Lake Erie Harmful Algal Bloom Early Warning System Operational Plan

Prepared by: Great Lakes Observing System

Project Partners:

NOAA Great Lakes Environmental Research Laboratory
NOAA National Centers for Coastal Ocean Science
LimnoTech
The Ohio State University
Cleveland Water Alliance
Cooperative Institute for Great Lakes Research
RPS Group

Prepared for:
NOAA Integrated Ocean Observing System Program Office
Final Report for Grant: NA17NOS0120175

Table of Contents

Table of Figures

Table of Tables

i. Preface

Harmful Algal Blooms (HABs) that occur during the summer in Lake Erie are a persistent annual problem that has threatened human health, affected the quality of life, and significantly degraded the ecosystem in Lake Erie. Roughly one-third of the total population of the Great Lakes basin is in the Lake Erie watershed and the lake is the primary source of drinking water for approximately eleven million people. Monitoring, research, and analysis increased in Lake Erie as a response to the 2014 drinking water crisis in Toledo, OH. However, the establishment of sustained observing assets, data integration pipelines, and optimized information products are in the early stages of maturity and require additional coordination, resources, and planning to ensure a viable transition to operational status. This project, aligned with the goals of the Ocean Technology Transition Program, advances the establishment of these foundational elements that are critical to a HABs Early Warning System (EWS) for Lake Erie. This Operational Plan is the final deliverable to the U.S. IOOS program under the IOOS Ocean Technology Transition (OTT) program.

Made possible through a three-year OTT grant awarded in 2017, the Great Lakes Observing System set out to understand and advance the vast network of local monitoring and data sharing activities underway in western Lake Erie to improve coordination and leverage resources in support of transitioning to a more mature HABs Early Warning System. As part of this effort, the project team:

- Engaged stakeholders in needs assessment and transition planning;
- Developed an inventory and assessment of core observing network assets and supported the upgrades and maintenance of assets for local water utilities;
- Procured and deployed an Environmental Sample Processor (ESP);
- Developed HABs EWS software in the GLOS IT platform; and
- Developed this Operational Plan as summary documentation with recommendations for future considerations.

We would like to acknowledge the project team: LimnoTech, Cleveland Water Alliance, Ohio State University and NOAA partners: Great Lakes Environmental Research Lab (GLERL), National Centers for Coastal Ocean Science (NCCOS), and the Cooperative Institute for Great Lakes Research (CIGLR). Advancing the Early Warning System is a collaborative endeavor and this project brings together the key organizations that are leading the annual monitoring, reporting, and dissemination of HABs data as well as representatives from the user groups affected by HABs, especially drinking water treatment plant operators.

This operational plan provides a high level documentation and guidance of the core observing network and the data management system. Additionally, the document summarizes the current status of the HABS EWS as well as recommended areas to improve the EWS. Understanding and communicating the impacts and risks associated with harmful algae blooms is critical for Lake Erie communities and we are pleased to be able to support the advancement of an early warning system.

1. Introduction and Background

1.1. Harmful Algal Blooms and Lake Erie

Harmful and nuisance algal blooms are a growing global threat for both sea- and freshwater systems. Nuisance algal blooms, while non-toxic, can impact economic activity by for example clogging industrial water intakes and fouling shorelines. They can also negatively impact ecosystem health by triggering low oxygen conditions that can lead to fish kills. Harmful algal blooms (HABs) negatively impact ecosystem health, economic activity, as well as public health since HABs can produce toxic substances that can cause irritation, illnesses and potentially death to humans and animals. The degradation of large HABs may also translate into hypoxic episodes.

Lake Erie has been impacted by algal blooms for many decades. Some have triggered hypoxic events with tens of thousands of dead fish, while others have produced toxins with concentrations that were many times over the World Health Organization limit for recreational contact and/or safe drinking water. The intensity of algal blooms has increased in more recent years, and climate change might exacerbate not just their intensity but also their toxicity in years to come.

Lake Erie borders the states of New York, Pennsylvania, Ohio, and Michigan, and the Canadian province of Ontario. Its 22,700 square miles watershed is highly urbanized and industrialized. In comparison to the other Great Lakes, its basin has the largest percentage of agricultural land use and is the most densely populated. The lake serves as a source of drinking water for about 12 million people in the U.S. and Canada. Lake Erie also supports the area's tourism, recreational fishing, and manufacturing industry, as well as several major shipping ports, and the largest commercial fishery of any Great Lake.

Lake Erie is the shallowest (62 ft average depth), and, by volume, the smallest of the Great Lakes. It is also the warmest, often exceeding 77°F during the summer and frequently freezing during the winter. According to the U.S. EPA[1] about 80% of Lake Erie's inflow comes from the Detroit River into the western basin of the lake, 11% comes from precipitation, and the rest from the other tributaries. Lake Erie receives the largest loads of phosphorus of the Great Lakes, and is also most subjected to sediment loading.

