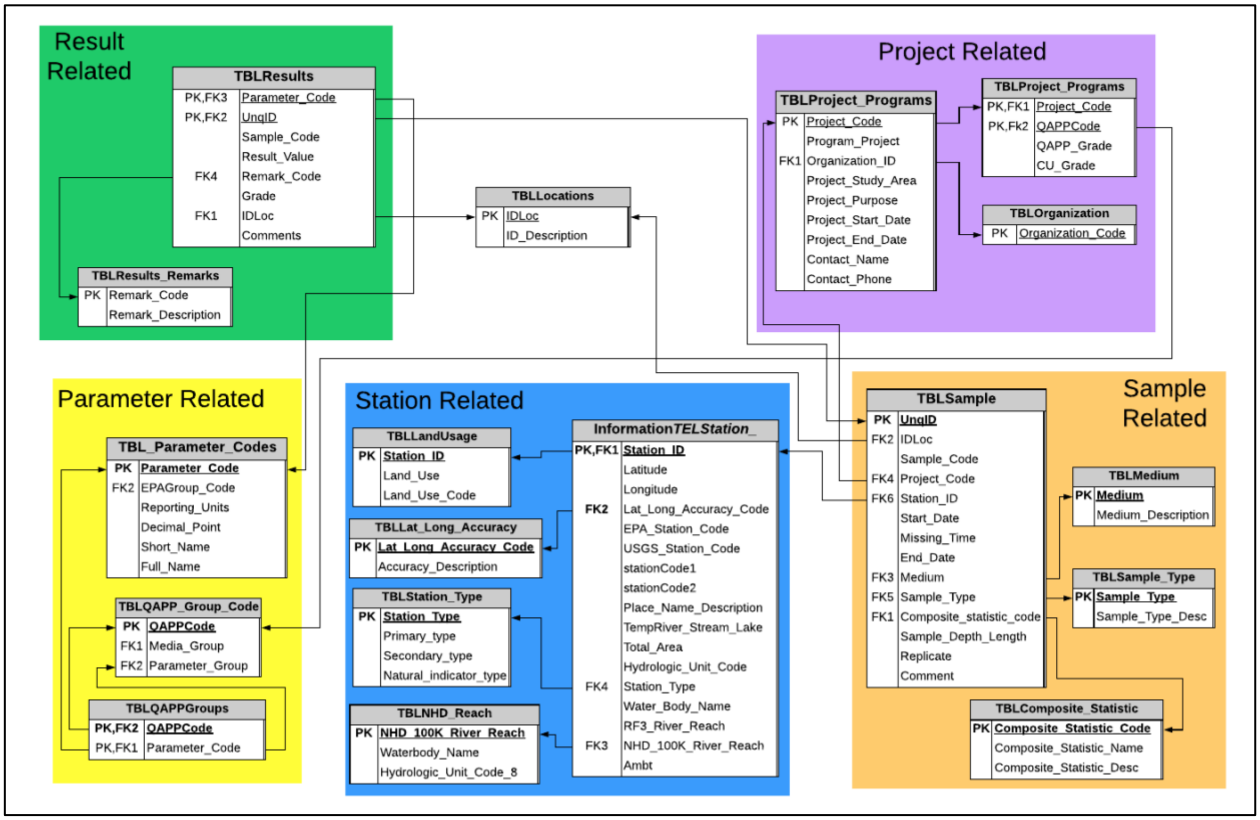


Figure 28. FoxDB schema (partial). A larger version is available in Appendix L. Source: Bartošová et al. 2005.

The FoxDB designers also had to combine data from different kinds of government entities, private parties and citizen scientists. Their projects varied in the analytes tested, geographic areas and time spans like the projects in this study. The scientists needed the combined database to perform different kinds of analysis and to provide input to a software model designed to understand the environment and identify point sources that are polluting the river system.

The key tables in the FoxDB design are conceptualized in the schematic representation in Figure 29 (which I adapted for my study area). The table names and purposes are introduced in Table 3.

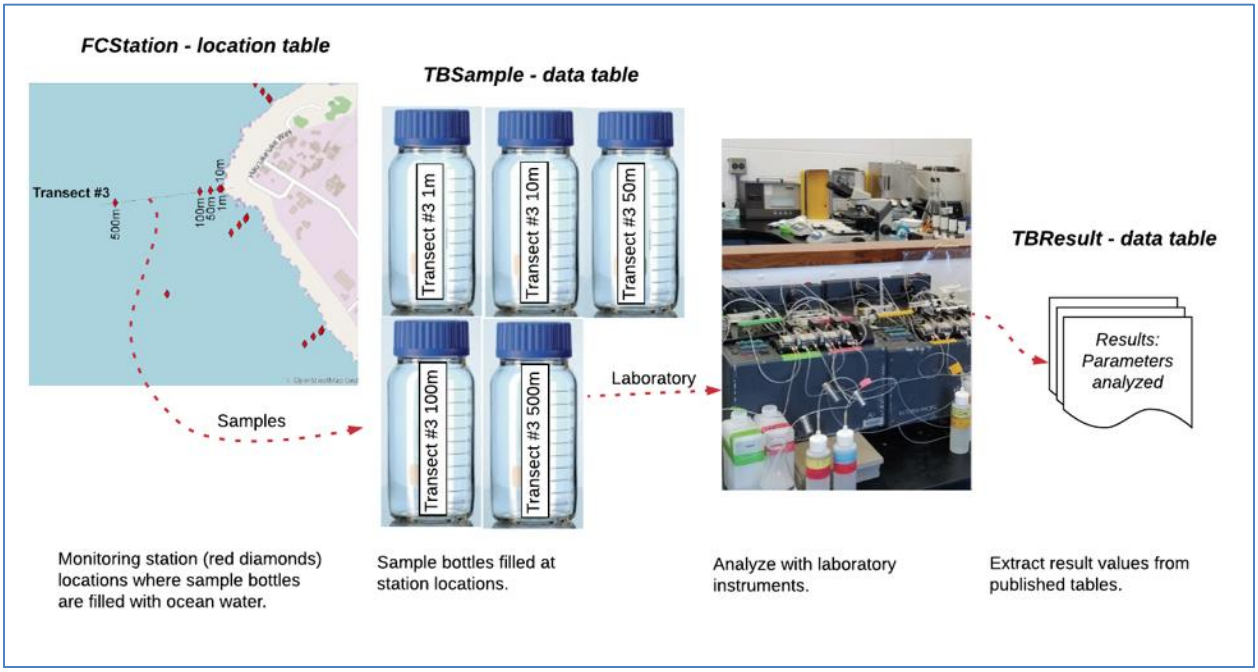


Figure 29. Schematic representation of the key tables in the water quality data model used in the FoxDB (illustrated here by Transect #3 at NELHA). Adapted from Bartošová et al. 2005, Laboratory Picture: Olson 2017.

Table 3. Key FoxDB Water Science Database Tables. Source: McConkey et al. 2004 and Bartošová et al. 2005.

Table	**Purpose**	**Notes**
FCStation	Represents a location in the ocean for use by ArcGIS.	*FC* prefix stands for "feature class" for a location table in ArcGIS terminology.
TBSample	Identifies an ocean water sample recorded at a station for a defined date range.	A sample is visualized as a bottle linked to a station and linked to all the analyzed result values from the instruments.
TBResult	Store the parameters analyzed by the laboratory instruments.	This is the table that contains the "analyte results" as displayed in the maps in Chapter 3.
FCProject (not shown in Figure 29)	Record the quality of the scientist and laboratory process.	Implements the QAPP as discussed in Section 2.1.3. Not developed in this study.

The adapted geodatabase schema used in this thesis is shown along the top of Figure 30 with the original FoxDB tables shown beneath. The FoxDB table and field names have been shortened (e.g. *TELStation_Information* to *FCStation*) and underscores removed. The "FC" prefix stands for feature class and the “TB” prefix stands for table. Domain fields have replaced the small FoxDB support tables with domain field names prefixed with “dm.”

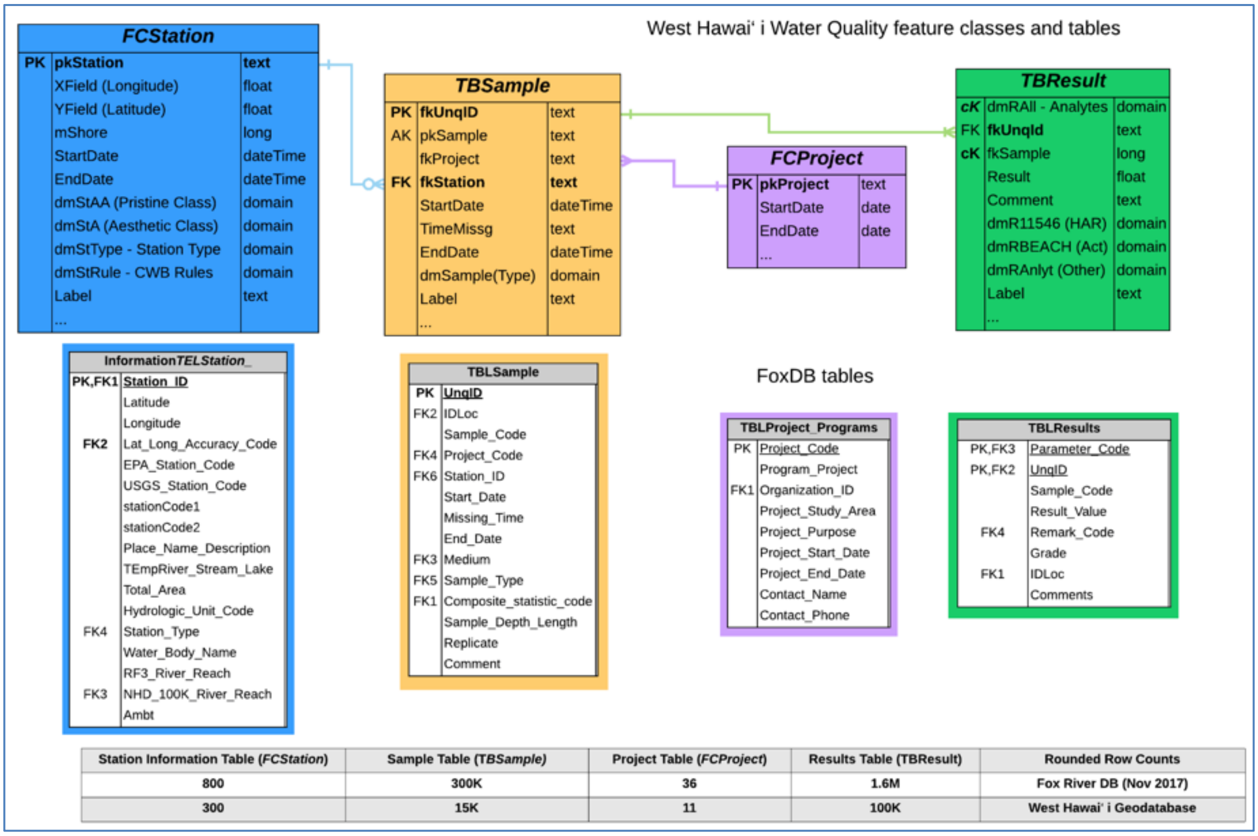


Figure 30. Matching geodatabase with FoxDB table designs. FoxDB tables adapted from Bartošová et al. 2005.

I implemented analyte domain fields to replace the “Parameter Related” section of FoxDB (see yellow box in Figure 28). The parameters refer to the large list of analytes known to the US EPA; this geodatabase uses only a small set of analytes unique to this study area. State Laws, as discussed in Chapter 2, drove the definition of domains (for example, see Section 4.1.3 for domain *dm11546*).

Figure 31 shows the full database design with the most important domains (in grey) next to the table with the domain field. Appendices M and N contain the entire list of fields in the data dictionary with the domains and assigned values.

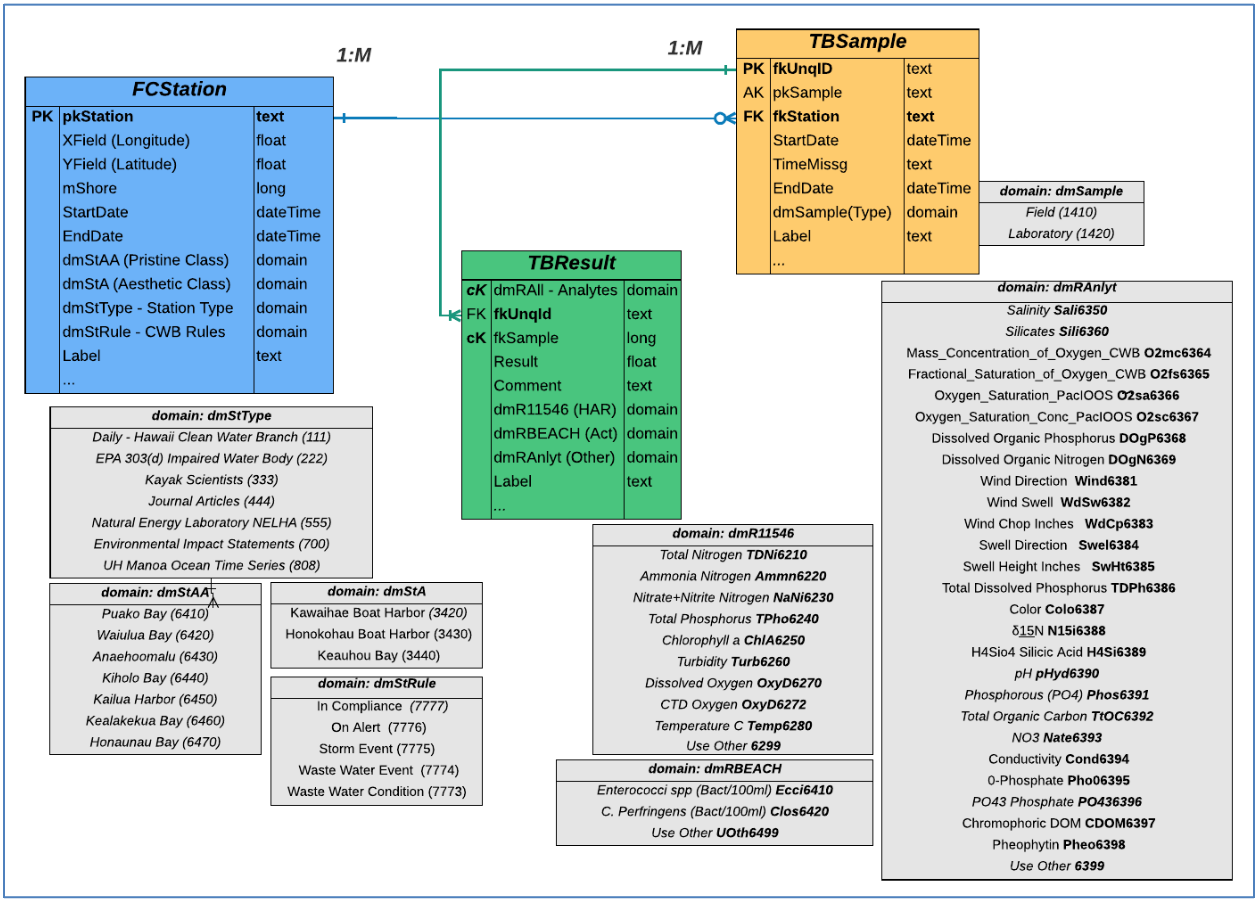


Figure 31. Geodatabase schema

The following tables (Table 4, 5, and 6) show the most important fields of the three tables used in this study (the full list of fields and domains is in Appendices M and N).

Table 4. *FCStation* table description (partial list)

Field Name	Description	SQL Data Type	Nullable
pkStation	Unique identification for stations	nvarchar(255)	No
XField	Longitude in decimal degrees	numeric(38,8)	No
YField	Latitude in decimal degrees	numeric(38,8)	No
mShore	Distance from shore (meters)	int	Yes
StartDate	Earliest date	datetime2(7)	No
EndDate	Ending date	datetime2(7)	No
dmStAA	Class AA domain	int	No
dmStA	Class A domain	int	No
dmStType	Station type	int	No
dmStRule	Rule domain	int	No
Label	For GIS labels	nvarchar(255)	No

Table 5 *TBSample* table description (partial list)

Field Name	**Description**	**SQL Data Type**	**Nullable**
pkSample	Unique sample identification number, number assigned if not provided by data originator. A sample is a monitoring activity (e.g., ambient samples, measurements, and observations) performed at a specific date, time, and location in order to characterize the environment	nvarchar(255)	No
fkStation	Unique number for the station, identifying the specific location, at which field work/sampling is conducted	nvarchar(255)	No
fkUnqID	Unique sample identification code for linking to *TBResult*	nvarchar(255)	No
StartDate	Date and time when field work/sample collection began	datetime2(7)	No
EndDate	Date and time when field work/sample collection ended	datetime2(7)	No
TimeMissg	Indicates whether sampling time was missing in the original data	nvarchar(255)	Yes
dmSample	Letter code describing the sampling method	int	No
Label	For GIS labels	nvarchar(255)	No

Table 6. *TBResult* table description (partial list)

Field Name	Description	SQL Data Type	Nullable
fkSample	Unique sample identification number. A sample is a monitoring activity performed at a specific date, time, and location in order to characterize the environment	nvarchar(255)	No
fkUnqID	Unique sample identification code for linking to *TBSample*	nvarchar(255)	No
Result	Data value for a sample result or a code representing an observation. Result values can be numeric or alphanumeric values	numeric(38,8)	No
Comment		nvarchar(255)	Yes
dmRAll	Analyte Domain containing all analytes	int	No
dmR11546	Analyte Domain for State Law	int	No
dmRBEACH	Analyte Domain for BEACH Act	int	No
dmRAnlyt	Analyte Domain for other than dmR11546 and dmRBEACH	int	No
Label	Analyte code, 4 letters+4 numbers	nvarchar(255)	No

4.1.2. Georeferencing Sampling Stations

The station location table needs the latitude and longitude of each sampling station. All stations were georeferenced using the Google MyMaps GIS. Using maps and other information provided in each project's documentation, I visually identified the location of each station or transect start point on the Google image. Even though some projects provided the XY coordinates, I georeferenced each station in this way for consistency. For projects that supplied marked up maps, I lined up the crevices, bays, houses, or other unique features in the map with the Google imagery and then drew transects and measured distances as indicated by the scientists. For the HDOH locations, they provided hand drawn maps with narrative explaining how to locate the station relative to nearby features. For finding the stations along transects, MyMaps has tools that allow a user to place pins in the ocean along transects at the distances specified by the scientists.