The western basin of Lake Erie is particularly shallow with an average depth of 24 ft. It receives about 61% of the lake's annual total phosphorus load, in comparison to 28% and 11% for the Central and Eastern basins respectively. These conditions, especially during the summer, make the western basin prone to algal blooms mostly dominated by the toxin-producing blue-green-algae (cyanobacteria) *Microcystis aeruginosa*. Some of those algal blooms can be carried by currents and waves towards the central basin where it might die and decompose. The conditions in the Eastern basin of Lake Erie promote the growth of a different type of non-toxic algae, Cladophora.

[1] https://www.epa.gov/greatlakes/lake-erie

In addition to the factors indicated above, the spread of invasive zebra and quagga mussels' compounds to make the algae problem in Lake Erie more difficult. In nearshore areas these mussels retain and recycle nutrients and increase water clarity which results in greater algae growth. The loss of wetlands once used to trap nutrients, the increasing temperatures, and higher occurrence of more intense spring storms, are also important contributing factors to algae growth.

1.2. Lake Erie HABs EWS

As a result of the many stressors mentioned earlier, Lake Erie has become the most prone of the Great Lakes to toxic cyanobacteria, especially in the western and central basins. The severity of Lake Erie's harmful algal blooms (HABs) has consistently increased since 2002 as indicated in Figure 1.1, which includes the 1 to 10 scaled bloom severity index.

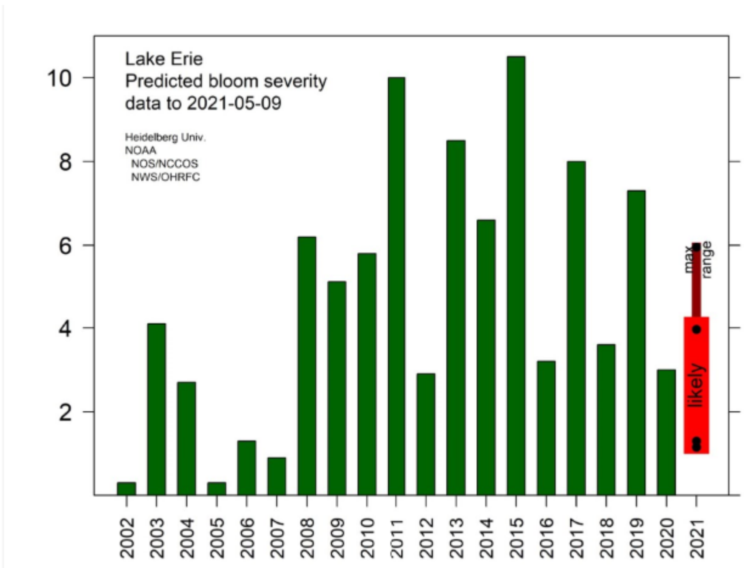


Figure 1.1: Western Lake Erie algae bloom severity index from 2002-2021.

In 2015 Ohio EPA provided funding to some water treatment plants (WTPs) in Lake Erie and the Great Lakes Observing System (GLOS) to procure and deploy water quality instrumentation with the goal of providing real-time data via GLOS' HABS Data Portal that came online in 2015. GLOS continued to provide support for this early warning system through 2016.

From 2017 through 2020, GLOS along with NOAA Great Lakes Environmental Research Laboratory (GLERL), the Cleveland Water Alliance (CWA), LimnoTech (LTI), and NOAA Center for Coastal Environmental Health and Biomolecular Research (CCEHBR) worked on a project to bring together the key organizations that are leading the annual monitoring, reporting, and dissemination of HAB data in Lake Erie, as well as representatives from the user groups affected by HABs, especially WTP operators. The project was funded by the Integrated Ocean Observing System (IOOS) Ocean Technology Transition Program (OTT) to stabilize and enhance 1) the in-situ monitoring capabilities, 2) the data management and communications support structure, and 3) the online HABs Data Portal for the Lake Erie early warning system.

An early warning system (EWS) is a set of technologies, policies and procedures designed to predict and mitigate the harm of natural and human-initiated disasters and other undesirable events. An EWS transforms multiple data sources into a signal forewarning users of possible threats as they arise, giving stakeholders more time to alleviate ecosystem and community impacts. The EWS for Lake Erie as defined in this study consists of a network of observing platforms, numerical models, and other data sources brought together in the GLOS Information Technology (IT) platform, where users can access data and set warnings based on thresholds selected for different parameters.

1.3. EWS components and Schema

The United Nations Office for Disaster Risk Reduction (UNDRR) International Strategy for Disaster Reduction[2] recommends that early warning systems have the following four components:

a) **Risk knowledge**: Data should be systematically collected and analyzed, with risk assessments performed.
b) **Monitoring and warning service**: Systems should be in place to monitor hazards and provide early warning services.
c) **Dissemination and communication**: Risk information and early warning messages must be delivered.
d) **Response capability**: Systems should be in place to respond to events.

For an end-to-end EWS as described above to be effective all four components need to be properly coordinated and re-assessed overtime. These types of systems can be implemented at various levels (national, state, company, community, etc.) depending on the risks and the ability to secure government and/or stakeholder involvement as pertinent. The Lake Erie EWS is not an end-to-end system that encompasses all elements as defined by UNDRR but an attempt at addressing parts of elements b) and c). The goal is to help provide stakeholders in Lake Erie, such as WTP managers, with timely and critical monitoring data to help them make informed decisions about activities and resources potentially impacted by the presence of algae blooms.

The Lake Erie HABs EWS has two tiers. The first tier, which is the current state of the EWS at the time of writing this report, is designed to deliver uninterpreted data to the practitioner/scientist. Trends in the water quality parameters sensed throughout the EWS network act as indicators of a HAB or hypoxic event, allowing the EWS user to take action. EWS users can set their own customized thresholds for various parameters to alert them as water quality changes. These features allow water managers and researchers in the Lake Erie region to have timely and user-friendly access to critical monitoring data to help them make informed decisions about activities and resources affected by HABs' presence and toxicity.

[2] https://www.unisdr.org/2006/ppew/whats-ew/basics-ew.htm Accessed July 15, 2021.

The second tier, which is an envisioned future system reachable through additional resources, is intended to deliver enhanced HABs information to the general public, commercial and recreational users of Lake Erie. Water quality parameters sensed throughout the network are assimilated into modeled outputs and analytics to calculate a HABs threat level index. This more sophisticated EWS is similar to the severe storm threat warnings issued by the National Weather Service. The second tier EWS has yet to be fully scoped; this operational plan contains recommendations to get the Great Lakes region closer to realizing it (see Chapter 4).

Table 1.1 compares the current EWS to the future desired EWS by the UNDRR defined components.

Table 1.1: Comparison of Tier 1 (existing) and Tier 2 (potential future) components for Lake Erie's harmful algal bloom early warning system.

EWS Components	Tier 1 Lake Erie HABs EWS	Potential Tier 2 Lake Erie HABs EWS
Risk knowledge	*Outside of EWS Scope*	Inclusion of seasonal forecast information
Monitoring and warning service	Core Observing Network; Data dashboard and customized alerts	Core Observing Network; Data dashboard and customized alerts EWS HABs threat level alerts
Dissemination and communication	Mobile alerts linked back to data dashboard	Mobile threat level alerts back to dashboard and disseminated through other weather related media
Response capability	*Outside of EWS Scope*	Users decide where/when it is safe to use Lake Erie

The Lake Erie HABs EWS can be further evolved beyond the two tiers described above. In order to realize an end-to-end EWS, there are several additional needs that need to be filled, particularly in the Risk Knowledge and Response Capability components. Addressing these needs will require involvement from regional partners who are aligned with these components, as well as additional resources to ensure proper implementation.

1.4. Partners' roles and responsibilities in operating the EWS

Operational partners are those that are responsible for operating a part of the EWS. They are composed of data contributors, model scientists, and GLOS to manage and deliver data (Figure 1.2).

Data Contributors

Serving the foundation of the EWS, 24 entities contribute data. Data types include real-time water quality data, physical attributes (i.e., wind, wave, surface temperature), in-situ toxicity data, and grab sample toxicity data. This group is comprised of the following:

- 17 water treatment utilities deploy monitoring equipment. In a few cases, buoys are deployed in Lake Erie close to where the raw source water enters the drinking water system. On the U.S. side, many of these monitoring stations are subsidized by the federal and state governments.
- 4 academic institutions deploy buoys in Lake Erie that monitor physical and water quality characteristics of Lake Erie. The buoys are located offshore, away from human development.
- The University of Michigan's Cooperative Institute of Great Lakes Research (CIGLR), in collaboration with the NOAA Great Lakes Environmental Research Lab (GLERL), deploy gliders, buoys, and Environmental System Processors (ESPs) to provide water quality and toxin data.
- 2 government agencies deploy a number of buoys in Lake Erie. The buoys are either research buoys, focused on collecting a wide subset of data, including meteorological and water quality, or operational buoys, focused primarily on physical parameters.

Modelers

Modelers use near real-time observational data to validate and refine model calculations. Models can estimate current and future conditions across Lake Erie, such as hypoxia and toxic blooms, creating spatial and temporal information that is not possible from monitoring efforts alone.

NOAA GLERL runs the Experimental Lake Erie Hypoxia Forecast Model, which provides the Lake Erie Hypoxia bulletin with information related to bottom water temperature and oxygen levels, as well as the predicted change of these parameters. NOAA's National Centers for Coastal Ocean Science provide an operational forecast for HABs in Lake Erie. Information for this is also distributed via a bulletin and focuses on the current satellite observed bloom position and the forecasted bloom position.