Once I had created pins in MyMaps for each station, I extracted and entered the latitude and longitude provided by MyMaps into the *POINT_Y* and *POINT_X* fields of the *FCStation* table. The geographic coordinate system of the geodatabase is WGS84 since that is what Google MyMaps uses.

4.1.3. Analyte Matching and Coding

The next step was to match the analytes across all projects by assigning common codes for use in the *TBResult* table. Table 7 shows the set of required analytes specified in the state law. The analytes are coded for the geodatabase using both a unique string and integer. I assigned the four number integer codes so they are unique to an analyte but the actual value is not important as they are just a number I picked for easy recognition. These are nominal data types with a limited set of distinct values but without meaning otherwise. The string code is constructed by prepending four characters to the four numbers of the nominal integer code. The four characters are also nominal data types but they were chosen to generally match the codes assigned by scientists in the heading rows of their data tables. In addition to the eight analytes mentioned in state law, many scientists, especially those who have undertaken research inquiries, add additional analyte types based on project specific goals and improved technologies, so in the final database there are over thirty analytes across all the included projects. All are coded in a similar manner to that described above.

Table 7. Analytes in State Law with assigned codes

Analyte Name	Integer Code	String Code
Total nitrogen	6210	TDNi6210
Ammonia nitrogen	6220	AmmN6220
Nitrate+nitrite nitrogen	6230	NaNi6230
Total phosphorus	6240	TPho6240
Chlorophyll a	6250	ChlA6250
Turbidity	6260	Turb6260
Dissolved oxygen	6270	OxyD6270
Temperature	6280	Temp6280

ArcGIS organizes and restricts attribute values through the use of domains in a geodatabase. When an attribute is associated with a domain, the software automatically limits the possible values to those specified in the domain. The domains for analyte results are *dmAll* (all analytes), *dm11546* (required by state law), *dmBEACH* (required by US BEACH Act), and *dmRAnlyt* (not in *dm11546* or *dmBEACH*).

4.1.4. SQL Queries and Data Extraction from Sources

Each project has a unique source data file format and different temporal properties; this means that extracting the data from the original source files into a set of data tables that could be combined into one integrated geodatabase required a unique set of queries for each project. These query sequences are recorded in script files where a script is a single text file that has a ".sql" extension. However, the query sequence script flow for each project is similar. The steps in this flow are outlined in Figure 32 using the NELHA project's SQL query file, ProjectID_LTS1_KeaholeNELHA.sql, as an example. An annotated and shortened copy of this script file is shown in Appendix A. The final geodatabase results from the concatenation of the SQL tables.

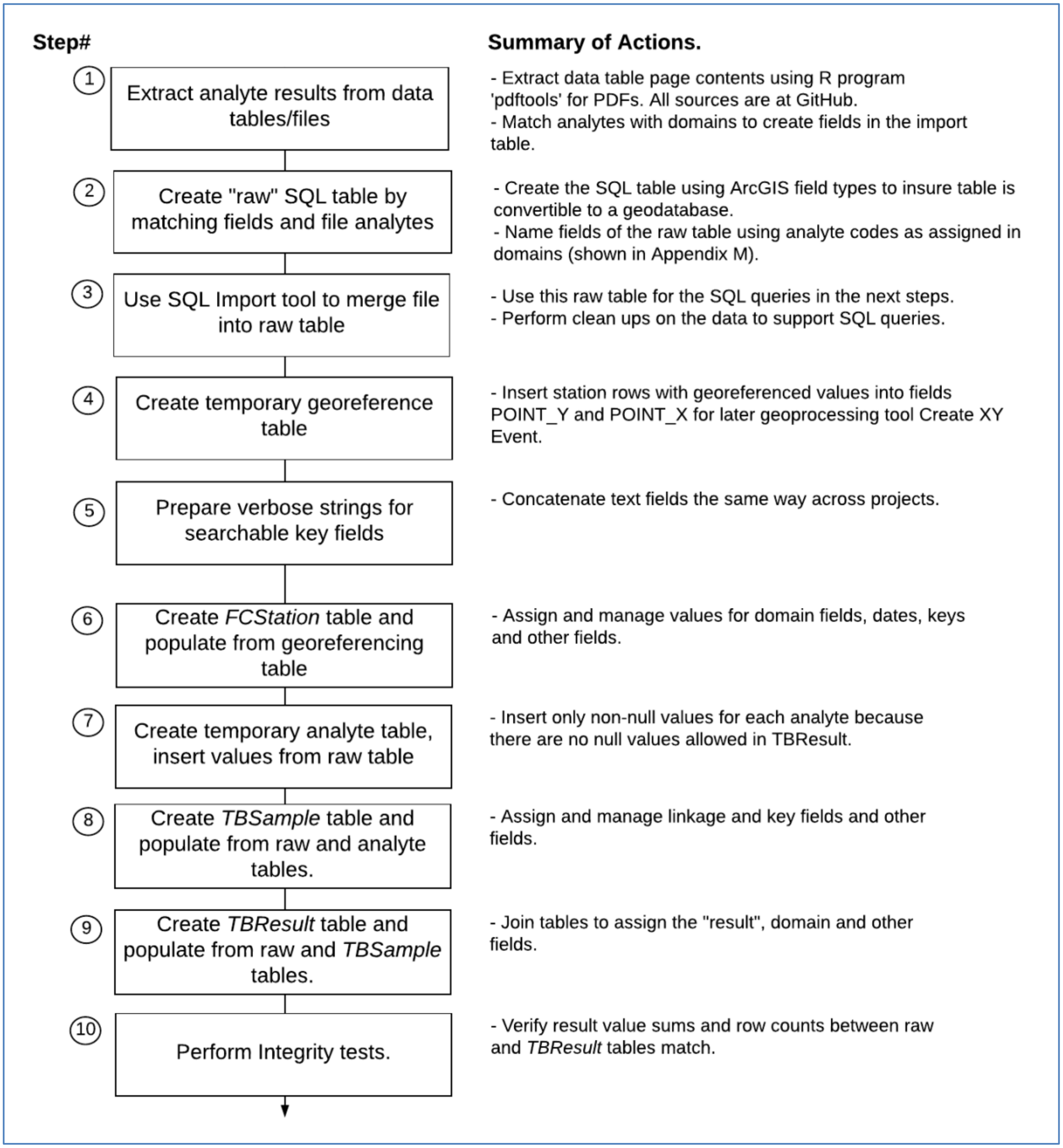


Figure 32. SQL Query Flow. Step numbers refer to markers in the script source code in Appendix A and file ProjectID_LTS1_KeaholeNELHA.sql at GitHub.

Steps #1-#3 are preparation steps to get the analyte results from a file into an SQL "Raw" table allowing manipulation with SQL queries. Step #4 creates a temporary table for inserting the values from the georeferencing step discussed in a previous section. Step #6 implements the creation of *FCStation* by copying from the temporary georeferencing table. Step #7 and #8 prepare for the creation of the table *TBSample* and *TBResult* by using a temporary table to pull analyte values from the Raw table. Step #9 creates and populates the *TBResult* table by pulling analytes from the temporary and Raw table.

For each project, the final output from this flow is a set of three tables: *FCStation*, *TBSample* and *TBResult* which represent the SQL version of the geodatabase. The next section describes the sequence of geoprocessing tools required to move these tables into ArcGIS and then to a file geodatabase. The next chapter explains where to find these files in this project's GitHub repository.

4.2. Geoprocessing Methods

Once the set of SQL tables was populated from the data tables of each project, it was possible to proceed to the final phase in which everything was merged into a single geodatabase. Since the ArcGIS interface to Microsoft SQL is for accessing stored databases only, this section outlines the sequence of ArcGIS geoprocessing methods necessary to convert the SQL tables into ArcGIS's native relational database format (the geodatabase). It is from this geodatabase that the maps included in Chapter 3 and the Appendices were produced. This section ends with a brief description of the ArcGIS methodology needed to traverse the links between the three geodatabase tables in order to perform a query to extract a specific set of analyte results.

4.2.1. Geodatabase Preparation

Three geoprocessing tools must be used to convert SQL tables to an Esri ArcGIS file geodatabase. ArcGIS Pro supports these SQL Server connections through an enterprise geodatabase license. Figure 33 details the three steps required. The first step is a simple table-to-table conversion (Table-To-Table), the second is converting the latitude and longitude fields in

FCStation to point geometry within the WGS84 coordinate system and the final step is writing the data in the file geodatabase. At this point, the geodatabase is ready to use in ArcGIS.

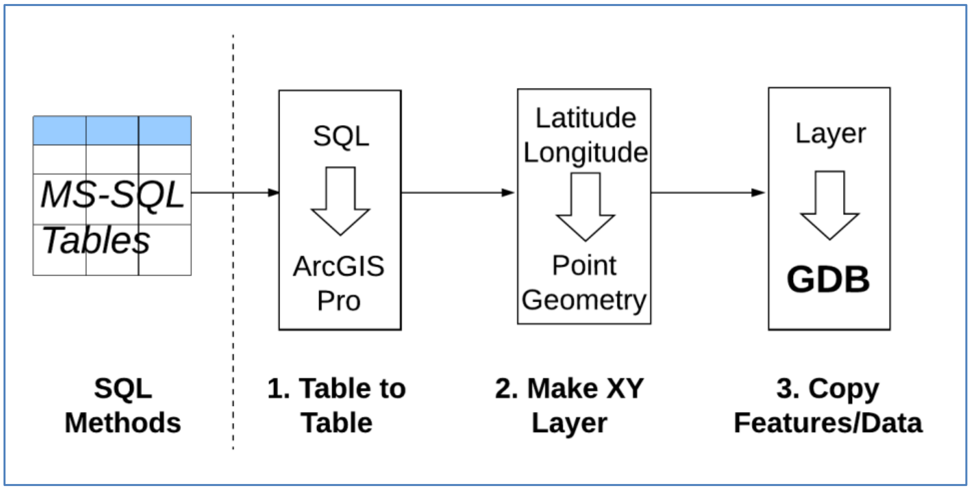


Figure 33. SQL tables to Esri ArcGIS processing steps

4.2.2. Esri Relate Use Case

To extract analyte results, an ArcGIS Pro user must execute a sequence “Relate” geoprocessing commands. Figure 34 shows both the Esri methods and the equivalent SQL commands necessary to extract analyte results from the geodatabase tables. At the bottom of the figure are the rounded counts of the number of rows for the combined data: *FCStation*=300, *TBSample*=15K and *TBResult*=100K. The maps of projects in this document show the points of the feature class *FCStation* which has links into matching rows of *TBSample* which then links to all the analyte result rows in *TBResult* for that sample.

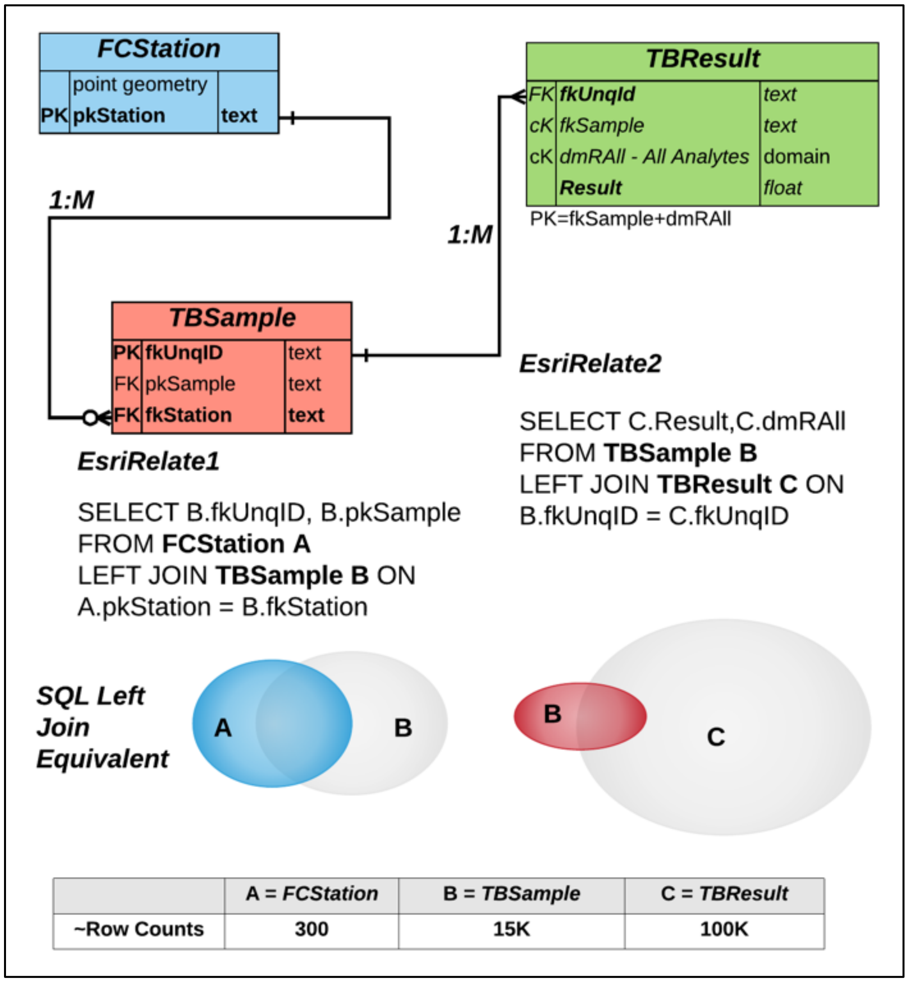


Figure 34. Relate feature class *FCStation* and tables *TBSample* and *TBResult*

The geodatabase is now ready to be used in ArcGIS and shared. The next chapter describes the final deliverables and its provision into GitHub for public distribution.

Chapter 5 Results: The GitHub Repository

The final results of the work completed in this thesis are a published geodatabase, a set of documented reproducible methods, a large collection of documents and other source materials, and ArcGIS maps and layouts that can be used to share and explore the data and to create multimedia presentations. All data were acquired from public sources (except the citizen scientist who gave me permission to share) and are returned to the public domain through a public repository at GitHub.

This geodatabase constructed from the data produced by each of the projects discussed in Chapter 3 is static. Future updates to the geodatabase should use standard GitHub processes (i.e. clone and fork) to ensure consistency and reproducibility.

5.1. Using GitHub

GitHub is a free online storage and retrieval repository which now holds the geodatabase, methods and ArcGIS Pro maps and layouts for this project. The GitHub account is WestHawaiiWaterQuality and its contents are summarized in Table 8.

Table 8. GitHub Public Repository Contents

Purpose	Repository Name	Contents
Geodatabase	GDB_Files	GDB and XLS files
Methods	SQL_Methods	SQL script for each project
ArcGIS Pro	Map_Layout_Files	Esri map and layout files
Data	WQData	Source data for each project
Library	WQLibrary	Dr. Bennett's research library

The next two figures show the steps necessary to view or download the repository data files. Use a browser to navigate to the Overview page of the *WestHawaiiWaterQuality* account to view all the repositories. Figure 35 shows the view of the Overview page which lists the repositories.

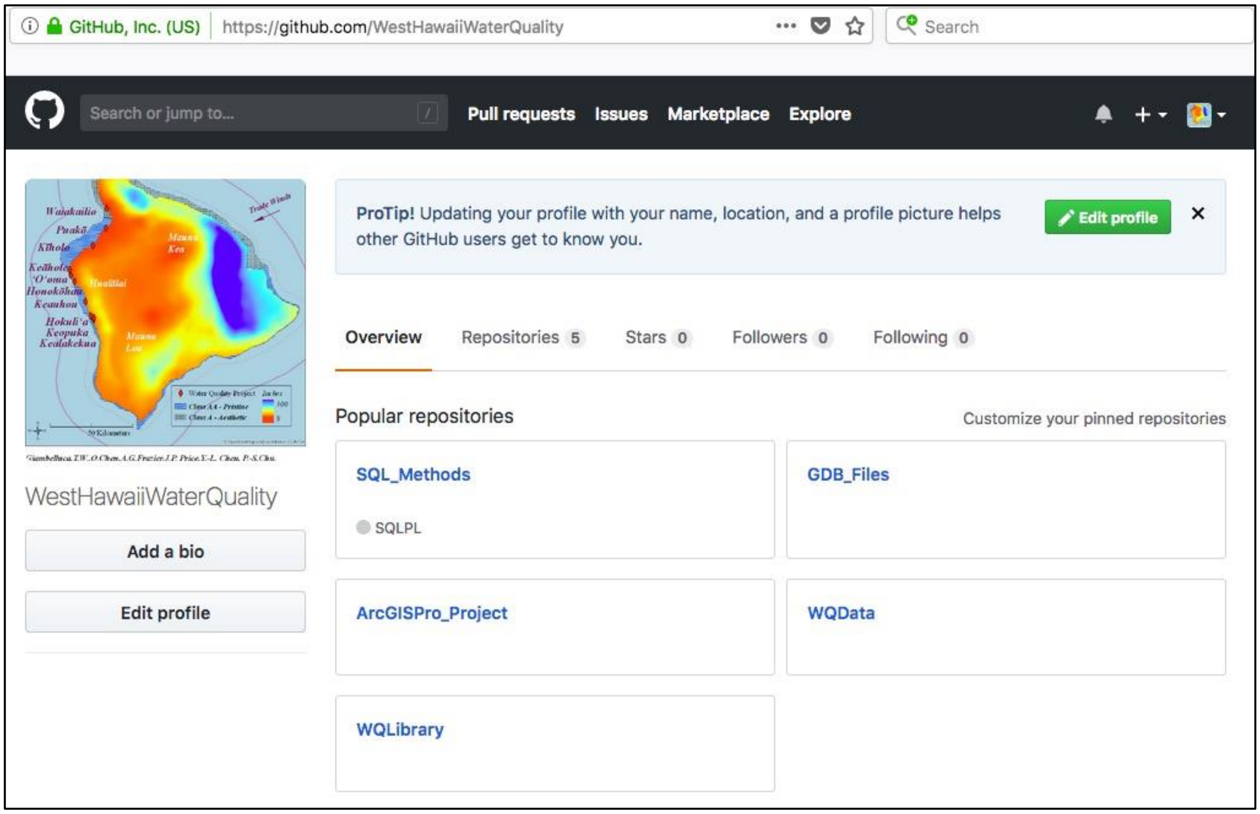


Figure 35. GitHub WestHawaiiWaterQuality Repositories

To view the contents of a repository, click on the relevant name. For example, choose the 'SQL_Methods' to the display query script files for each of the projects, as shown in Figure 36. To download or view the files in a repository, GitHub provides a large green button visible in the figure. The README file contains more details and versioning information related to the files in the repository. This is normally the starting point in any GitHub repository provided to enable users to understand the latest information about the files.