GLOS

Management and delivery of data and information is provided by GLOS and its IT platform. At the Tier 1 level, data and information is composed of direct measurements of Lake Erie water quality and alerts if measurements exceed a determined threshold. GLOS will provide these data to any entity well positioned to develop and operate the Tier 2 EWS. The Tier 2 EWS provider has yet to be determined.

Supporting Partners

Additional supporting partners are those not directly involved with the operation of the EWS, but may be involved in supplying hardware or software, maintaining components of the system, promoting technology innovation within the EWS, and advocating for the continued operations of the system. Some of the current, key supporting partners and their roles are:

- Cleveland Water Alliance: tracking innovative technologies in HABs detection
- Fondriest: fabrication and maintenance of monitoring equipment
- LimnoTech: deployment and maintenance of monitoring equipment
- SpinDance: backend IT platform development
- Dig: frontend user experience development

1.5. Stakeholder Needs

The EWS will serve the needs of a variety of stakeholders: drinking water plant managers, residents, charter captains, decision makers, lakeside business owners, media representatives, and marina owners. As part of this IOOS-OTT funded project, a stakeholder assessment[3] was conducted to identify stakeholder needs to help optimize the Lake Erie HABs EWS. Table 1.2 summarizes the results of those stakeholder outreach efforts.

Additional Stakeholder EWS Considerations

Preferred information channel and frequency from stakeholder groups (other than drinking water managers and residents) is online or email on a weekly basis. Residents would like to receive basic severity of HABs information as necessary, specifically for drinking water conditions, that is free through TV news or other local media outlets. In the EWS Tier 2, stakeholder groups would like to receive HABs and lake condition alerts in three modes: the drinking water watch mode, recreation alert mode, and drinking water alert mode on a daily basis with residents wanting recreation alert mode on an hourly basis.

For stakeholders who chose to consider the warning systems with more detailed information, there was a strong preference for systems that contained: current conditions, mixing forecast, transport forecast and satellite imagery (this mix of information was particularly important for decision makers)[3]. However, some stakeholders just wanted information on current conditions (i.e., business owners and the media).

[3] Tellez C, Wilson RS. 2020. Stakeholder assessment for a HABs early warning system. Columbus, OH: The Ohio State University, School of Environment and Natural Resources.

Table 1.2: An overview of the interviewed western Lake Erie stakeholder characteristics and identified needs related to harmful algal blooms.

STAKEHOLDER GROUP	*CHARACTERISTICS*	*NEEDS*
DRINKING WATER PLANT MANAGERS	Affected by harmful algal blooms (HABs)	Data curation, lag before public release, and regulatory firewall for initial data
	Vary in treatment plant size, environmental threats, size and expertise of staff	Predictive data analysis to serve for better and more cost effective plant operations
	Must comply with Ohio EPA standards	Receive alerts and view data that triggered them
	Rely on source water quality data from a wide array of sources	View water quality data displayed across regional map and/or in tabular format
		Observe changes in water quality as water body moves
		Be alerted to sudden changes in certain parameters
		Verify accuracy of alerts to researchers
		Share alerts with water treatment plant staff
CHARTER CAPTAINS	Average experience is 18 years with ~60 trips/year	Physical parameters (e.g. water temperature)
	Average trip is 2 hours each way with 6 anglers, with 50-75% of clients as repeat clients	Severity of bloom (e.g. size, toxin concentration) and health-related information (e.g. advisories)
	Majority of captains charter a 30-foot vessel	Short-term (i.e. 3 to 5-day) forecasts
	~1/3 of captains experienced a decreased in expected revenue since Toledo water crisis in 2014, with most indicating 20% or less decrease	Biological and water quality parameters (e.g. nutrients, fish, birds)
	Target fish species are walleye and perch	
DECISION MAKERS	Public sector representing municipalities and states	Short term (i.e. 3 to 5-day) forecasts
	Hear from constituents regarding HABs at variety of frequencies (e.g. daily, weekly, monthly)	Severity of bloom (e.g. size, toxin level) and health-related information (e.g. human exposure cases)

LAKESIDE BUSINESS OWNERS	Represent lodging or recreation and entertainment industry	Severity of bloom (e.g. size, toxin level) and health-related information (e.g. human exposure cases)
	Less than 20 employees	No preference in long-term vs. short-term forecasts
	Experienced decrease in expected revenue due to HABs in years since 2014	
	Half are seasonal vs. year-round businesses	
MARINA OWNERS	Average 30 years of ownership with 191 boat slips	Physical parameters
	Less than one-third experienced decrease in expected revenue due to HABs since 2014	Severity of bloom (e.g. size, toxin level) and health-related information (e.g. human exposure cases)
		Long-term (i.e. 10-day) forecasts
MEDIA REPRESENTATIVES	Represent newspapers and broad journalism	Short term (i.e. 3 to 5-day) forecasts
		Severity of bloom (e.g. size, toxin level) and health-related information (e.g. human exposure cases)
		Need trusted sources and content of information about HABs that will be shared with public
RESIDENTS	30% live between 1-5 miles and 40% live between 5-20 miles from the nearest shore	Do not want to receive as much early warning system information
	Small percentage experienced decrease in property value since purchase due to HABs	Short term (i.e. 3 to 5-day) forecasts
	Believe their home's tap water comes from Lake Erie or are unsure of its source	

2. Core Observation Network

2.1 Network Status

2.1.1 Overview of existing network

The Lake Erie HABs EWS is composed of many observing assets, and other monitoring and forecasting tools. The deployment of monitoring buoys in Lake Erie started in 2009. Approximately 5% of the existing Lake Erie assets were deployed in 2011, 40.5% in 2015 following the 2014 Toledo water crisis, 14.3% in 2016, and 33% between 2018 and 2020. The monitoring network focuses on the western and central Lake Erie basins, and provides measurements of physical, biological, and chemical characteristics of lake water. Most monitoring platforms are deployed seasonally, usually from April through early November.

The network's main purpose is to monitor near real-time water quality conditions (approximately every 10 -15 mins) related to seasonal HABs and hypoxia, two major threats to the Lake Erie ecosystem that can lead to important public health and safety concerns. The network provides value to water treatment plant (WTP) managers by providing information that helps monitor lake conditions in the western basin in general, as well as near real-time raw water quality data closer to the WTP's intakes. The network also provides value to those interested in recreational activities in Lake Erie.

2.1.2 Monitoring Network

The Lake Erie HABs and hypoxia monitoring network was formed by existing monitoring platforms deployed from 2009 through 2016, owned and operated by municipalities and/or WTPs, research and academic institutions, non-profit and private organizations, as well as the provincial government of Ontario, as indicated in Table 2.1. Funding for these monitoring stations came at the time from a variety of federal, state/provincial, and local sources, while the maintenance of each observing asset was the responsibility of the owners. In 2017 and 2018, Ohio State University and Bowling Green State University deployed two additional buoys and sondes. The University of Windsor, ON, deployed four buoys in 2019. This near real-time network of observing assets measures the parameters indicated in Table 2.1, and the data is shared via GLOS data portals.

From 2017 through 2020, GLOS, Cleveland Water Alliance (CWA), Ohio State University (OSU), NOAA GLERL and LimnoTech, Inc (LTI) worked with WTPs to, among other things, upgrade the core HABs and hypoxia monitoring network and to expand its geographical footprint by including other WTPs along the shores of Lake Erie with a three-year grant from the IOOS Ocean Technology Transition (OTT) project.

In 2018 and 2019, five additional US water treatment facilities, located mostly in the western Lake Erie basin, were added to the network. Sondes were deployed at these treatment plants' water wells, intakes, and/or remote pump stations depending on the location. In 2020, the first Canadian utility, Union Water Supply System (UWSS), was also added to the network.

The telecommunication data service and the hardware needed by the five WTPs that were added to the network in 2018-2019 were funded by the IOOS-OTT funded project. Costs to maintain the new utilities' equipment were covered by the project, as well. The OTT project supported calibrations, servicing, and upgrades to most of the monitoring equipment owned by the WTPs that form part of the original Lake Erie HABs and hypoxia observing network.

Table 2.2 lists all observing assets that were added to the core observing network from 2017 to 2020, along with the measurements they collect.

Overall as of 2020, data from thirty-six monitoring buoys and/or shore-based sondes was available in near real-time (approximately every 10-15 minutes) via the GLOS HABs data portal. The monitoring equipment was owned and/or operated by fifteen WTPs located in the US and one in Canada, several academic institutions located in Michigan, Ohio, and Ontario, NOAA GLERL, the Ontario Ministry of the Environment and Climate Change, and several private organizations.

It should be noted that by the end of the OTT project on September 30, 2020, the expectation was that the additional five water treatment plants would choose to keep the monitoring equipment and support its maintenance going forward. The maintenance includes yearly manufacturer tune-ups, occasional probe replacements, and monthly to quarterly calibrations, depending on the time of year and in-situ environmental conditions. However, three of the five utilities, those located in Carroll Township, the City of Vermillion, and the City of Huron, eventually chose not to keep the equipment.

Another goal of the IOOS-OTT funded project was to establish a reliable and operational network of Monterey Bay Aquarium Research Institute (MBARI) Second Generation (2G) Environmental Sample Processors (ESP). The first was deployed in Lake Erie in 2017 to provide near real-time microcystin toxin data. Currently, the network consists of three 2G ESPs deployed seasonally at the following locations: about five miles west of the Toledo, Ohio WTP intake, and nearby the Monroe, Michigan WTP intake (at GLERL's WE8 monitoring station; see Figures 2.1 and 2.2). The purpose of this effort was twofold: 1) to enhance the spatial and near-real time resolution of offshore microcystin toxin measurements in the western Lake Erie basin, and 2) to display microcystin toxin concentrations at the ESP sites on the GLOS Lake Erie HABs Data Portal. GLOS is working to provide visualizations of the ESP data in the portal in 2021.