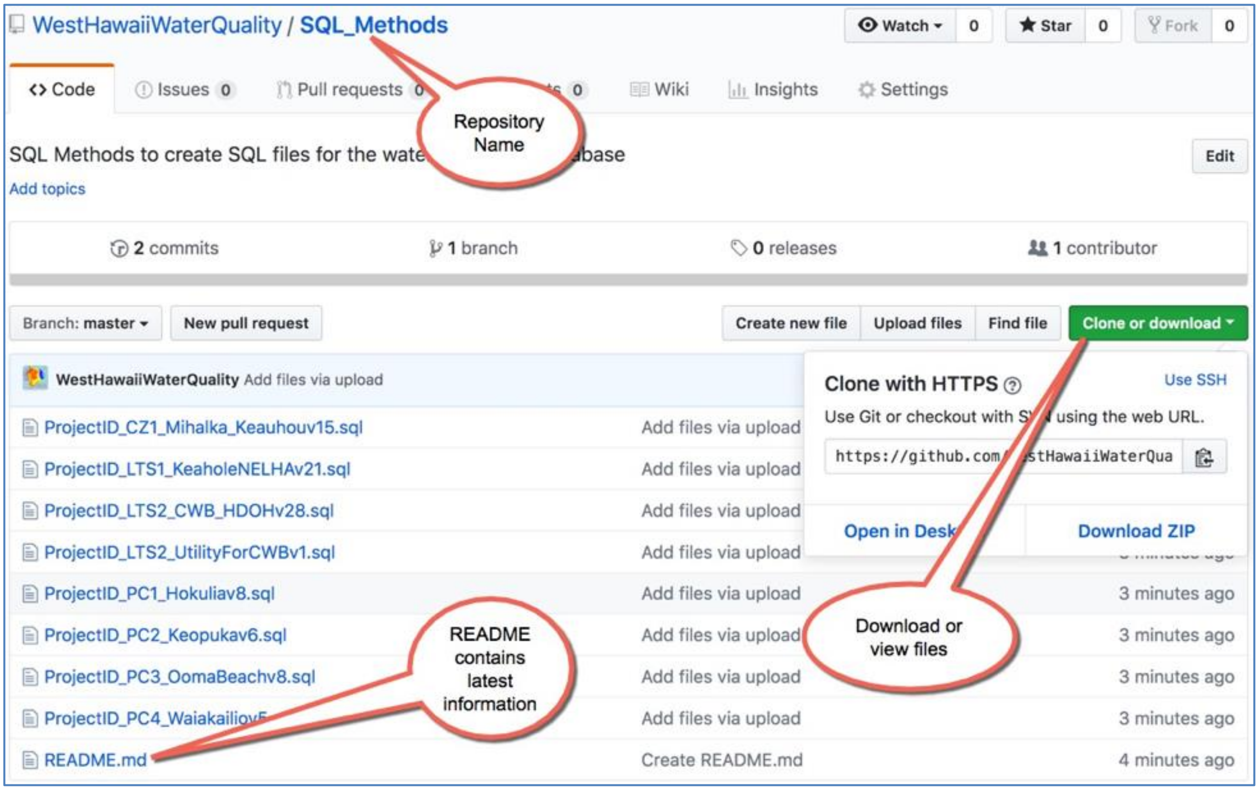
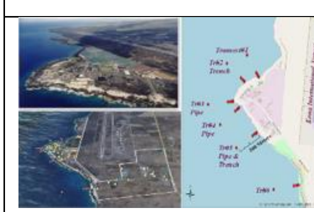


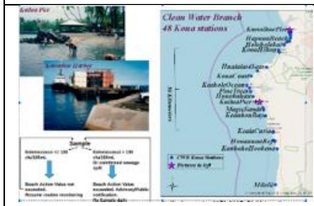
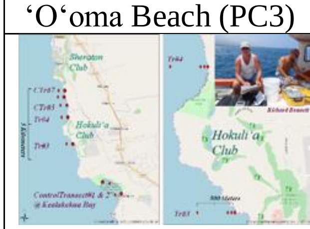
Figure 36. Downloading SQL_Methods from GitHub

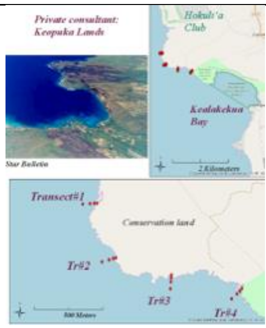
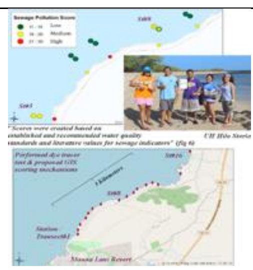
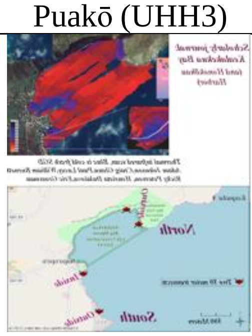
5.2. Repository Details

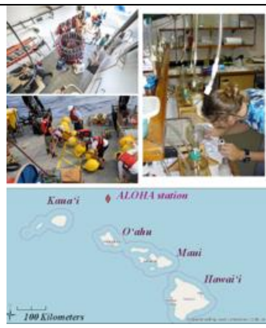
The GitHub repository contains all the source data, the resulting geodatabase, associated documents, and a large collection of maps and map layouts produced from the geodatabase. Table 9 lists the key components of this repository as it relates to each project included.

Table 9. Project Row Counts with GitHub Repository Names

Project	**GDB_Files Row Counts**	**SQL_Method Scripts**	**ArcGISPro_Project and WQData**
 Keāhole Point (LTS1)	FCStation=60 TBSample=2,940 TBResults=29,698	ProjectID_LTS1_KeaholeTransect[1-6]_ImportXLS.sql ProjectID_LTS1_KeaholeNELHA.sql	**Map:** KeaholeNELHA **Layout:** Ch3 ProjectIDLTS1_Keahole **WQData:** 19_7287_LTS1_Keahole_KOlson

Project	GDB_Files Row Counts	SQL_Method Scripts	ArcGISPro_Project and WQData
 HDOH Clean Water Branch (LTS2)	FCStation=48 TBSample=10,655 TBResults=50,014	ProjectID_LTS2_CWB_HDOHsql	**Map:** CWB **Layout:** Ch3 ProjectIDLTS2_CWB7 **WQData:** CleanWaterBranch*
 Waiakailio Bay (PC4)	FCStation=30 TBSample=30 TBResults=419	ProjectID_PC4_Waiakailio.sql	**Map:** Waiakailio **Layout:** Ch3 ProjectIDPC4_Waiakailio **WQData:** 20_0747_PC4_Waiakailo_SDollar
 ‘O‘oma Beach (PC3)	FCStation=69 TBSample=69 TBResults=940	ProjectID_PC3_OomaBeach.sql	**Map:** Ooma Beach **Layout:** Ch3 ProjectIDPC3_Ooma **WQData:** 19_7084_PC3_Ooma_SDollar
 Hokuli‘a (PC1)	FCStation=35 TBSample=35 TBResults=380	ProjectID_PC1_Hokulia.sql	**Map:** Hokulia **Layout:** Ch3 ProjectIDPC1_Hokulia **WQData:** 19_5139_PC1_Hokulia_DZiemann

Project	GDB_Files Row Counts	SQL_Method Scripts	ArcGISPro_Project and WQData
 Keopuka Lands (PC2)	FCStation=24 TBSample=24 TBResults=316	ProjectID_PC2_Keopuka.sql	**Map:** Keopuka **Layout:** Ch3 ProjectIDPC2_Keopuka **WQData:** 19_4872_PC2_Keopuka_RBrock
 Kīholo Bay Buoy (UHH1)	FCStation=1 TBSample=588 TBResults=5831	ProjectID_UHH1_KiholoBuoy.sql	**Map:** KiholoBuoy **Layout:** Ch3 ProjectIDUHH1_KiholoBuoy **WQData:** 19_8631_UHH1_KiholoBay_JAdolf
 Puakō (UHH3)	FCStation=16 TBSample=16 TBResults=160	ProjectID_UHH3_Abaya_Puako.sql	**Map:** Puako **Layout:** Ch3 ProjectIDUHH3_Puako **WQData:** 19_9831_UHH3_Puako_LAbaya
 Honokōhau Harbor and Kealakekua Bay (UHH2)	FCStation=8 TBSample=8 TBResults=80	ProjectID_UHH2_Parsons.sql	**Maps:** ParsonsHono, ParsonsKela **Layouts:** Ch3 ProjectIDUHH2h_ParsonsHono.sql Ch3 ProjectIDUHH2k_ParsonsKeala **WQData:** 19_4824_UHH2_KealaHono_MParsons

Project	GDB_Files Row Counts	SQL_Method Scripts	ArcGISPro_Project and WQData
 Keauhou Bay (CZ1)	FCStation=7 TBSample=280 TBResults=2048	ProjectID_CZ1_Mihalka_Keauhou.sql	**Map:** Keauhou Bay **Layout:** Ch3 ProjectIDCZ1_Keauhou Bay **WQData:** 19_5620_UHH2_Keauhou_DMihalka
 Hawai‘i Ocean Time series (Control)	FCStation=1 TBSample=1 TBResults=7123	ProjectID_Control_HOTUHManoa.sql	**Map:** HOTS **Layout:** Ch3 ProjectIDControl_HOT **WQData:** 22_7500_Control_HOTs_ALOHA

Chapter 6 Conclusion

This thesis has met all its original goals. As a result of the work completed in this thesis, the local environmental scientists who motivated this work can now: 1) share and conduct research with a combined set of data on West Hawaii's coastal water quality packaged into a single integrated geodatabase, 2) refer to the reproducible methods at GitHub so that others can continue to build and maintain the repository while meeting the external submission requirements of the state, and 3) use all of the materials provided in the repository as a foundation for generating content for a Web GIS that will be able to share this information effectively with the public.

Consultations with the group of environmental scientists confirmed that the output products have potential (and immediate) value. For example, Dr. Bennett is already using the data visualizations and site maps in community and government presentations and, along with Dr. Mihalka, plans to create Esri story maps to produce content for local community television. Mr. Olson intends to compile data from the geodatabase for government presentations. Additionally, Dr. Tracy Wiegner from UHH Marine Science Laboratory has expressed interest in the data for use in her ongoing research.

Future researchers can use these deliverables to design new surveys and as a baseline for analysis. For example, the FoxDB database (from which this geodatabase was derived) is a source of input to water quality models, pollutant screening tools, and simulations that are used to better understand the environment.

6.1. Lessons Learned

Some of lessons learned include data handling for reproducibility and handling georeferenced data. While not mentioned in the workflow above, pdftools, the R package by

Jeroen Ooms (2018), was invaluable to minimize the transcription time and minimize errors for text extraction from PDF documents. See Appendix O for the use of the program in the ‘O‘oma Beach project. I performed minimal editing of the extracted text files, just enough for the SQL Import tool to recognize the rows and columns properly. For georeferenced data, I should have used the KML output of Google MyMaps into ArcGIS Pro then into SQL tables instead of hand entering all latitude/longitude values into SQL scripts.

Another lesson learned was that it would have been good to design the geodatabase so that it could incorporate other factors that are important in the study area, ones that may not have been important in the Fox River watershed. For this study region, given that submarine ground discharge is so important, I should have incorporated the SGD flow rate data that has already been collected (e.g. by Knee et al. 2010) into the schema design.

6.2. Future Opportunities

Given the time constraint of this thesis project, I did not get to write a story map nor work with Dr. Bennett to develop ways of visualizing the analysis of the data. Instead, I have provided detailed descriptions of the transects, stations, time series, methods of collection, etc., for each project and this information can be transformed into a format more easily understood when presented to the public.

The inherent challenge of how to use a GIS to visualize water quality data is yet to be addressed and offers great potential for the next stage of work with this geodatabase. The maps developed for this thesis are limited to the sample locations along transects and their spatial context at regional and local scales. Further development with regards to the visualizations of the integrated water quality data contained within this database will offer stakeholders and scientists a means to gain in-depth and comprehensive understanding about water quality trends in the region.

References

Abaya, Leilani M. 2016. "Identifying Hotspots of Sewage Pollution in Coastal Areas with Coral Reefs." Master's thesis. University of Hawaiʻi at Hilo.

Bartošová, Alena, Ramanitharan Kandiah, and Vladimir Novotny. 2005. *Development of a relational database for studying ecological response of streams to anthropogenic watershed stresses and stream modifications*. Northeastern University, Boston, MA: Center for Urban Environmental Studies.

Brock, Richard. 2000. *Status of Nearshore Marine Communities and Coastal Water Quality Front the Keopuka Lands Parcel, South Kona, Hawaii Preconstruction Baseline Report*. By Pacific Star LLC. In the Office of Environmental Quality Control EA and EIS Library. Accessed December 1, 2017. http://oeqc2.doh.hawaii.gov/PDR3 2000-07-23-HA-DEIS-KEOPUKA- Water qual appen C -VOL-2.pdf.

Dollar, S.J. 2008. *Assessment Of Marine Water Chemistry 'O'oma Beachside Village North Kona*. By 'O'oma Beachside Village, LLC. In the Office of Environmental Quality Control EA and EIS Library. Accessed December 1, 2017. http://oeqc2/doh/hawaii.gov/PDR2/Appendix/Ooma Marine Chemistry.pdf.

Dollar, S.J. 2010. *Assessment Of Marine Water Chemistry and Marine Biotic Communities*. By Kohala Shoreline, LLC. In the Office of Environmental Quality Control EA and EIS Library. Accessed December 1, 2017. http://oeqc2.doh.hawaii.gov/EA_EIS_Library/2015-07-08-HA-DEA-Kohala-Shoreline-LLC-Project.pdf.

Dollar, S. J. and M. J. Atkinson. 1992. "Effects of Nutrient Subsidies from Groundwater to Nearshore Marine Ecosystems Off the Island of Hawaii." *Estuarine, Coastal and Shelf Science* 35 (4): 409-424. doi:10.1016/S0272-7714(05)80036-8.

Giambelluca, T. W., Q. Chen, A. G. Frazier, J. P. Price, Y. -L Chen, P. -S Chu, J. K. Eischeid, and D. M. Delparte. 2013. "Online Rainfall Atlas of Hawaiʻi " *Bulletin of the American Meteorological Society* (94): 313-316. doi:10.1175/BAMS-D-11-00228.1.

Hawaiʻi Department of Health (HDOH). 1997. *Rationale of the Development of Area-Specific Water Quality Criteria for the West Coast of the Island of Hawai'i and Procedures for their Use*. By Environmental Planning Office. Honolulu.

Hawaiʻi Department of Health (HDOH). 2012a. *Beach Monitoring Quality Assurance Project Plan CWBMONQAPP002*, Revision 0, May 07, 2012. By Environmental Management Division, Clean Water Branch, Monitoring and Analysis Section. Honolulu. https://www.pacioos.hawaii.edu/wp-content/uploads/2016/08/Beach_Monitoring_QAPP_CWBMONQAPP002_120507.pdf.

Hawai‘i Department of Health (HDOH). 2012b. *2012 State of Hawaii Water Quality Monitoring and Assessment Report: Integrated Report to the U.S. Environmental Protection Agency and the U.S. Congress Pursuant to Sections 303(D) and 305(B), Clean Water Act (P.L.97-117);2014 SRI S2065-3*. Honolulu. http://health.hawaii.gov/cwb/files/2013/04/IntegragedReport.pdf.

Hawai‘i Department of Health (HDOH). 2014. *Amendment and Compilation of Chapter 11-54 Hawaii Administrative Rules, Water Quality Standards*. Honolulu. https://health.hawaii.gov/cwb/files/2013/04/Clean_Water_Branch_HAR_11-54_20141115.pdf.

Hawai‘i Department of Health (HDOH). 2017a. *Hawaii Beach Monitoring Program*. By Clean Water Branch. Honolulu. https://health.hawaii.gov/cwb/files/2017/05/Hawaii-Beach-Monitoring-Program-170508.pdf.