*Table STYLEREF 1 \s 2. SEQ Table * ARABIC \s 1 1: Monitoring platforms and parameters measured by the original Lake Erie HABs Observing Network (2009-2016).*

Owner / Operator	Platform type	Station ID	Start Year	Main Parameters Measured																						
				Chlorophyll	Water Temp.	Specific Conductivity	Dissolved Oxygen	Dissolved Oxygen at Saturation	Fluorescent Dissolved Oxygen Matter	Turbidity	pH	Blue Green Algae	Significant Wave Height	Wave Period	Max. Wave Height	Wave Direction	Wind Direction	Wind Speed	Wind Gust	Air Temp.	Air Pressure	Water Current
City of Oregon, OH	Sonde	LEORGN	2015	✓	✓	✓			✓	✓	✓	✓										
Ottawa County Regional Water Treatment Plant, OH	Sonde	LEOC	2015	✓	✓	✓	✓	✓	✓	✓	✓	✓										
City of Avon Lake, OH	Sonde	LEAVON	2015	✓	✓	✓	✓	✓	✓		✓	✓										
Aqua Ohio - City of Ashtabula, OH	Sonde	LEASH	2015	✓	✓	✓	✓	✓		✓	✓	✓										
Aqua Ohio - City of Mentor, OH	Sonde	LEMENTOR	2015	✓	✓	✓	✓	✓	✓	✓	✓	✓										
Village of Marblehead, OH	Sonde	LEMRBHD	2015	✓	✓	✓					✓	✓										
City of Elyria, OH	Sonde	LEELYRIA	2015	✓	✓	✓				✓	✓	✓										
City of Toledo, OH/LTI	Moored buoy	45165	2009	✓	✓	✓				✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
City of Toledo, OH/LTI	Moored buoy	TOLCRIB	2015	✓	✓	✓				✓	✓	✓					✓	✓	✓	✓	✓	
City of Toledo, OH/LTI	Sonde	TOLLSPS	2015	✓	✓	✓			✓	✓	✓	✓										
City of Cleveland, OH/LTI	Moored buoy	45176	2016	✓	✓ ^	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
City of Cleveland, OH/LTI	Sonde	45176b	2016	✓	✓	✓	✓	✓	✓	✓	✓	✓										
City of Cleveland, OH/LTI	Moored buoy	45164	2011		✓ ^	✓	✓	✓	✓				✓	✓	✓	✓	✓	✓	✓	✓	✓	
City of Cleveland, OH/LTI	Moored buoy	45169	2015		✓ ^	✓	✓	✓	✓				✓	✓	✓	✓	✓	✓	✓	✓	✓	
NOAA-GLERL	Moored buoy	GLERLWE13	2016	✓	✓	✓	✓	✓	✓	✓		✓					✓	✓	✓	✓	✓	
NOAA-GLERL	Moored buoy	GLERLWE2	2015	✓	✓	✓	✓	✓	✓	✓		✓					✓	✓	✓	✓	✓	
NOAA-GLERL	Moored buoy	GLERLWE4	2015	✓	✓	✓	✓	✓	✓	✓		✓					✓	✓	✓	✓	✓	
NOAA-GLERL	Moored buoy	GLERLWE8	2015	✓	✓	✓	✓	✓	✓	✓		✓					✓	✓	✓	✓	✓	
Ohio State University (OH)	Moored buoy	OSUGI	2015	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓			✓	✓		✓	✓	
Bowling Green State University (OH)	Sonde	BGSDB*	2015																			
Bowling Green State University (OH)	Moored buoy	BGSUSD*	2015																			
University of Toledo (OH)	Moored buoy	UTLCP	2015	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓			✓	✓		✓	✓	
Great Lakes Research Consortium (NY)	Moored Buoy	ESF3	2009		✓ ^	✓	✓	✓									✓	✓		✓	✓	
Ontario Ministry of the Environment and Climate Change (ON, Canada)	Moored buoy	OMOECC_E1	2016	✓^	✓^	✓^	✓^	✓^		✓^		✓^										✓^
Ontario Ministry of the Environment and Climate Change (ON, Canada)	Moored buoy	OMOECC_01	2016	✓^	✓^	✓^	✓^	✓^		✓^		✓^										✓^
Regional Science Consortium (PA)	Moored buoy	45167	2011		✓ ^		✓	✓	✓	✓	✓		✓	✓		✓	✓	✓		✓	✓	
Regional Science Consortium (PA)	Moored buoy	PA-DEP-1538	2016	✓	✓ ^	✓	✓	✓		✓	✓											

* Decommissioned

^ Measurements at various depths

Table 2.2: Monitoring platforms and parameters measured by the extended Lake Erie HABs Observing Network (2017-2020).