Hawai‘i Department of Health (HDOH). 2017b. *Data Submittal Requirement For External Entities*. By Clean Water Branch. Honolulu. https://health.hawaii.gov/cwb/files/2017/06/Data-Submittal-Requirements-for-External-Entities.pdf.

Hunt, Chip. 2018. *HIDOH / USGS Cooperative Program, Wastewater and Nutrient Source Tracking – Results of Reconnaissance Chemical Mapping at Kualoa and Kahana, Oahu*. USGS Pacific Islands Water Science Center. Accessed on April 26, 2018. https://hi.water.usgs.gov/studies/beachmonitoring/pdf/USGS_DOH_Sourcetracking_slides.pdf.

Johnson, A., C. Glenn, P. Lucey, W. Burnett, R. Peterson, H. Dulaiova, E. Grossman. 2015. *Thermal Infrared Surveys and Nutrients Reveal Substantial Submarine Groundwater Discharge Systems Emanating from the Kona Coast of Hawaii*. Honolulu, HI: University of Hawai‘i. Accessed April 28, 2018. www.soest.hawaii.edu/GG/FACULTY/glenn/Glen_Infrared_POSTER_11x17.pdf.

Knee, Karen L., Joseph H. Street, Eric E. Grossman, Alexandria B. Boehm, and Adina Paytan. 2010. "Nutrient Inputs to the Coastal Ocean from Submarine Groundwater Discharge in a Groundwater-dominated System: Relation to Land use (Kona Coast, Hawaii, U.S.A.)." *Limnology and Oceanography* 55 (3): 1105-1122. doi:10.4319/lo.2010.55.3.1105.

Lau, L. Stephen and John F. Mink. 2006. *Hydrology of the Hawaiian Islands*. Honolulu: University of Hawai‘i Press.

Laws, Edward A. 1993. *Aquatic Pollution*. 2nd ed. New York: Wiley.

McConkey, S., A. Bartošová, L. Lian-Shin, K. Andrew, M. Machesky, and C. Jennings. 2004. *Fox River Watershed Investigation – Stratton Dam to the Illinois River: Water Quality Issues and Data Report to the Fox River Study Group, Inc*. Champaign, IL: Illinois State

Water Survey. Accessed April 28, 2018. http://ilrdss.isws.illinois.edu/fox/fox_report_phase1.asp

Montello, Daniel R. and Paul Sutton. 2012. *An Introduction to Scientific Research Methods in Geography and Environmental Studies*. 2nd ed. Los Angeles: Sage Publications Ltd. http://arrow.unisa.edu.au:8081/1959.8/124508.

Olson, Keith. 2017. *Annual Report for The Comprehensive Environmental Monitoring Program Covering the Period: July 24, 1982 through June 30, 2016*. Kailua-Kona, HI: Natural Energy Laboratory of Hawaiʻi Authority.

Ooms, Jeroen. 2018. *Package 'pdftools'*. Accessed August 22, 2018. https://cran.r-project.org/web/packages/pdftools/pdftools.pdf.

Parsons, Michael L., Walsh William J., Settlemier, Chelsie J. et al. 2008. "A Multivariate Assessment Of The Coral Ecosystem Health Of Two Embayments On The Lee Of The Island Of Hawaiʻi." *Marine Pollution Bulletin* 56, no. 6 (2008): 1138-1149. Accessed April 26, 2018. doi:10.1016/j.marpolbul.2008.03.004.

US Congress. 2000. *Beaches Environmental Assessment and Coastal Health Act of 2000, Bill Text* 2000. https://www.congress.gov/106/plaws/publ284/PLAW-106publ284.pdf

US Congress. 2002. *Federal Water Pollution Control Act, As Amended Through P.L. 107-303,* November 27, 2002. https://www.epa.gov/sites/production/files/2017-08/documents/federal-water-pollution-control-act-508full.pdf.

USGS. 2018. 'The USGS Water Science School'. Last modified 2018. Accessed August 10, 2018. https://water.usgs.gov/edu.

Wymer, Larry J. 2007. "The Evolution of Water Quality Criteria in the United States, 1922–2003." In *Statistical Framework for Recreational Water Quality Criteria and Monitoring*, 1-11. Chichester, UK: John Wiley & Sons, Ltd.

Zeiler, M. and M. Murphy. 2010. *Modeling our World: The Esri Guide to Geodatabase Concepts*. 2nd ed. Redlands, CA: Esri Press.

Ziemann, David A. and Karen Klein. 2003. *Hokulia Water Monitoring Program Distributions and Relationships between Water Quality Parameters at Hokulia and Control Stations*. Waimanalo and Kailua-Kona: Oceanic Institute Makapuʻu Point and AECOS Laboratories of Hawaiʻi.

Appendix A Keāhole Point (LTS1)

Input Table:

Mr. Olson adds a new row to five worksheets in six different spreadsheet files every three months. The files represent the six transects and the worksheets represent the five stations on each transect. The stations lengths changed in 2007 so I called the 1993-2007 the “old” data, the “new” data is 2007-2017 (see Figure 37). Table 10 and 11 show the analyte results resulting from the SQL query shown below.

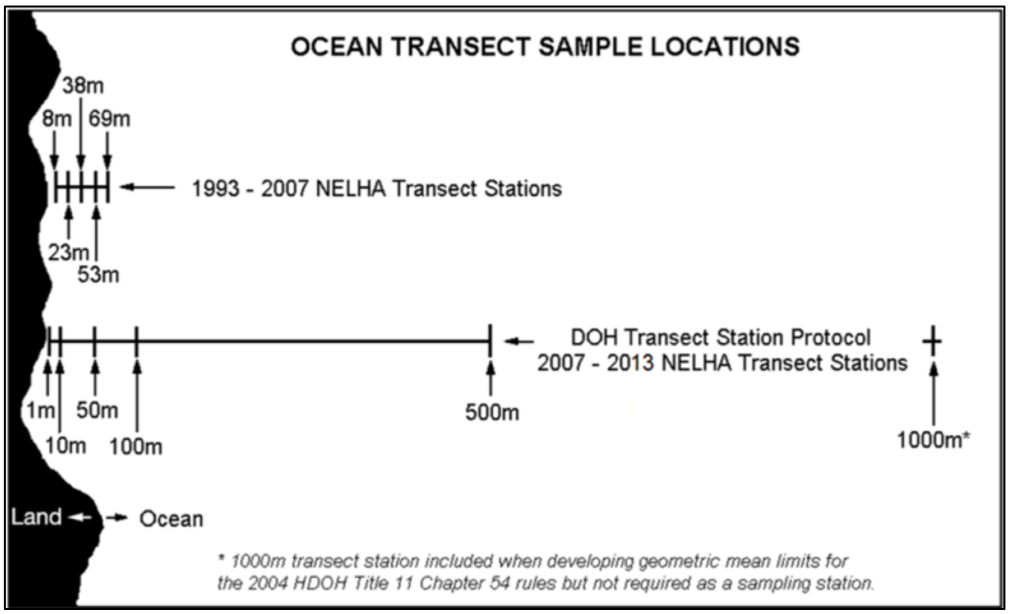


Figure 37. NELHA 2007 station distances change. Source: Olson 2017.

Table 10. Analyte Results: Keāhole Point Old 1993-2007. Source: Olson 2017.

Label	#rows	~gmean	stdev	min	max
AmmN6220 - Ammonia Nitrogen	1680	2.34	2.61	.14	67.93
ChlA6250 - Chlorophyll a	332	.09	.08	.01	.84
Clos6420 - Clostridium perfringens	29	1.59	5.20	1.00	29.00
Ecci6410 - Enterococci	44	1.45	1.45	1.00	9.00
NaNi6230 - Nitrate+Nitrite Nitrogen	1680	2.73	6.82	.14	71.99
OxyD6270 - Dissolved Oxygen	328	6.99	.46	6.10	9.86
Phos6391 - Phosphorus PO4	1680	3.51	4.64	.93	183.98
pHyd6390 - pH	336	8.25	.04	8.18	8.48
Sali6350 - Salinity	1678	34.58	.43	24.90	35.85
Sili6360 - Silicates	1680	118.92	137.04	14.60	1851.40

Label	#rows	~gmean	stdev	min	max
TDNi6210 - Total Nitrogen	340	77.38	14.36	37.40	145.53
TDPh6386 - Total Dissolved Phosphorus	336	10.83	2.30	6.50	33.14
Temp6280 - Temperature C	1676	26.08	1.14	23.50	29.30
TPho6240 - Total Phosphorus	336	11.18	2.97	7.12	48.32
TtOC6392 - Total Organic Carbon	275	1.06	.11	.80	1.48
Turb6260 - Turbidity	1620	.09	.10	.02	2.64
Total Count TBResult #rows	14,623				

```
SELECT 'Old 1993-2007' as 'NELHA 56 Quarters', Count(*) as 'TBResult #rows' FROM
NELHATBResult WHERE fkSample LIKE '%T56Quarters%'
SELECT Label, Count(dmRAll) as '#rows', round(exp(avg(log(Result))),3) as
'~gmean',round(stdev(Result),2) as 'stdev',cast(min(Result) as decimal (10,2)) as
'min',cast(max(Result) as decimal (10,2)) as 'max'
FROM NELHATBResult WHERE Result != 0 AND fkSample LIKE '%T56Quarters%'
GROUP by Label, dmRAll ORDER by Label
```

Table 11. Analyte Results: Keāhole Point New 2007-2017. Source: Olson 2017.

Label	#rows	~gmean	stdev	min	max
AmmN6220 - Ammonia Nitrogen	1230	2.52	3.11	.05	48.60
ChlA6250 - Chlorophyll a	1169	.12	.08	.01	.65
NaNi6230 - Nitrate+Nitrite Nitrogen	1259	2.95	5.44	.01	48.10
OxyD6270 - Dissolved Oxygen	1260	6.24	.62	4.05	9.92
Phos6391 - Phosphorus PO4	1259	2.67	2.36	.10	43.80
pHyd6390 - pH	1260	8.21	.06	7.95	8.70
Sali6350 - Salinity	1260	34.88	.25	33.81	35.81
Sili6360 - Silicates	1259	74.67	172.93	.10	2612.00
TDNi6210 - Total Nitrogen	1259	72.93	23.09	24.00	203.80
TDPh6386 - Total Dissolved Phosphorus	1259	11.56	3.47	.30	47.70
Temp6280 - Temperature C	1260	25.87	.98	23.82	28.73
Turb6260 - Turbidity	1260	.07	.04	.02	.50
Total Count TBResult #rows	15,075				

```
SELECT 'New 2007-2017' as 'NELHA 42 Quarters', Count(*) as 'TBResult #rows'
FROM NELHATBResult WHERE fkSample LIKE '%T42Quarters%'
SELECT Label, Count(dmRAll) as '#rows', round(exp(avg(log(Result))),3) as
'~gmean',round(stdev(Result),2) as 'stdev',cast(min(Result) as decimal (10,2)) as
'min',cast(max(Result) as decimal (10,2)) as 'max'
FROM NELHATBResult WHERE Result != 0 and fkSample LIKE '%T42Quarters%'
GROUP by Label, dmRAll ORDER by Label
```

SQL Query Script:
This is the script file "ProjectID_LTS1_KeaholeNELHA.sql" but edited to remove table creates (found in Appendix L) and repetitive sequences to save space.
-- **Step #1 and #2** Create "Raw" table for NELHA append in each of 6 XLS files

```
-- (ProjectID_LTS1_KeaholeTransect[1-6]_ImportXLSv3.sql)
CREATE TABLE NELHARaw (
 [EsriDateTime] [datetime2](7) NOT NULL, -- pk
 [AMorPM] [nvarchar](3) NOT NULL,
 [Transect] [int] NOT NULL,
 [mShore] [int] NOT NULL,
 [NewOrOld] [nvarchar](3) NOT NULL,
 [SampTimeHHMM] [nvarchar](255) NOT NULL,
 [LatituteN] [nvarchar](255) NULL,
 [LongitudeN] [nvarchar](255) NULL,
 [SiteId] [nvarchar](255) NOT NULL,
 [Phos6391] [numeric](38,8) NULL,--P043-
 [NaNi6230] [numeric](38,8) NULL,
 [AmmN6220] [numeric](38,8) NULL,
 [Sili6360] [numeric](38,8) NULL,
 [TDPh6386] [numeric](38,8) NULL,--
 [TDNi6210] [numeric](38,8) NULL,
 [TPho6240] [numeric](38,8) NULL,
 [TtOC6392] [numeric](38,8) NULL,
 [Turb6260] [numeric](38,8) NULL,
 [Sali6350] [numeric](38,8) NULL,
 [Temp6280] [numeric](38,8) NULL,
 [pHyd6390] [numeric](38,8) NULL,
 [OxyD6270] [numeric](38,8) NULL,
 [ChlA6250] [numeric](38,8) NULL,
 [Coli6420] [nvarchar](255) NULL,
 [Ecci6410] [nvarchar](255) NULL,
 POINT_Y [numeric](38,8) NULL,
 POINT_X [numeric](38,8) NULL,
 MyLatLong [nvarchar](255) NULL,
 MyLatitude [nvarchar](255) NULL,
 MyLabel [nvarchar](255) NULL,
 MyLocationLabel [nvarchar](255) NULL,
 MyCoreName [nvarchar](255) NULL,
 MyStartDate [datetime2](7) NULL,
 MyEndDate [datetime2](7) NULL,
 MyAnalyteList [nvarchar](255) NULL,
 MyTransectName [nvarchar](255) NULL,
 MyDateRange [nvarchar](255) NULL,
 MyQuarterCount [nvarchar](255) NULL,
 RawUniq [uniqueidentifier] NOT NULL,
 GeoUniq [uniqueidentifier] NULL,
 RawIdentityI [int] NOT NULL,
 RawIdentityA [int] NOT NULL,
PRIMARY KEY (EsriDateTime))
```

-- **Step #3** The SQL Server Import tool populated the “raw” table.

```
-- create our own identity attribute
Declare @myRow varchar(255), @MyRawUniq varchar(255)
Declare @cursorInsert CURSOR
set @cursorInsert = CURSOR FOR
Select ROW_NUMBER() OVER(ORDER BY Esridatetime),RawUniq from NELHARaw
OPEN @cursorInsert
FETCH NEXT FROM @cursorInsert into @myRow, @MyRawUniq
WHILE @@FETCH_STATUS = 0
BEGIN
UPDATE NELHARaw set RawIdentityA = @myRow where RawUniq = @MyRawUniq
FETCH NEXT FROM @cursorInsert INTO @myRow, @MyRawUniq
END
CLOSE @cursorInsert
DEALLOCATE @cursorInsert
-- Step #3 clean up strange text fields
UPDATE NELHARaw SET Ecci6410='0' WHERE Ecci6410 = '<1' -- 279 rows -- 'no sample' is just null
UPDATE NELHARaw SET Ecci6410=NULL WHERE Ecci6410 = 'no sample' -- 2 rows
UPDATE NELHARaw SET Ecci6410=NULL WHERE Ecci6410 = 'NULL' -- 0 rows ok
-- same for coliform field (Coli6420)
--------------------------------------------------------
-- Step #4 Create temporary table for georeferenced stations
CREATE TABLE NELHAMyGeocodedLocations_temp (
  Transect [int] NOT NULL,--pk+
  mShore [int] NOT NULL,--pk
  mShoreT [nvarchar](4) NULL,
  POINT_Y [numeric](38,8) NOT NULL,
  POINT_X [numeric](38,8) NOT NULL,
  MyLatLong [nvarchar](255) NULL,
  MyLatitude [nvarchar](255) NULL,
  GeoUniq [uniqueidentifier] NOT NULL,
  NewOrOld [nvarchar](3) NULL,
  StartDate [datetime2](7) NULL,
  EndDate [datetime2](7) NULL,
PRIMARY KEY (Transect, mShore))
-- Step #5 Assemble name strings and populate other fields
-- Insert rows with location information (only two rows shown here)
INSERT INTO NELHAMyGeocodedLocations_temp VALUES (1, 8, '008m', '19.73209', '-156.05717', NULL,NULL,newid(),NULL,NULL,NULL)
INSERT INTO NELHAMyGeocodedLocations_temp VALUES (6, 500, '500m', '19.71140', '-156.05440', NULL,NULL,newid(),NULL,NULL,NULL)
-- get dates for new series
UPDATE NELHAMyGeocodedLocations_temp SET NewOrOld = 'New', StartDate='2007-09-26 08:32:01', EndDate='2017-11-30 10:39:01' WHERE mShore IN (1,10,50,100,500) -- 30 half, good
-- get dates for old series
```

```
UPDATE NELHAMyGeocodedLocations_temp SET NewOrOld = 'Old', StartDate='1993-07-26
09:26:01', EndDate='2007-04-16 11:30:01' WHERE mShore IN (8,23,38,53,69) -- 30 other half,
good
-- create location names
UPDATE NELHAMyGeocodedLocations_temp SET
MyLatLong  = CAST(POINT_Y as nvarchar) + ', ' + CAST(POINT_X as nvarchar),
MyLatitude = CAST( REPLACE( str( CAST(POINT_Y AS DECIMAL (7,4)),7,4),'.','_') AS
varchar) -- 60
-- do this join manually
Declare @myRow varchar(255), @MyGeoUniq varchar(255), @MyPOINT_Y numeric(38,8),
@MyPOINT_X numeric(38,8)
Declare @MyTransect int, @MymShore int, @DMyLatLong varchar(255), @DMyLatitude
varchar(255)
Declare @cursorInsert CURSOR
set @cursorInsert = CURSOR FOR
 Select ROW_NUMBER() OVER (ORDER BY Transect,mShore),
       GeoUniq,
                  POINT_Y,POINT_X,
                  Transect,mShore,
                  MyLatLong,MyLatitude
                  FROM NELHAMyGeocodedLocations_temp
OPEN @cursorInsert
FETCH NEXT FROM @cursorInsert into @myRow, @MyGeoUniq, @MyPOINT_Y,
@MyPOINT_X,@MyTransect,@MymShore,@DMyLatLong,@DMyLatitude
WHILE @@FETCH_STATUS = 0
BEGIN
 UPDATE NELHARaw SET GeoUniq=@MyGeoUniq,
                POINT_Y=@MyPOINT_Y, POINT_X=@MyPOINT_X,
                                          MyLatLong=@DMyLatLong,
MyLatitude=@DMyLatitude
 WHERE
   Transect = @MyTransect AND mShore = @MymShore
  select @myRow, @MyGeoUniq,@MyPOINT_Y,@MyPOINT_X,@MyTransect,@MymShore
FETCH NEXT FROM @cursorInsert INTO @myRow, @MyGeoUniq, @MyPOINT_Y,
@MyPOINT_X,@MyTransect,@MymShore,@DMyLatLong,@DMyLatitude
END
-- file in name fields for later usage constructing the pkStation and pkSample
-- transect name/number
UPDATE NELHARaw SET
MyTransectName = 'NELHATransect' + LTRIM(str(Transect)) FROM NELHARaw raw -- 2940
-- location/distance from shore
UPDATE NELHARaw SET  MyLocationLabel = RIGHT('000' + LTRIM(str(mShore)), 3) + 'm'
FROM NELHARaw raw -- 2940
-- set quarters by new/old setting, old data 93-2007 was for 56 quarters, new 2007-2017 42
quarters
```

```
UPDATE NELHARaw SET MyQuarterCount = '56Quarters' FROM NELHARaw WHERE
NewOrOld = 'Old' -- 1680
UPDATE NELHARaw SET MyQuarterCount = '42Quarters' FROM NELHARaw WHERE
NewOrOld = 'New' -- 1260  1260+1680=2940 good
---------------------------------------------------------------------------------
-- transect and mshore are the label
UPDATE NELHARaw SET MyLabel = MyTransectName + ' ' + MyLocationLabel +
NewOrOld
FROM NELHARaw raw -- 2940
-- date/time, since NELHA so highly structured we just really have only one date
-- but that is ok since he took all samples quarterly
UPDATE NELHARaw SET  MyStartDate = Esridatetime, MyEndDate   = Esridatetime FROM
NELHARaw raw -- 2940
UPDATE NELHARaw SET
MyDateRange = SUBSTRING(convert(varchar, MyStartDate, 107),1,3) +
SUBSTRING(convert(varchar, MyStartDate, 107),9,4)
FROM NELHARaw raw -- 2940
-- create the corename
 UPDATE NELHARaw SET MyCoreName =
  MyLatitude + '_' +
  RIGHT('000'+LTRIM(str(mShore)),3) + 'm' + '_' +
  'KeaholePoint'    + '_' +
  'KOlson'   + '_' +
  'Transect' + str(Transect,1) + '_' +
  'T' + MyQuarterCount         + '_' +
  'Y' +  MyDateRange
FROM NELHARaw --2940
UPDATE NELHARaw SET MyCoreName = REPLACE(REPLACE(MyCoreName, ' ',''),'-','')
-- 2940
-----------------------------------------------------------------------
```

-- **Step #6** create and populate FCStation

```
CREATE TABLE NELHAFCStation (
        [pkStation] [nvarchar](255) NOT NULL,  (see Appendix L for full CREATE TABLE …)
PRIMARY KEY (pkStation))
INSERT INTO NELHAFCStation
SELECT
MyLatitude + '_' + mShoreT + '_KeaholePoint_KOlson_Transect' + str(Transect,1),-- pkStation,
temp unique
replace(str(transect),' ','') + '/' + mShoreT,   -- Label
NULL,                -- TablePage
NULL,                -- fkStation1
NewOrOld,            -- fkStation2 something anyway
mShore,              -- mShore
3499,              -- dmStA
6499,              -- dmStAA
3799,              -- dmStBott
```

```
7110,              -- dmStClas not applicable value
4660,              -- dmAccuracy from Google MyMaps geocoding
8855,              -- dmReef - not applicable value
7777,              -- dmRule - not applicable value
555,               -- dmStType
NULL,              -- fkEPA
NULL,              -- fkUSGS
'No',              -- Embayment
NULL,              -- VolumeBay
NULL,              -- CrossArea
197320,            -- AttachSt
NULL,              -- dmAccu4610
CAST(IIF(NewOrOld='New', MyLatLong, NULL) as nvarchar),  -- dmAccu4620, used
documentation from NELHA for new stations, but I converted to decimal since that is all I use
NULL,                -- dmAccu4630
NULL,                -- dmAccu4640
CAST(IIF(NewOrOld='New', NULL, MyLatLong) as nvarchar),  -- dmAccu4650 I geocoded all
the old stations
NULL,                -- dmAccu4660
NULL,                -- dmAccu4670
StartDate,           -- StartDate fill in later
EndDate,             -- EndDate
NULL,NULL,NULL,NULL,NULL,    -- 5 extra fields
POINT_X,POINT_Y,             -- XField, YField
IIF(mShore in (1,8),90,IIF(mShore in (10,23),180,IIF(mShore in (38,50),270,IIF(mShore in
(53,100),-90,0)))),
0,                   -- Normalize
0,                   -- Weight
0,NULL,                   -- PageQuery/S
GeoUniq, -- tracer to geocode table
newid(), -- StatUniq is assigned here since making a new record
GeoUniq  -- RawUniq tracer backwards
FROM NELHAMyGeocodedLocations_temp
ORDER by Transect, mShore -- 60
---------------------------------------------------------------------------
-- Step #7 Create temporary analyte table
CREATE TABLE NELHADateWithAnalyte_temp (
[Transect] [int] NOT NULL,
[mShore] [int] NOT NULL,
[NewOrOld] [nvarchar](3) NOT NULL,
[tPhos6391] [nvarchar](255) NULL,
[tNaNi6230] [nvarchar](255) NULL,
[tAmmN6220] [nvarchar](255) NULL,
[tSili6360] [nvarchar](255) NULL,
[tTDPh6386] [nvarchar](255) NULL,--5
[tTDNi6210] [nvarchar](255) NULL,
```

```
[tTPho6240] [nvarchar](255) NULL,
[tTtOC6392] [nvarchar](255) NULL,
[tTurb6260] [nvarchar](255) NULL,
[tSali6350] [nvarchar](255) NULL,--10
[tSali9999] [nvarchar](255) NULL,
[tTemp6280] [nvarchar](255) NULL,
[tpHyd6390] [nvarchar](255) NULL,
[tOxyD6270] [nvarchar](255) NULL,
[tChlA6250] [nvarchar](255) NULL,--15
[tColi6420] [nvarchar](255) NULL,
[tEcci6410] [nvarchar](255) NULL,--17
A_RawUniq [uniqueidentifier] NOT NULL,
A_NewUniq [uniqueidentifier] NOT NULL)
-- Step #7 populate temporary analyte table
INSERT into NELHADateWithAnalyte_temp
SELECT
Transect,mShore, NewOrOld,
NULL,NULL,NULL,NULL,NULL,
NULL,NULL,NULL,NULL,NULL,
NULL,NULL,NULL,NULL,NULL,
NULL,NULL, RawUniq, -- pk
newid() FROM NELHARaw – 2950
-- Step #7 Update the temporary analyte table from the raw table (only non-null values) for each
analyte
 UPDATE NELHADateWithAnalyte_temp SET tPhos6391 = 'Phos6391' FROM NELHARaw
raw
WHERE raw.Phos6391 IS NOT NULL AND
raw.RawUniq=NELHADateWithAnalyte_temp.A_RawUniq
-- perform update for each analyte (not shown here)
-- this is the final assemblage from above effort
 UPDATE NELHARaw SET MyAnalyteList =
RTRIM(ISNULL(tPhos6391,")+ISNULL(tNaNi6230,")+
ISNULL(tAmmN6220,")+ISNULL(tSili6360,")+ISNULL(tTDPh6386,")+ISNULL(tTDNi6210,"
)+
ISNULL(tTPho6240,")+ISNULL(tTtOC6392,")+ISNULL(tTurb6260,")+ISNULL(tSali6350,")+
ISNULL(tSali9999,")+ISNULL(tTemp6280,")+ISNULL(tpHyd6390,")+ISNULL(tOxyD6270,")
+
ISNULL(tChlA6250,")+ISNULL(tColi6420,")+ISNULL(tEcci6410,"))
FROM NELHADateWithAnalyte_temp
WHERE NELHARaw.RawUniq = NELHADateWithAnalyte_temp.A_RawUniq --2940
---------------------------------------------------------------------------
-- Step #8 Create and populate TBSample
CREATE TABLE NELHATBSample (
        [pkSample] [nvarchar](255) NOT NULL, see Appendix L for full CREATE TABLE …)
PRIMARY KEY (pkSample))
INSERT INTO NELHATBSample
```

```
SELECT
MyCoreName + '_' + MyAnalyteList,
MyLabel,             -- LabelName
NULL,                -- TitlePage
0,                   -- Alert not applicable at this site
'KOlson',            -- fkIDLoc
Transect,
mShore,
MyLatitude + '_' + RIGHT('000' + LTRIM(str(mShore)), 3) + 'm' +
'_KeaholePoint_KOlson_Transect' + str(Transect,1),-- fkStation, temp unique          --
NULL,                -- fkStatio2
'KeaholePoint',      --
newid(),             -- fkUniqID is the new if generated in this new row creation of the insert
'<<fkOrg>>',         -- *** GET ****
1420,                -- laboratory method domain field
MyStartDate,MyEndDate,
'Has Local Time',    -- and make such comment in TimeMissg field, we have time - local
NULL,NULL,
NULL,                -- no Comment yet
197320,              -- attach is same as FCStation and is Lat of Transect#1, Station#1 1m (new)
NULL,NULL,NULL,NULL,NULL,
0,0,0,1,NULL,        -- cartography fields
newid(),
RawUniq              -- for tracing to save source row id
FROM NELHARaw
ORDER BY RawIdentityA -- maintain order (if that means anything) --2940
```

-- **Step #9** Create and populate TBSample

```
CREATE TABLE NELHATBResult (
        [fkSample] [nvarchar](255) NOT NULL, see Appendix L for full CREATE TABLE …)
PRIMARY KEY (fkSample,dmRAll))
 -- 1/16
 INSERT INTO NELHATBResult
SELECT
samp.pkSample,
'Phos6391 - Phosphorus PO4',
samp.fkIDLoc,
samp.fkUnqID,
raw.Phos6391,
NULL, --stdev
6391, -- dmRAll
6464, -- dmRAMethod
6310, -- dmR11546
6391, -- dmRAnlyt
6430, -- dmRBEACH
NULL, -- Grade
NULL, --Comments
```

```
NULL, NULL, NULL, NULL, NULL, NULL,
raw.RawUniq, newid(),
samp.SampUniq  -- SampUniq
FROM NELHARaw raw
INNER JOIN NELHATBSample samp
ON raw.RawUniq = samp.RawUniq and raw.Phos6391 IS NOT NULL
ORDER BY raw.RawIdentityA -- 2939
```
-- **Step #9** perform the above insert for each analyte [Phos6391]- [Ecci6410]
-- **Step #10** QC Quality Checks (not shown here).
```
select sum(NaNi6230) FROM NELHARaw raw
WHERE raw.NaNi6230 IS NOT NULL
select sum(Result) FROM NELHATBResult
WHERE dmRall = 6230
--15538.63763000  both
```

Appendix B HDOH Clean Water Branch (LTS2)

Input File from PacIOOS:

The input file for HDOH Clean Water Branch was download from PacIOOS as a comma separated text file with a header indicating the analyte result types (www.pacioos.hawaii.edu, accessed January 17, 2018) (see Figure 38 for first few rows).

```
time (UTC),longitude (degrees_east),latitude (degrees_north),location_id,location_name,number_concentration_of_enterococcus_in_sea_water (CFU/100-milliliters),number_concentration_of_clostridium_perfringens_in_sea_water (CFU/100-milliliters),sea_water_quality_alert (1),
sea_water_temperature (Celsius),sea_water_salinity (PSU),sea_water_turbidity (NTU),sea_water_ph_reported_on_total_scale (pH),
mass_concentration_of_oxygen_in_sea_water (milligrams/liter),fractional_saturation_of_oxygen_in_sea_water (%)
1973-06-12T22:38:00Z,-155.90416,20.186945,1216,MAHUKONA LANDING,NaN,NaN,NaN,25.0,NaN,0.3,9.4,7.4,NaN
1973-06-12T23:25:00Z,-155.82944,20.014723,1217,MAUNA KEA BEACH HOTEL-OUTFALL,NaN,NaN,NaN,26.5,NaN,1.6,8.1,7.4,NaN
1973-06-12T23:55:00Z,-155.8228,20.024029,1225,Spencer Beach Park,NaN,NaN,NaN,27.5,NaN,1.1,9.1,6.9,NaN
1973-06-13T17:33:00Z,-155.97212,19.594566,1215,Magic Sands,NaN,NaN,NaN,24.0,NaN,0.1,7.8,6.7,NaN
```

Figure 38. HDOH Clean Water Branch Input File. Source: PacIOOS, www.pacioos.hawaii.edu/metadata/cwb_water_quality.html.

Table 12. Analyte Results: HDOH Clean Water Branch. PacIOOS, www.pacioos.hawaii.edu/metadata/cwb_water_quality.html.

Label	#rows	~gmean	stdev	min	max
Clos6420 - Clostridium perfringens	5647	1.13	8.33	.20	300.00
Ecci6410 - Enterococci	9804	4.10	114.70	.30	2005.00
O2fs6365 - Fractional_Saturation_of_Oxygen	4463	93.29	11.28	9.87	155.80
O2mc6364 - Mass_Concentration_of_Oxygen	5709	6.43	5.91	3.47	107.00
pHyd6390 - pH	5311	8.18	.20	6.70	9.40
Sali6350 - Salinity	7707	31.11	3.94	.79	38.00
Temp6280 - Temperature C	5961	25.92	1.66	20.12	34.40
Turb6260 - Turbidity	5366	1.26	6.62	.10	192.00
Total Count TBResult #rows	50,014				

```
SELECT '1983-2017 EPA&BEACHAct' as 'Clean Water Branch HDOH', Count(*) as 'TBResult #rows' FROM CWBTBResult; SELECT Label, Count(dmRAll) as '#rows', round(exp(avg(log(Result))),3) as '~gmean',round(stdev(Result),2) as 'stdev',cast(min(Result) as decimal (10,2)) as 'min',cast(max(Result) as decimal (10,2)) as 'max' FROM CWBTBResult
WHERE Result<>0 GROUP by Label, dmRAll ORDER by Label
```

Each row of the CWB data file contains a True/False flag indicating whether a Alert was issued base on the analyte results of that sample. Table 13 shows the top locations where alerts have been recorded.

Table 13. HDOH Clean Water Branch alert counts by location. Source: PacIOOS, www.pacioos.hawaii.edu/metadata/cwb_water_quality.html.

CWB#	Location	AlertCount
1208	KailuaPierStationD	35
1203	KahaluuBeachPark	31
1205	KailuaPierStationA1	18
1222	PuakoMiddleofLot	17
1235	BanyansSurfingArea	17
1204	KailuaPierStationA	16
1236	AnaehoomaluBay	12
1225	SpencerBeachPark	10
1246	Pelekane	10

SELECT CWBLocationCode, MyLabel, AlertCount FROM CWBFCStation WHERE AlertCount >= 10 ORDER by AlertCount DESC

Appendix C Waiakailio Bay (PC4)

Input Table:

There was one input data table for Waiakailio Bay (see Figure 39). Dollar produced this format for this project and ‘O‘oma Beach. He adjusted the distances from shore (DFS in second column) of the stations along the transect across the two projects based on a survey based on the location of SGD.

TABLE 2. Water chemistry measurements (in μg/L at three locations off of the Kohala coastline collected December 17, 2009 as part of the Kohala Shoreline LLC project. Data from two groundwater wells located upslope from the ocean sampling area are also presented. Abbreviations as follows: S=surface; D=deep; DFS=distance from shore, BDL= below detection limit. Red line separates samples with salinity of 32 ‰. For sampling transect locations, see Figure 1.

TRANSECT	DFS (m)	DEPTH (ft)	PO_4^{3-} (μg/L)	NO_3^- (μg/L)	NH_4^+ (μg/L)	Si (μg/L)	TOP (μg/L)	TON (μg/L)	TP (μg/L)	TN (μg/L)	TURB (NTU)	SALINITY (o/oo)	CHL a (μg/L)	TEMP (deg.C)	O2 (%sat.)	pH
	0 S	0	1.24	1.68	11.90	137.7	10.54	156.2	11.78	169.8	0.49	34.557	0.50	25.35	97.40	8.18
	1 S	1	0.62	3.08	4.76	128.4	9.30	134.7	9.92	142.5	0.65	34.549	0.24	25.46	99.43	8.19
	2 S	1	0.93	0.70	1.96	135.7	8.99	112.6	9.92	115.2	0.20	34.543	0.14	25.48	99.34	8.21
	3 S	2	1.24	3.92	3.22	128.7	8.68	113.0	9.92	120.1	0.25	34.594	0.12	25.53	100.23	8.16
	4 S	3	1.55	2.66	0.84	99.8	9.92	133.1	11.47	136.6	0.19	34.575	0.09	25.51	100.53	8.17
ORTH)	5 S	1	1.86	7.42	6.30	150.6	10.85	164.6	12.71	178.4	0.24	34.576	0.09	25.53	99.53	8.16

Figure 39. Waiakailio Bay Table 2. Source: Dollar 2015.