Owner / Operator	Platform type	Station ID	Start Year	Main Parameters Measured																		
				Chlorophyll	Water Temp.	Specific Conductivity	Dissolved Oxygen	Dissolved Oxygen at Saturation	Fluorescent Dissolved Oxygen Matter	Turbidity	pH	Blue Green Algae	Significant Wave Height	Wave Period	Max. Wave Height	Wave Direction	Wind Direction	Wind Speed	Wind Gust	Air Temp.	Air Pressure	Current
Bowling Green State University (OH)	Moored Buoy	BGUSD2	2017	✓	✓	✓	✓	✓	✓	✓	✓	✓					✓	✓		✓	✓	
Ohio State University (OH)	Moored Buoy	OSUSS	2018	✓	✓	✓				✓	✓	✓										
Bowling Green State University (OH)	Sonde	SBEDISON	2018	✓	✓	✓	✓	✓	✓	✓	✓	✓										✓
City of Lorain, OH	Sonde	LELORAIN	2018	✓	✓	✓	✓	✓	✓	✓	✓	✓										
City of Sandusky, OH / BGSU	Sonde	LEBIWW	2018	✓	✓	✓	✓	✓		✓	✓	✓										
City of Huron, OH	Sonde	LEHURON	2019	✓	✓	✓					✓	✓										
City of Vermilion, OH	Sonde	LEVERM	2019	✓	✓	✓				✓	✓	✓										
Carrol Township, OH	Sonde	LECARR	2019	✓	✓	✓				✓	✓	✓										
Village of Put-In-Bay, OH	moored-buoy	SBIPIB	2019	✓	✓	✓				✓	✓	✓										✓
Middle Bass, OH^^	Sonde	TBD	TBD																			
Univ. of Windsor - RAEON	moored-buoy	UWRAEON1*	2019	✓	✓ ^		✓	✓				✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Univ. of Windsor - RAEON	moored-buoy	UWRAEON2**	2019	✓	✓ ^		✓	✓				✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Univ. of Windsor - RAEON	moored-buoy	UWRAEON3***	2019	✓	✓ ^		✓	✓				✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Univ. of Windsor - RAEON	moored-buoy	UWRAEON4	2019	✓	✓ ^		✓	✓				✓	✓	✓	✓	✓	✓	✓		✓	✓	

^ Measurements at various depths

^^ Pending installation

* Renamed UWSS RAEON 1 in 2020

** Renamed UWSS RAEON 2 in 2020

*** Renamed UWSS RAEON 3 in 2020

In addition to the data gathered by the network of moored buoys, shore-based sondes, and ESPs, GLERL and CIGLR also conduct:

- weekly airborne campaigns to capture high-resolution, hyperspectral HABs imagery in the western and central basins of Lake Erie.
- Uncrewed underwater system deployments, as needed, during the bloom season to capture observations at specific locations or over desired transects. GLERL, CIGLR, NOAA National Centers for Coastal Ocean Science (NCCOS), and MBARI have developed mobile HAB detection technology that can be integrated in autonomous underwater vehicles or autonomous surface vehicles, and some preliminary testing has been conducted in Lake Erie.
- weekly field sampling at eight stations in western Lake Erie as indicated in Figure 2.2. Four of those stations correspond to the NOAA GLERL buoys indicated in Table 2.1 that are gathering water quality data every 15 minutes. Field sampling includes temperature, Secchi disk transparency, algal parameters (chlorophyll, phycocyanin, phytoplankton abundance, toxin-producing cyanobacterial populations), algal toxins (microcystins and saxitoxins), and nutrients (phosphorus, and nitrogen). Laboratory genetic tests are also conducted to detect toxic populations of HABs and investigate the relationship between water temperature, nutrients, and HABs.

Additional field samples are collected by WTPs weekly. Other organizations like Ohio EPA, Ohio State University-Stone Laboratory, University of Toledo, and USGS also collect monthly-to-weekly field samples.