Table 14. Analyte Results: Waiakailio Bay. Source: Dollar 2015.

Label	#rows	~gmean	stdev	min	max
AmmN6220 - Ammonia Nitrogen	29	2.45	2.99	.28	11.90
ChlA6250 - Chlorophyll a	30	.09	.10	.02	.50
DogN6369 - Dissolved Organic Nitrogen	30	123.08	22.89	87.50	183.30
DogP6368 - Dissolved Organic Phosphorus	30	9.21	1.59	7.13	15.81
OxyS6367 - O2 Saturation	30	99.59	1.49	95.34	101.32
pHyd6390 - pH	30	8.18	.06	8.05	8.29
PO436396 - PO43 Phosphate	30	1.72	.99	.62	4.03
Sali6350 - Salinity	30	34.13	1.42	27.47	34.67
Sili6360 - Silicates	30	165.92	1005.86	43.56	5101.80
TDNi6210 - Total Nitrogen	30	141.69	73.31	95.70	465.50
Temp6280 - Temperature C	30	25.49	.15	24.99	25.76
TPho6240 - Total Phosphorus	30	11.14	1.77	8.99	17.05
Turb6260 - Turbidity	30	.29	10.94	.08	60.24
Total Count TBResult #rows	419				

```
SELECT 'Waiakailio Bay December 2009' as 'S.J.Dollar', Count(*) as 'TBResult #rows'
FROM WaiakailioTBResult; SELECT Label, Count(dmRAll) as '#rows',
```

```
round(exp(avg(log(Result))),3) as '~gmean',round(stdev(Result),2) as
'stdev',cast(min(Result) as decimal (10,2)) as 'min',cast(max(Result) as decimal (10,2)) as
'max' FROM WaiakailioTBResult WHERE Result>0 AND dmRAll not in (6393,6397)
GROUP by Label, dmRAll ORDER by Label
```

Appendix D ‘O‘oma Beach (PC3)

Input Table:

There were three input data tables for ‘O‘oma Beach (see Figure 40 for the table header from 2006).

TRANSECT SITE	STA. NO.	DFS (m)	PO_4^{3-} (μg/L)	NO_3^- (μg/L)	NH_4^+ (μg/L)	Si (μg/L)	DOP (μg/L)	DON (μg/L)	TDP (μg/L)	TDN (μg/L)	TURB (NTU)	SAL (ppt)	CHL a (μg/L)	TEMP (deg.C)	O2 (%sat)	pH
	1S	0	5.27	67.76	7.14	1,231	8.37	91.70	13.64	166.60	0.20	32.917	0.69	26.64	106.4	8.25
	2S	1	15.50	125.30	4.76	2,105	11.78	111.30	27.28	241.36	0.29	31.482	0.74	26.76	105.3	8.27
	3S	2	3.41	109.34	8.26	1,850	9.92	96.04	13.33	213.64	0.21	31.952	0.56	26.89	107.0	8.27
	4S	5	8.37	104.72	5.60	1,810	8.37	94.22	16.74	204.54	0.13	31.987	0.91	26.99	107.5	8.27

Figure 40. ‘O‘oma Beach 2006 Data Table 2. Source: Dollar 2008.

Table 15. Analyte Results: ‘O‘oma 1990-1992 analyte results. Source: Dollar 2008.

Label	#rows	~gmean	stdev	min	max
Ammn6220 - Ammonia Nitrogen	18	2.24	1.64	.82	7.19
ChlA6250 - Chlorophyll a	18	.08	.02	.04	.14
DOgN6369 - Dissolved Organic Nitrogen	18	86.09	12.91	68.00	111.40
DOgP6368 - Dissolved Organic Phosphorus	18	5.79	1.66	1.58	8.45
Nate6393 - NO3	18	7.49	260.49	.54	1076.94
pHyd6390 - pH	18	8.17	.03	8.08	8.22
PO436396 - PO43 Phosphate	18	4.46	17.22	2.21	73.78
Sali6350 - Salinity	18	31.18	5.93	11.87	34.59
Sili6360 - Silicates	18	284.27	3479.75	71.20	14299.72
TDNi6210 - Total Nitrogen	18	128.17	258.77	75.70	1153.30
Temp6280 - Temperature C	18	26.11	.82	23.60	26.90
TPho6240 - Total Phosphorus	18	12.56	16.07	8.25	76.07
Turb6260 - Turbidity	18	.13	.02	.09	.16
Total Count TBResult #rows	234				

```
SELECT '''O''oma Oct1990&May1991&Nov1991&Mar1992 GeoMean' as 'S.J.Dollar', Count(*) as 'TBResult #rows' FROM OomaTBResult WHERE fkSample LIKE '%TableA2Page43%'
SELECT Label, Count(dmRAll) as '#rows', round(exp(avg(log(Result))),3) as '~gmean',round(stdev(Result),2) as 'stdev',cast(min(Result) as decimal (10,2)) as 'min',cast(max(Result) as decimal (10,2)) as 'max' FROM OomaTBResult
WHERE Result<>0 AND fkSample LIKE '%TableA2Page43%' GROUP by Label, dmRAll ORDER by Label
```

Table 16. Analyte Results: ‘O‘oma 2002 analyte results. Source: Dollar 2008.

Label	#rows	~gmean	stdev	min	max
Ammn6220 - Ammonia Nitrogen	13	.80	1.39	.12	3.73
ChlA6250 - Chlorophyll a	21	.15	.10	.09	.45
DOgN6369 - Dissolved Organic Nitrogen	21	191.77	16.13	159.47	218.35
DOgP6368 - Dissolved Organic Phosphorus	21	7.50	1.33	5.02	10.79
Nate6393 - NO3	21	1.59	4.12	.29	18.33
OxyS6367 - O2 Saturation	21	93.01	5.85	83.00	101.00
pHyd6390 - pH	21	8.17	.05	8.08	8.28
PO436396 - PO43 Phosphate	21	1.60	1.29	.46	4.70
Sali6350 - Salinity	21	34.80	.10	34.48	34.89
Sili6360 - Silicates	21	149.62	89.80	68.19	408.37
TDNi6210 - Total Nitrogen	21	195.52	17.92	160.15	224.98
Temp6280 - Temperature C	21	26.82	.43	26.30	27.90
TPho6240 - Total Phosphorus	21	9.57	1.08	8.04	12.05
Turb6260 - Turbidity	21	.13	.17	.04	.80
Total Count TBResult #rows	286				

SELECT '''O''oma November 1,2002' as 'S.J.Dollar', Count(*) as 'TBResult #rows' FROM OomaTBResult WHERE fkSample LIKE '%TableA4Page45%';

SELECT Label, Count(dmRAll) as '#rows', round(exp(avg(log(Result))),3) as '~gmean',round(stdev(Result),2) as 'stdev',cast(min(Result) as decimal (10,2)) as 'min',cast(max(Result) as decimal (10,2)) as 'max' FROM OomaTBResult WHERE Result<>0 AND fkSample LIKE '%TableA4Page45%' GROUP by Label, dmRAll ORDER by Label

Table 17. Analyte Results: ‘O‘oma 2006 analyte results. Source: Dollar 2008.

Label	#rows	~gmean	stdev	min	max
Ammn6220 - Ammonia Nitrogen	30	6.03	4.17	.84	17.36
ChlA6250 - Chlorophyll a	30	.44	.38	.12	2.00
DOgN6369 - Dissolved Organic Nitrogen	30	98.98	17.11	59.36	147.56
DOgP6368 - Dissolved Organic Phosphorus	30	8.18	1.97	.62	12.71
Nate6393 - NO3	30	28.93	121.15	1.96	642.18
OxyS6367 - O2 Saturation	30	105.58	1.61	101.20	108.80
pHyd6390 - pH	30	8.23	.06	8.07	8.34
PO436396 - PO43 Phosphate	30	1.61	4.64	.31	19.53
Sali6350 - Salinity	30	31.93	3.40	17.15	34.73
Sili6360 - Silicates	30	811.58	1924.33	70.00	10342.00
TDNi6210 - Total Nitrogen	30	163.06	122.01	88.48	715.68
Temp6280 - Temperature C	30	27.06	.18	26.54	27.27
TPho6240 - Total Phosphorus	30	11.27	4.46	8.68	27.28
Turb6260 - Turbidity	30	.13	.13	.07	.78
Total Count TBResult #rows	420				

```
SELECT '''O''oma November 1,2006' as 'S.J.Dollar', Count(*) as 'TBResult #rows'
FROM OomaTBResult WHERE fkSample LIKE '%Table2Page34%';
SELECT Label, Count(dmRAll) as '#rows', round(exp(avg(log(Result))),3) as
'~gmean',round(stdev(Result),2) as 'stdev',cast(min(Result) as decimal (10,2)) as
'min',cast(max(Result) as decimal (10,2)) as 'max' FROM OomaTBResult WHERE
Result<>0 AND fkSample LIKE '%Table2Page34%' GROUP by Label, dmRAll ORDER
by Label
```

Appendix E Hokuli‘a (PC1)

Input Tables:

There were seven input data tables for Hokuli‘a (see Figure 41 for one of them). This format is flipped compared to other data files. The rows are the analyte results and columns are the stations.

Table 2 Control A-North Boundary

Site DFS Analyte(unit)	1 1S	2 10S	3 10B	4 50S	5 50B	6 100S	7 100B	8 500S	DL	%Rec
pH	---	---	---	---	---	---	---	---	0.01	NA
T(C)	---	---	---	---	---	---	---	---	0.10	NA
DO(mg/L)	---	---	---	---	---	---	---	---	0.01	NA
Salinity (ppt)	35.0	35.2		35.4		35.5		35.4	0.10	NA
Turbidity NTU)	0.57	0.44		0.30		0.17		0.28	0.01	NA
Ammonia-N (ug/L)	4	< 1		1		< 1		< 1	1.0	94
Total Dissolved Nitrogen(ug/L)	129	145		135		105		94	1.0	96
Nitrate+Nitrite Nitrogen(ug/L)	7	15		14		4		1	1.0	95
Total Dissolved Phosphorus(ug/L)	9	8		5		5		7	1.0	95
0-Phosphate (ug/L)	6	7		3		2		2	1.0	95
Chlorophyll *a* (ug/L)	0.50	0.39		0.27		0.16		0.15	0.01	NA
Silicates (ug/L)	216	328		309		142		75	1.0	98
Total Suspended Solids(mg/L)	3.0	2.9		2.8		3.4		2.3	1.0	

Figure 41. Hokuli‘a transect data table. Source: Ziemann and Klein 2003.

Table 18. Analyte Results: Hokuli‘a. Source: Ziemann and Klein 2003.

Label	#rows	~gmean	stdev	min	max
AmmN6220 - Ammonia Nitrogen	35	1.54	2.82	.50	10.00
ChlA6250 - Chlorophyll a	35	.31	.61	.10	3.24
NaNi6230 - Nitrate+Nitrite Nitrogen	35	8.53	64.09	1.00	272.00
Nate6393 - NO3	35	3.78	2.03	1.80	10.20
OxyD6270 - Dissolved Oxygen	10	6.78	.12	6.59	6.95
Pho06395 - 0-Phosphate	35	3.88	3.36	1.00	15.00
pHyd6390 - pH	10	8.27	.05	8.23	8.35
Sali6350 - Salinity	35	34.81	1.19	30.20	35.60
Sili6360 - Silicates	35	260.64	685.92	70.00	3223.00
TDNi6210 - Total Nitrogen	35	141.16	68.68	68.00	378.00
Temp6280 - Temperature C	10	27.77	.19	27.40	28.10

Label	#rows	~gmean	stdev	min	max
TPho6240 - Total Phosphorus	35	7.49	3.05	4.00	16.00
Turb6260 - Turbidity	35	.28	.10	.14	.57
Total Count TBResult #rows	380				

```
SELECT 'Hokulia Aug2003' as 'D.Ziemann', Count(*) as 'TBResult #rows' FROM
HokuliaTBResult;
SELECT Label, Count(dmRAll) as '#rows', round(exp(avg(log(Result))),3) as '~gmean',
round(stdev(Result),2) as 'stdev',cast(min(Result) as decimal (10,2))
as 'min',cast(max(Result) as decimal (10,2)) as 'max' FROM HokuliaTBResult
WHERE Result<>0 GROUP by Label, dmRAll
```

Appendix F Keopuka Lands (PC2)

Input Tables:

There were two input data tables for Keopuka Lands (see Figure 42 for the table header and first few rows from the 2000 survey).

STATION	Nitrate N	Ammonia N	Total N	Ortho-P	Total P	Silica	DOP	DON	Salinity (o/oo)	Turbidity ntu	Chl-a ug/l	Temp	Oxygen % Sat.	pH
1-S	27.66	3.64	125.58	8.68	16.43	688.60	7.75	94.08	33.966	0.16	0.108	26.2	102	8.21
2-S	8.26	2.80	111.72	6.20	13.95	202.72	7.75	100.66	34.526	0.09	0.117	25.8	101	8.20
3-B	4.48	1.82	115.06	7.75	14.26	114.24	6.51	106.78	34.623	0.09	0.114	25.8	101	8.20

Figure 42. Keopuka data table. Source: Brock 2000.

Table 19. Analyte Results: Keopuka from 1991. Source: Brock 2000.

Label	#rows	~gmean	stdev	min	max
AmmN6220 - Ammonia Nitrogen	7	3.16	1.77	1.96	7.14
ChlA6250 - Chlorophyll a	7	.09	.02	.07	.14
DogN6369 - Dissolved Organic Nitrogen	7	106.58	11.10	96.18	126.52
Nate6393 - NO3	7	2.36	16.45	.14	45.78
OxyD6270 - Dissolved Oxygen	7	9.73	1.44	7.44	11.16
OxyS6367 - O2 Saturation	7	101.42	.98	100.00	103.00
Pho06395 - 0-Phosphate	7	2.49	1.61	1.24	6.20
pHyd6390 - pH	7	8.24	.03	8.19	8.28
Sali6350 - Salinity	7	34.68	.08	34.58	34.76
Sili6360 - Silicates	7	155.62	393.90	45.08	1153.68
TDNi6210 - Total Nitrogen	7	96.08	21.24	81.20	143.22
Temp6280 - Temperature C	7	25.59	.12	25.50	25.80
TPho6240 - Total Phosphorus	7	8.95	1.37	7.44	11.78
Turb6260 - Turbidity	7	.07	.01	.07	.08
Total Count TBResult #rows	98				

```
SELECT 'Old 12/1991' as 'Keopuka,R.Brock', Count(*) as 'TBResult #rows' FROM
KeopukaTBResult WHERE fkSample LIKE '%YDec1991%'
SELECT Label, Count(dmRAll) as '#rows', round(exp(avg(log(Result))),3) as '~gmean',
round(stdev(Result),2) as 'stdev',
cast(min(Result) as decimal (10,2)) as 'min',
cast(max(Result) as decimal (10,2)) as 'max'
FROM KeopukaTBResult WHERE Result != 0 AND fkSample LIKE '%YDec1991%'
GROUP by Label, dmRAll ORDER by Label
```

Table 20. Analyte Results: Keopuka from 2000. Source: Brock 2000.

Label	#rows	~gmean	stdev	min	max
AmmN6220 - Ammonia Nitrogen	17	2.30	.68	1.54	3.64
ChlA6250 - Chlorophyll a	17	.10	.01	.07	.13
DogN6369 - Dissolved Organic Nitrogen	17	110.86	16.99	93.80	155.40
Nate6393 - NO3	16	2.64	10.79	.14	27.66
OxyD6270 - Dissolved Oxygen	17	8.57	1.42	6.82	11.47
OxyS6367 - O2 Saturation	17	101.47	.80	100.00	103.00
Pho06395 - 0-Phosphate	17	6.24	1.60	4.34	9.30
pHyd6390 - pH	17	8.17	.11	7.95	8.47
Sali6350 - Salinity	17	34.52	.26	33.97	34.71
Sili6360 - Silicates	17	120.99	224.80	42.28	688.80
TDNi6210 - Total Nitrogen	17	121.25	16.91	97.72	159.46
Temp6280 - Temperature C	17	25.32	.63	24.50	26.20
TPho6240 - Total Phosphorus	17	15.02	1.66	13.33	19.22
Turb6260 - Turbidity	17	.10	.02	.07	.16
Total Count TBResult #rows	238				

```
SELECT 'New 04/2000' as 'Keopuka,R.Brock', Count(*) as 'TBResult #rows' FROM
KeopukaTBResult WHERE fkSample LIKE '%YApr2000%'
SELECT Label, Count(dmRAll) as '#rows', round(exp(avg(log(Result))),3) as '~gmean',
round(stdev(Result),2) as 'stdev',
cast(min(Result) as decimal (10,2)) as 'min',
cast(max(Result) as decimal (10,2)) as 'max'
FROM KeopukaTBResult WHERE Result != 0 AND fkSample LIKE '%YApr2000%'
GROUP by Label, dmRAll ORDER by Label
```

Appendix G Kīholo Bay Buoy (UHH1)

Input File from PacIOOS:

The input file for Kīholo Bay Buoy was download from PacIOOS as a comma separated file with a header indicating the analyte result types (www.pacioos.hawaii.edu, accessed January 17, 2018) (see Figure 43 for first few rows).

```
time,temperature,conductivity,salinity,oxygen,oxygen_saturation,oxygen_saturation_concentration,turbidity,chlorophyll,cdom,nitrate
UTC,Celsius,S m-1,1e-3,kg m-3,1,kg m-3,NTU,kg m-3,QSDE,mmol m-3
2014-03-01T10:00:00Z,25.46,5.16,33.53,0.00728076,1.0792,0.006752548,0.45,3.7E-7,1.44,185.8
2014-03-01T11:00:00Z,25.58,5.21,33.83,0.007180828,1.0676,0.006723996,0.42,3.3999999E-7,1.73,196.8
2014-03-01T12:00:00Z,25.54,5.23,34.01,0.007123724,1.0603,0.006723996,0.38,3.0E-7,1.61,194.3
2014-03-01T13:00:00Z,25.52,5.29,34.45,0.006681168,0.9959,0.00670972,0.42,3.9E-7,1.21,168.6
2014-03-01T14:00:00Z,25.59,5.32,34.66,0.006652616,0.9943,0.006695444,0.42,3.8E-7,1.1,190.1
2014-03-01T15:00:00Z,25.53,5.31,34.64,0.00649558,0.9695,0.006695444,0.38,3.3999999E-7,1.06,184.2
2014-03-01T16:00:00Z,25.4,5.29,34.56,0.006152956,0.9163,0.006723996,0.44,3.5E-7,1.29,184.7
2014-03-01T17:00:00Z,25.35,5.28,34.52,0.006053024,0.8995,0.006723996,0.41,3.2E-7,1.33,184.3
2014-03-01T18:00:00Z,25.35,5.26,34.42,0.005539088,0.824,0.006723996,0.46,3.0E-7,1.59,175.8
```

Figure 43. Input file for Kīholo Bay Buoy project. Source: PacIOOS, www.pacioos.hawaii.edu/metadata/WQB03agg.html.

Table 21. Analyte Results: Kīholo Buoy. Source: PacIOOS, www.pacioos.hawaii.edu/metadata/WQB03agg.html

Label	#rows	~gmean	stdev	min	max
Clos6420 - Clostridium perfringens	535	.90	.61	.02	10.21
Cond6394 - Conductivity	588	4.84	.76	.04	5.39
O2sa6366 - Oxygen_Saturation	588	1.00	.05	.79	1.25
O2sc6367 - Oxygen_Saturation_Conc	588	6.78	.28	6.58	8.85
OxyD6270 - Dissolved Oxygen	588	6.75	.37	5.54	8.75
Sali6350 - Salinity	588	31.48	4.99	.23	35.12
Temp6280 - Temperature C	588	25.25	.63	21.12	27.10
Turb6260 - Turbidity	588	.35	.52	.19	10.21
Total Count TBResult #rows	5,831				

```
SELECT 'Kiholo Bay Buoy March 2014' as 'J.Adolf PacIOOS mmol', Count(*) as
'TBResult #rows' FROM KiholoTBResult
SELECT Label, Count(dmRAll) as '#rows', round(exp(avg(log(Result))),3) as
'~gmean',round(stdev(Result),2) as 'stdev',cast(min(Result) as decimal (10,2)) as
'min',cast(max(Result) as decimal (10,2)) as 'max'
FROM KiholoTBResult WHERE Result>0 AND dmRAll not in (6393,6397) GROUP by
Label, dmRAll ORDER by Label
```

Appendix H Puakō (UHH3)

Input Tables:

There were two input data tables, the microbiological table is shown in Figure 44.

Table 1. Average ± SE and [range] of *Enterococcus* (MPN/100 mL), *C. perfringens* (CFU/100 mL), and $\delta^{15}N$ in macroalgal tissue (‰) for shoreline stations at Puakō, Hawaiʻi. Superscript letters indicate significant groupings from One-way ANOVA and post-hoc Tukey's tests. $\alpha = 0.05$; n = 4.

Station	*Enterococcus*	*Clostridium perfringens*	$\delta^{15}N$
1	18 ± 8^{b} [9-43]	2 ± 1 [0-4]	$6.38 \pm 0.15^{a\text{-}c}$ [6.03-6.65]
2	74 ± 25^{ab} [37-143]	2 ± 1 [0-4]	$7.54 \pm 0.18^{a\text{-}c}$ [7.04-7.90]

Figure 44. Table 1 (partial). Source: Abaya 2016.

Table 22. Analyte Results: Puakō. Source: Abaya 2016.

Label	#rows	~gmean	stdev	min	max
Ammn6220 - Ammonia Nitrogen	16	1.39	4.92	.49	20.83
Clos6420 - Clostridium perfringens	16	4.80	3.01	2.00	12.00
Ecci6410 - Enterococci	16	307.42	681.26	18.00	2777.00
H4Si6389 - H4Sio4 Silicic Acid	16	262.20	171.50	95.00	652.00
N15i6388 - 15N	16	6.52	2.17	4.23	11.88
NaNi6230 - Nitrate+Nitrite Nitrogen	16	51.91	55.42	13.37	196.05
Phos6391 - Phosphorus PO4	16	1.28	1.80	.39	7.42
Sali6350 - Salinity	16	18.64	7.51	6.43	30.77
TDNi6210 - Total Nitrogen	16	68.20	56.07	22.00	221.00
TPho6240 - Total Phosphorus	16	1.66	1.94	.57	8.25
Total Count TBResult #rows	160				

SELECT 'Puako Nov2014-July2015' as 'L.Abaya', Count(*) as 'TBResult #rows' FROM AbayaTBResult; SELECT Label, Count(dmRAll) as '#rows', round(exp(avg(log(Result))),3) as '~gmean', round(stdev(Result),2) as 'stdev',cast(min(Result) as decimal (10,2)) as 'min',cast(max(Result) as decimal (10,2)) as 'max' FROM AbayaTBResult WHERE Result<>0 GROUP by Label, dmRAll ORDER by Label

Appendix I Honokōhau Harbor and Kealakekua Bay (UHH2h and UHH2k)

Input Table:

There was one table for both sites (see the first two columns in Figure 45). The format reflects the high aggregation nature of this data. For example, 'HNI' refers to the North Inside Honokōhau location where they had two fifty meter transects, these values are the geometric mean of all individual samples for the duration of the project.

Parameter	HNI
Salinity	34.5 ± 0.04
(8.39, 0.0002, 0.11)	AB
Temperature*	26.3 ± 0.27
(16.53, <0.0001, 0.19)	C
Nitrate + nitrite	0.36 ± 0.035
(2.62, 0.0526, 0.08)	AB
Phosphate	0.00 ± 0.005
(8.97, 0.2545, na)	AB
Silicate	1.51 ± 0.300
(2.94, 0.0351, 1.10)	C
Ammonium*	0.02 ± 0.016
(8.06, 0.0003, 0.21)	C
TDN	3.29 ± 0.201
(1.04, 0.4415, 0.67)	AB
$\delta^{15}N^*$	2.46 ± 0.492
(3.25, 0.0240, 0.34)	AB
Chlorophyll-*a*	0.11 ± 0.012
(1.87, 0.1426, 0.03)	A
Pheophytin	0.14 ± 0.005
(12.95, <0.0001, 0.02)	C

Figure 45. Table 1 (partial) input data table. Source: Parsons et al. 2008.

Table 23. Analyte Results: Honokōhau Harbor. Source: Parsons et al. 2008.

Label	#rows	~gmean	stdev	min	max
AmmN6220 - Ammonia Nitrogen	4	.05	.04	.02	.09
ChlA6250 - Chlorophyll a	4	.11	.01	.10	.13
N15i6388 - 15N	4	2.00	.56	1.29	2.50
NaNi6230 - Nitrate+Nitrite Nitrogen	4	.37	.02	.35	.40
Pheo6398 - Pheophytin	4	.15	.01	.14	.17
Sali6350 - Salinity	4	34.45	.13	34.30	34.60
Sili6360 - Silicates	4	1.77	.30	1.51	2.11
TDNi6210 - Total Nitrogen	4	3.45	.14	3.29	3.63

Label	#rows	~gmean	stdev	min	max
Temp6280 - Temperature C	4	26.48	.13	26.30	26.60
TPho6240 - Total Phosphorus	1	.01		.01	.01
Total Count TBResult #rows	40				

```
SELECT 'Honokohau Harbor' as 'M.Parsons', Count(*) as 'TBResult #rows' FROM
ParsonsTBResult WHERE fkSample LIKE '%Honokahou%'
SELECT Label, Count(dmRAll) as '#rows', round(exp(avg(log(Result))),3) as
'~gmean',round(stdev(Result),2) as 'stdev',cast(min(Result) as decimal (10,2)) as
'min',cast(max(Result) as decimal (10,2)) as 'max' FROM ParsonsTBResult
WHERE Result<>0 AND fkSample LIKE '%Honokahou%' GROUP by Label, dmRAll
ORDER by Label
```

Table 24. Analyte Results: Kealakekua Bay. Source: Parsons et al. 2008.

Label	#rows	~gmean	stdev	min	max
AmmN6220 - Ammonia Nitrogen	4	.30	.08	.23	.38
ChlA6250 - Chlorophyll a	4	.12	.01	.11	.13
N15i6388 - 15N	4	2.72	.29	2.33	3.01
NaNi6230 - Nitrate+Nitrite Nitrogen	4	.38	.06	.31	.45
Pheo6398 - Pheophytin	4	.16	.03	.13	.20
Sali6350 - Salinity	4	34.63	.10	34.50	34.70
Sili6360 - Silicates	4	2.83	.31	2.58	3.17
TDNi6210 - Total Nitrogen	4	1.84	1.57	.29	3.45
Temp6280 - Temperature C	4	27.00	.08	26.90	27.10
TPho6240 - Total Phosphorus	2	.02	.01	.01	.03
Total Count TBResult #rows	40				

```
SELECT 'Kealakekua Bay' as 'M.Parsons', Count(*) as 'TBResult #rows' FROM
ParsonsTBResult WHERE fkSample LIKE '%Kealakekua%'
SELECT Label, Count(dmRAll) as '#rows', round(exp(avg(log(Result))),3) as
'~gmean',round(stdev(Result),2) as 'stdev',cast(min(Result) as decimal (10,2)) as
'min',cast(max(Result) as decimal (10,2)) as 'max' FROM ParsonsTBResult
WHERE Result<>0 AND fkSample LIKE '%Kealakekua%' GROUP by Label, dmRAll
ORDER by Label
```

Appendix J Keauhou Bay (CZ1)

Table 25. Analyte Results: Keauhou Bay. Source: D. Mihalka 2017.

Label	#rows	~gmean	stdev	min	max
Colo6387 - Color	189	2.33	.68	1.00	4.00
Sali6350 - Salinity	259	53.52	7.24	29.10	66.00
Swel6384 - Swell Direction	257	265.06	20.96	225.00	315.00
SwHt6385 - Swell Height Inches	191	7.36	8.47	.10	56.00
Temp6280 - Temperature C	248	26.34	1.65	19.80	29.00
Turb6260 - Turbidity	202	.71	5.88	.01	51.00
WdCp6383 - Wind Chop Inches	196	1.52	5.38	.10	28.80
WdSw6382 - Wind Swell	196	1.23	2.44	.10	14.00
Wind6381 - Wind Direction	258	234.04	53.97	18.00	360.00
Total Count TBResult #rows	2048				

```
SELECT 'May-Dec2014 Weekly' as 'Keauhou Bay,D.Mihalka', Count(*) as 'TBResult
#rows' FROM MihalkaTBResult; SELECT Label, Count(dmRAll) as '#rows',
round(exp(avg(log(Result))),3) as '~gmean',round(stdev(Result),2) as
'stdev',cast(min(Result) as decimal (10,2)) as 'min',cast(max(Result) as decimal (10,2)) as
'max' FROM MihalkaTBResult WHERE Result!=0 GROUP by Label, dmRAll ORDER
by Label
```

Appendix K Hawai‘i Ocean Time series (Control Site)

Table 26. Analyte Results: Selected yearly analytes by count from station ALOHA (0-5m depth)

Sample Count	Year	Analyte	~gmean	stdev	min	max
30	1996	Sili6360-Silicates	1.431	0.33	1.05	2.58
24	2000	Sili6360-Silicates	1.334	0.42	0.71	2.67
23	1995	Sili6360-Silicates	1.337	0.46	0.64	2.63
22	2001	Sili6360-Silicates	1.482	0.27	1.10	2.21
20	1997	Sili6360-Silicates	1.31	0.34	0.82	1.93
19	2004	Sili6360-Silicates	0.96	0.13	0.77	1.21
18	2001	NaNi6230-Nitrate+Nitrite Nitrogen	0.045	0.12	0.01	0.40
17	2003	Sili6360-Silicates	0.957	0.22	0.65	1.62
16	2004	NaNi6230-Nitrate+Nitrite Nitrogen	0.042	0.04	0.01	0.14
16	1999	Sili6360-Silicates	1.576	0.53	1.22	2.78
15	2003	NaNi6230-Nitrate+Nitrite Nitrogen	0.036	0.03	0.01	0.14
13	1994	Sili6360-Silicates	1.399	0.46	0.98	2.35
13	2000	TDNi6210-Total Nitrogen	5.438	0.39	4.71	6.03
13	2000	TPho6240-Total Phosphorus	0.304	0.05	0.22	0.39
11	2004	ChlA6250-Chlorophyll a	90.397	37.49	47.00	185.00
11	2001	ChlA6250-Chlorophyll a	99.068	47.49	54.00	223.00
11	2002	NaNi6230-Nitrate+Nitrite Nitrogen	0.037	0.04	0.01	0.13
11	2015	NaNi6230-Nitrate+Nitrite Nitrogen	0.023	0.04	0.01	0.14
11	1995	pHyd6390-pH	8.094	0.01	8.09	8.11
11	2002	Sili6360-Silicates	1.147	0.15	0.92	1.39

Appendix L FoxDB Database Schema

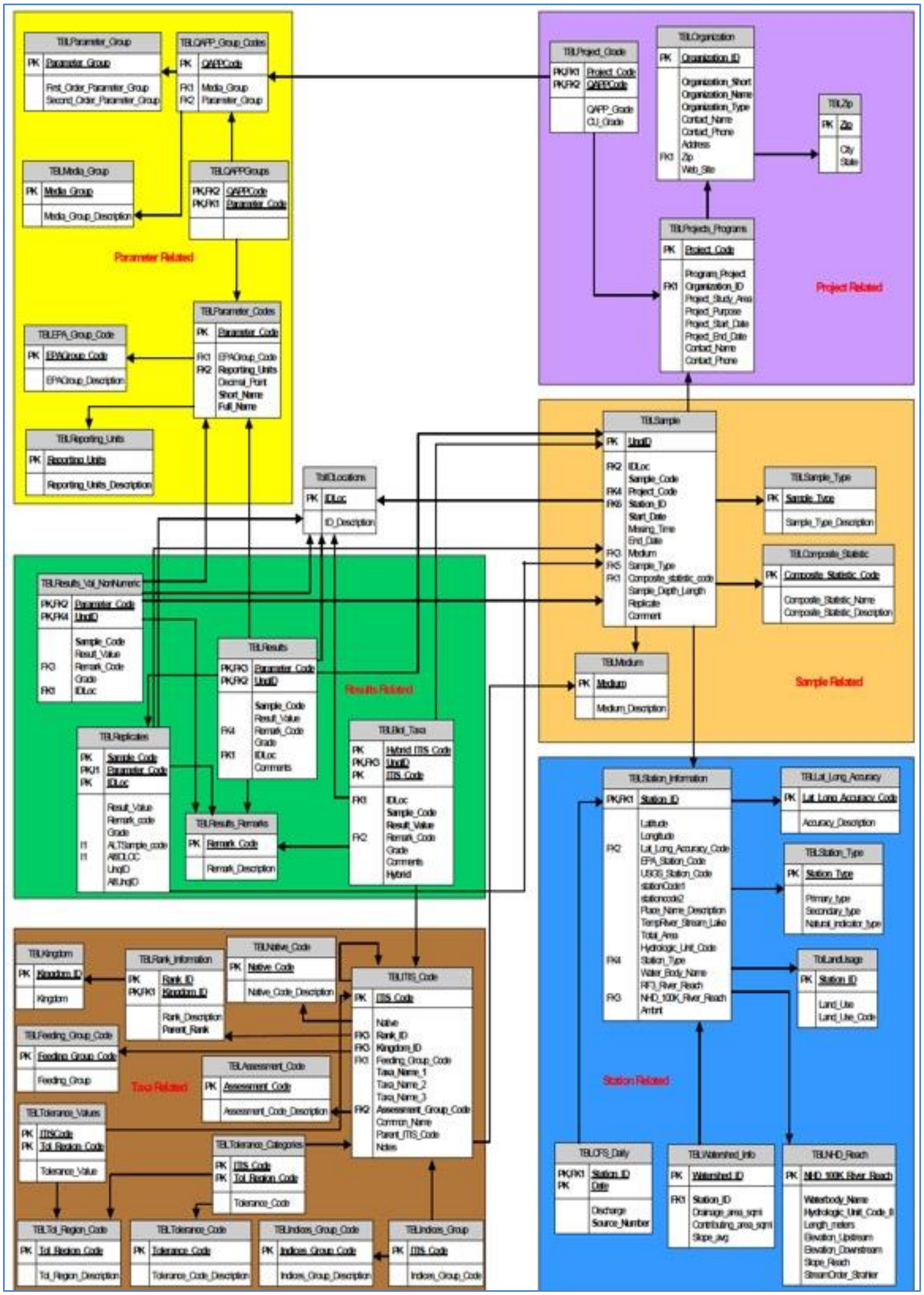


Figure 46. FoxDB Schema. Source: Bartošová et al. 2005.

Appendix M Geodatabase Data Dictionary

Table 27. *FCStation* description (all fields). McConkey et al. 2004 and Bartošová et al. 2005.

Field Name	Description	SQL Data Type	Nullable
pkStation	Unique identification for stations	nvarchar(255)	No
XField	Longitude in decimal degrees	numeric(38,8)	No
YField	Latitude in decimal degrees	numeric(38,8)	No
mShore	Distance from shore (meters)	int	Yes
StartDate	Earliest date	datetime2(7)	No
EndDate	Ending date	datetime2(7)	No
dmStAA	Class AA domain	int	No
dmStA	Class A domain	int	No
dmStType	Station type	int	No
dmStRule	Rule domain	int	No
Label	For GIS labels	nvarchar(255)	No
Other Fields			
dmStClas	Class domain	int	No
TablePage	Table and Page in source document	nvarchar(255)	Yes
dmAccuracy	Accuracy domain	int	No
dmStReef	Reef domain	int	Yes
AttachSt	Code for GIS attachment image	int	Yes
RawUniq	Identifies the row in source data	uniqueidentifier	No
fkStation1	Reserved for other organizations	nvarchar(255)	Yes
fkStation2	Reserved for other organizations	nvarchar(255)	Yes
fkEPA	The station code used by the USEPA and State. Not all stations have a code assigned by USEPA	nvarchar(255)	Yes
fkUSGS	The station code used by the USGS. Not all stations have a code assigned by USGS	nvarchar(255)	Yes
Embayment	The embayment name	nvarchar(255)	Yes
VolumeBay	Volume of the bay	numeric(38,8)	Yes
CrossArea	Cross section of embayment	numeric(38,8)	Yes
dmStBott	Bottom domain	int	Yes
dmAccu4610	GPS from PacIOOS	nvarchar(255)	Yes
dmAccu4620	GPS from documentation	nvarchar(255)	Yes
dmAccu4630	GPS from world geocoder	nvarchar(255)	Yes
dmAccu4640	GPS from Esri XY geocoder	nvarchar(255)	Yes
dmAccu4650	GPS from Google MyMaps	nvarchar(255)	Yes
dmAccu4660	GPS override	nvarchar(255)	Yes
dmAccu4670	GPS from Google Maps	nvarchar(255)	Yes

Field Name	Description	SQL Data Type	Nullable
StFloat1	Extra numeric	numeric(38,8)	Yes
StFloat2	Extra numeric	numeric(38,8)	Yes
StDate3	Extra date	datetime2(7)	Yes
StDate4	Extra date	datetime2(7)	Yes
StLong5	Extra integer	int	Yes
Rotation	GIS symbology field	int	No
Normalize	GIS symbology field	int	No
Weight	GIS symbology field	int	No
PageQuery	GIS feature not implemented	int	Yes
PageQueryS	GIS feature not implemented	nvarchar(255)	Yes
GeoUniq	Development field	uniqueidentifier	No
StatUniq	Development field	uniqueidentifier	No

***FCStation* SQL create command:**

```
CREATE TABLE FCStation (
	[pkStation] [nvarchar](255) NOT NULL,
	[Label] [nvarchar](255) NOT NULL,
	[TablePage] [nvarchar](255) NULL,
	[fkStation1] [nvarchar](255) NULL,
	[fkStation2] [nvarchar](255) NULL,
	[mShore] [int] NULL,
	[dmStA] [int] NOT NULL,
	[dmStAA] [int] NOT NULL,
	[dmStBott] [int] NULL,
	[dmStClas] [int] NOT NULL,
	[dmAccuracy] [int] NOT NULL,
	[dmStReef] [int] NULL,
	[dmStRule] [int] NOT NULL,
	[dmStType] [int] NOT NULL,
	[fkEPA] [nvarchar](255) NULL,
	[fkUSGS] [nvarchar](255) NULL,
	[Embayment] [nvarchar](255) NULL,
	[VolumeBay] [numeric](38, 8) NULL,
	[CrossArea] [numeric](38, 8) NULL,
	[AttachSt] [int] NULL,
	[dmAccu4610] [nvarchar](255) NULL,
	[dmAccu4620] [nvarchar](255) NULL,
	[dmAccu4630] [nvarchar](255) NULL,
	[dmAccu4640] [nvarchar](255) NULL,
	[dmAccu4650] [nvarchar](255) NULL,
	[dmAccu4660] [nvarchar](255) NULL,
	[dmAccu4670] [nvarchar](255) NULL,
	[StartDate] [datetime2](7) NOT NULL,
	[EndDate] [datetime2](7) NOT NULL,
	[StFloat1] [numeric](38, 8) NULL,
	[StFloat2] [numeric](38, 8) NULL,
	[StDate3] [datetime2](7) NULL,
	[StDate4] [datetime2](7) NULL,
	[StLong5] [int] NULL,
	[XField] [numeric](38, 8) NOT NULL,
	[YField] [numeric](38, 8) NOT NULL,
	[Rotation] [int] NOT NULL,
	[Normalize] [int] NOT NULL,
	[Weight] [int] NOT NULL,
	[PageQuery] [int] NULL,
	[PageQueryS] [nvarchar](255) NULL,
	[GeoUniq] [uniqueidentifier] NOT NULL,
	[StatUniq] [uniqueidentifier] NOT NULL,
	[RawUniq] [uniqueidentifier] NOT NULL
PRIMARY KEY (pkStation))
```

Table 28. *TBSample* description (all fields) McConkey et al. 2004 and Bartošová et al. 2005.

Field Name	Description	SQL Data Type	Nullable
pkSample	Unique sample identification number, number assigned if not provided by data originator. A sample is a monitoring activity (e.g., ambient samples, measurements, and observations) performed at a specific date, time, and location in order to characterize the environment	nvarchar(255)	No
fkStation	Unique number for the station, identifying the specific location, at which field work/sampling is conducted	nvarchar(255)	No
fkUnqID	Unique sample identification code for linking to *TBResult*	nvarchar(255)	No
StartDate	Date and time when field work/sample collection began	datetime2(7)	No
EndDate	Date and time when field work/sample collection ended	datetime2(7)	No
TimeMissg	Indicates whether sampling time was missing in the original data.	nvarchar(255)	Yes
dmSample	Letter code describing the sampling method	int	No
Label	For GIS labels	nvarchar(255)	No
Other Fields:			
Transect	Transect number	int	Yes
Medium	Not implemented	nvarchar(255)	Yes
CompStat	Not implemented	nvarchar(255)	Yes
AlertTrue	For CWB only, boolean indicates with sample caused Alert	int	Yes
TablePage	Table and page from source document	nvarchar(255)	Yes
Comment	Any comment	nvarchar(255)	Yes
RawUniq	Identifies the row in source data	uniqueidentifier	Yes
AttachSa	GIS attachment code	int	Yes
mShore	Distance from shore (meters)	int	Yes
fkOrg	Organization table link, not implemented	nvarchar(255)	No
fkStation2	Not implemented	nvarchar(255)	Yes
fkProject	Unique number, lookup is *TBProjects*	nvarchar(255)	No
fkIDLoc	FoxDB link (not implemented)	nvarchar(255)	No
SaFloat1	Extra numeric	numeric(38,8)	Yes
SaFloat2	Extra numeric	numeric(38,8)	Yes
SaDate3	Extra date	datetime2(7)	Yes

Field Name	**Description**	**SQL Data Type**	**Nullable**
SaDate4	Extra date	datetime2(7)	Yes
SaLong5	Extra integer	int	Yes
Rotation	GIS symbology field	int	No
Normalize	GIS symbology field	int	No
Weight	GIS symbology field	int	No
PageQuery	GIS feature not implemented	int	No
PageQueryS	GIS feature not implemented	nvarchar(255)	Yes
SampUniq	Development field	uniqueidentifier	Yes

***TBSample* SQL create command:**

```
CREATE TABLE TBSample (
	[pkSample] [nvarchar](255) NOT NULL,
	[Label] [nvarchar](255) NOT NULL,
	[TablePage] [nvarchar](255) NULL,
	[AlertTrue] [int] NULL,
	[fkIDLoc] [nvarchar](255) NOT NULL,
	[Transect] [int] NULL,
	[mShore] [int] NULL,
	[fkStation] [nvarchar](255) NOT NULL,
	[fkStation2] [nvarchar](255) NULL,
	[fkProject] [nvarchar](255) NOT NULL,
	[fkUnqID] [nvarchar](255) NOT NULL,
	[fkOrg] [nvarchar](255) NOT NULL,
	[dmSample] [int] NOT NULL,
	[StartDate] [datetime2](7) NOT NULL,
	[EndDate] [datetime2](7) NOT NULL,
	[TimeMissg] [nvarchar](255) NULL,
	[Medium] [nvarchar](255) NULL,
	[CompStat] [nvarchar](255) NULL,
	[Comment] [nvarchar](255) NULL,
	[AttachSa] [int] NULL,
	[SaFloat1] [numeric](38, 8) NULL,
	[SaFloat2] [numeric](38, 8) NULL,
	[SaDate3] [datetime2](7) NULL,
	[SaDate4] [datetime2](7) NULL,
	[SaLong5] [int] NULL,
	[Rotation] [int] NOT NULL,
	[Normalize] [int] NOT NULL,
	[Weight] [int] NOT NULL,
	[PageQuery] [int] NOT NULL,
	[PageQueryS] [nvarchar](255) NULL,
	[SampUniq] [uniqueidentifier] NULL,
	[RawUniq] [uniqueidentifier] NULL,
PRIMARY KEY (pkSample))
```

Table 29. *TBResult* description (all fields). McConkey et al. 2004 and Bartošová et al. 2005.

Field Name	**Description**	**SQL Data Type**	**Nullable**
fkSample	Unique sample identification number. A sample is a monitoring activity performed at a specific date, time, and location in order to characterize the environment.	nvarchar(255)	No
fkUnqID	Unique sample identification code for linking to *TBSample*.	nvarchar(255)	No
Result	Data value for a sample result or a code representing an observation. Result values can be numeric or alphanumeric values.	numeric(38,8)	No
Comment		nvarchar(255)	Yes
dmRAll	Analyte Domain containing all analytes	int	No
dmR11546	Analyte Domain for State Law	int	No
dmRBEACH	Analyte Domain for BEACH Act	int	No
dmRAnlyt	Analyte Domain for other than dmR11546 and dmRBEACH	int	No
Label	For GIS labels (full analyte code).	nvarchar(255)	No
Other Fields:			
dmRAMeth	Domain for methods	int	No
Stddev	Standard deviation	numeric(38,8)	Yes
RawUniq	Identifies the row in source data	uniqueidentifier	Yes
AttachR	GIS attachment code	int	Yes
Grade	Not implemented	nvarchar(255)	Yes
fkIDLoc	FoxDB link (not implemented)	nvarchar(255)	No
SampUniq	Development field	uniqueidentifier	Yes
ResUniq	Development field	uniqueidentifier	Yes
RFloat1	Extra numeric	numeric(38,8)	Yes
RFloat2	Extra numeric	numeric(38,)	Yes
RDate3	Extra date	datetime2(7)	Yes
RDate4	Extra date	datetime2(7)	Yes
RLong5	Extra integer	int	Yes

***TBResult* SQL create command:**

```
CREATE TABLE TBResult (
	[fkSample] [nvarchar](255) NOT NULL,
	[Label] [nvarchar](255) NOT NULL,
	[fkIDLoc] [nvarchar](255) NOT NULL,
	[fkUnqID] [nvarchar](255) NOT NULL,
	[Result] [numeric](38, 8) NOT NULL,
	[Stddev] [numeric](38, 8) NULL,
	[dmRAll] [int] NOT NULL,
	[dmRAMeth] [int] NOT NULL,
	[dmR11546] [int] NOT NULL,
	[dmRAnlyt] [int] NOT NULL,
	[dmRBEACH] [int] NOT NULL,
	[Grade] [nvarchar](255) NULL,
	[Comments] [nvarchar](255) NULL,
	[AttachR] [int] NULL,
	[RFloat1] [numeric](38, 8) NULL,
	[RFloat2] [numeric](38, 8) NULL,
	[RDate3] [datetime2](7) NULL,
	[RDate4] [datetime2](7) NULL,
	[RLong5] [int] NULL,
	[RawUniq] [uniqueidentifier] NULL,
	[ResUniq] [uniqueidentifier] NULL,
	[SampUniq] [uniqueidentifier] NULL,
PRIMARY KEY (pkSample, dmRAll))
```

Appendix N Analyte Domains and Codes

dmR11546– type LONG (REQUIRED, NON_NULLABLE), *TBResult*
TDNi6210,6210,Total Nitrogen
AmmN6220,6220,Ammonia Nitrogen
NaNi6230,6230,Nitrate+Nitrite Nitrogen
TPho6240,6240,Total Phosphorus
ChlA6250,6250,Chlorophyll a
Turb6260,6260,Turbidity
OxyD6270,6270,Dissolved Oxygen
COxy6272,6272,CTD Oxygen
Temp6280,6280,Temperature C
NotdmR11546,6299,Not applicable,Use other domain
dmRAnlyt– type LONG (REQUIRED, NON_NULLABLE), *TBResult*
Sali6350,6350,Salinity
Sili6360,6360,Silicates
O2mc6364,6364,Mass_Concentration_of_Oxygen_CWB
O2fs6365,6365,Fractional_Saturation_of_Oxygen_CWB
O2sa6366,6366,Oxygen_Saturation_PacIOOS
O2sc6367,6367,Oxygen_Saturation_Conc_PacIOOS
DOgP6368,6368,Dissolved Organic Phosphorus
DOgN6369,6369,Dissolved Organic Nitrogen
Wind6381,6381,Wind Direction
WdSw6382,6382,Wind Swell
WdCp6383,6383,Wind Chop Inches
Swel6384,6384,Swell Direction
SwHt6385,6385,Swell Height Inches
TDPh6386,6386,Total Dissolved Phosphorus
Colo6387,6387,Color
N15i6388,6388,15N
H4Si6389,6389,H4Sio4 Silicic Acid
pHyd6390,6390,pH
Phos6391,6391,Phosphorus PO4
TtOC6392,6392,Total Organic Carbon
Nate6393,6393,NO3
Cond6394,6394,Conductivity
Pho06395,6395,0-Phosphate
PO436396,6396,PO43 Phosphate
CDOM6397,6397,Chromophoric Dissolved Organic Matter
Pheo6398,6398,Pheophytin
NotdmAnlyt,6399,Not applicable,Use other domain
dmRBEACH– type LONG (REQUIRED, NON_NULLABLE), *TBResult*
Ecci6410,6410,Enterococci spp.
Clos6420,6420,Clostridium perfringens
NotRBEACH,6499,Not applicable,Use other domain

dmRAll– type LONG (REQUIRED, NON_NULLABLE), *TBResult*
dmR11546+dmRAnlyt+dmRBEACH
dmRAMeth – type LONG (REQUIRED, NON_NULLABLE), *TBResult*
6450 Method#1
6451 CWB
6460 Method#2
6462 Method#3 HOTs
6464 Method#4 NELHA
6465 Method#4 UHHilo
dmSample– type LONG (REQUIRED, NON_NULLABLE), *TBSample*
1410 Field
1420 Laboratory
dmStA – type LONG (REQUIRED, NON_NULLABLE), *FCStation*
Class A, Marine Water and Embayments from HAR 11-54
3420 Kawaihae Boat Harbor
3430 Honokohau Boat Harbor
3440 Keauhou Bay
3499 Class A Not Applicable
dmStAA– type LONG (REQUIRED, NON_NULLABLE), *FCStation*
Class AA, Marine Water and Embayments from HAR 11-54
6410 Puako Bay
6420 Waiulua Bay
6430 Anaehoomalu Bay
6440 Kiholo Bay
6450 Kailua Harbor
6460 Kealakekua Bay
6470 Honaunau Bay
6499 Class AA Not Applicable
dmStClas – type LONG (REQUIRED, NON_NULLABLE), *FCStation*
Classification from HAR 11-54
7110 Class AA Marine Waters and Embayments
7120 Class AA DLNR Embayments
7130 Class AA Marine Sanctuaries
7140 Class AA US Fish and Wildlife Service
7150 Class AA LeleiwiPt to WaiulaulaPt
7160 Class A   Marine Waters and Embayments
7170 Not Specified
dmAccuracy – type LONG (REQUIRED, NON_NULLABLE), *FCStation, FCProject*
4610 Pacific Islands Ocean Observing System
4620 From Documentation
4630 Esri World Geocoder
4640 Esri XY provider
4650 Google Maps
4660 Google Override
dmStRule – type LONG (REQUIRED, NON_NULLABLE), *FCStation*
From QAPP

7777 In Compliance
7776 On Alert
7775 Storm Event
7774 Waste Water Event
7773 Waste Water Condition
dmStType – type LONG (REQUIRED, NON_NULLABLE)**,** *FCStation*
Basic type of Station
111 Daily - Hawai‘i Clean Water Branch
222 EPA 303(d) Impaired Water Body
333 Kayak Scientists
444 Journal Articles
555 National Energy Laboratory of Hawai‘i Authority
808 UH Hawai‘i Ocean Time Series
700 Environmental Impact Reports
755 Buoy

Appendix O R Code to Extract Text from PDF File

This R script uses the package ‘pdftools’ provided free by Jeroen Ooms. (https://cran.r-project.org/web/packages/pdftools/pdftools.pdf, accessed May 27, 2018) to parse the text on each page looking for a specific table on the pages with water quality data. In the ‘O‘oma example below, the parse utility finds the page with the text “TABLE 1” and then writes the entire page to a text file.

```
# Ooma EIS appendix with water values
pdf_file <- "~/WestHawaiiWaterQuality4/WQData/19_7084_PC3_OomaBeach_SDollar/PDR2
 Appendix Ooma Marine Chemistry.pdf"

# pull in PDF using 'pdftools' pkg
text <- pdf_text(pdf_file)

w <- grep("TABLE 1.", text)

# ...grep found this on page 53 and 57, need to mine the page
# - not coded yet, send to console
cat(text[w])

write.table(text,
 "~/WestHawaiiWaterQuality4/WQData/19_7084_PC3_OomaBeach_SDollar/Ooma_RStudio_p
 dftools_Table1.txt", sep="\t")
```