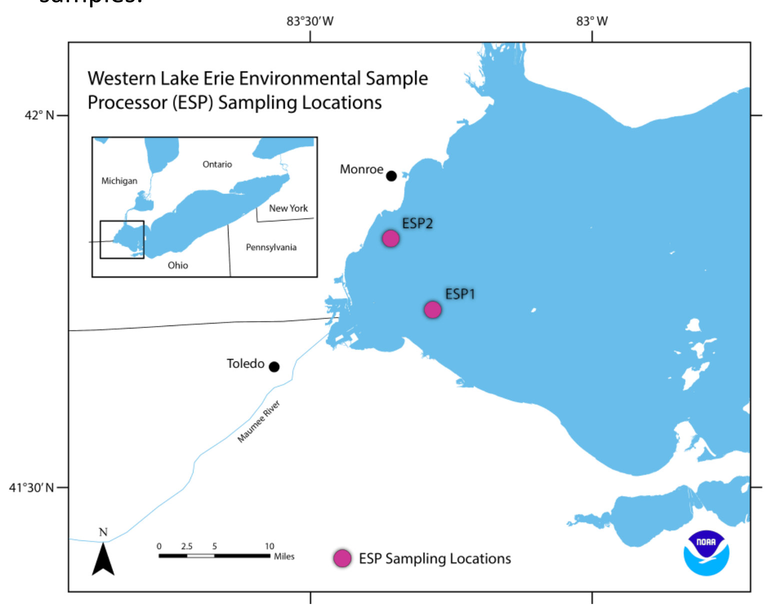
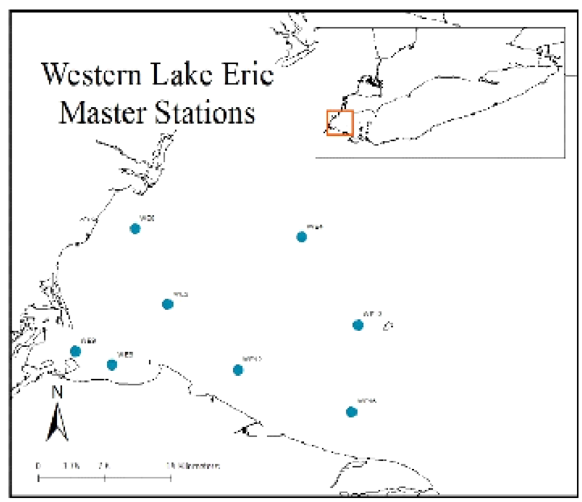


Figure .: NOAA GLERL deployment locations for the Environmental Sample Processors (ESPs). Figure courtesy of NOAA GLERL.



2.1.3 Forecasting

HABs location, size, duration, and toxicity are critical to understand the impact of a bloom and to plan accordingly. The network of western Lake Erie monitoring buoys, shore-based sondes, field sampling, and ESPs, as well as data from remote sensing and hyperspectral HAB imagery have advanced NOAA's operational HAB forecasting system. For the duration of this project the forecast included:

- a seasonal HAB forecast of bloom severity that NOAA issued in July,
- a five-day forecast of bloom location and movement updated twice weekly from July through October, and
- a summary of current conditions that provided main findings based on the analysis of HAB imagery, field samples, and forecasted information of bloom travel both horizontally and vertically.

2.2 OTT project contributions to the network

Network geographical footprint expansion

Water treatment plants in the western and central Lake Erie basins were interviewed by LTI as part of the OTT project to determine 1) the magnitude of the HABs impacts at each location, 2) each plant's methods for dealing with algae, and 3) what additional equipment or information could be beneficial to each plant.

The original Lake Erie HABs and hypoxia observing network included among other assets, instrumented buoys and/or shore-based sondes owned by water treatment plants located in nine municipalities that treat water and subsequently distribute clean drinking water through an extensive network that serves a population of over two million people according to Ohio EPA records as of 7/16/2020 (http://dww.epa.ohio.gov/). The IOOS-OTT funded project added five WTPs to the network. These utilities are located in the western basin of Lake Erie between Toledo and Cleveland. One of these utilities is on the east side of South Bass Island for the city of Put-in-Bay. According to Ohio EPA records from 7/16/2020, these additional WTPs serve over 240,000 people.

Table 2.3 includes a break-down of the cities and drinking water entities along with their corresponding population. The cells in green denote the WTPs that were part of the original Lake Erie HABs and hypoxia monitoring network. The locations of all observing assets corresponding to the Lake Erie Harmful Algal Bloom Early Warning System are shown in Figure 2.3.

Upgrades to the core existing network

During 2019 and 2020, an assessment of the status of the instrumentation, telemetry, and cellular provider costs for each of the WTP observing assets that form part of the EWS network was conducted. Upgrades were proposed by LTI and a summary of those performed is included in Table 2.4.
For the most part, the WTPs operate the same monitoring equipment. This strategy simplified logistics and costs related to equipment upkeep and maintenance throughout the length of this project. All stations have a multiparameter YSI EXO2 sonde with individual sensors to measure some, or all, of the following parameters: temperature, conductivity, pH, oxidation reduction potential, turbidity, chlorophyll, phycocyanin, and dissolved oxygen. All sondes have a central wiper to minimize biofouling.

Besides the sondes, the monitoring systems deployed at each WTP also include a datalogger, modem, and antenna. The data loggers were upgraded to ensure compatibility with the GLOS IT platform data ingestion protocols, and modems were upgraded to 4G. Overall, the current network is mostly composed of Campbell Scientific CR6 or CR1000x data loggers, and Sierra Wireless 4G LTE modems.

Incorporation of new, and more advanced technology

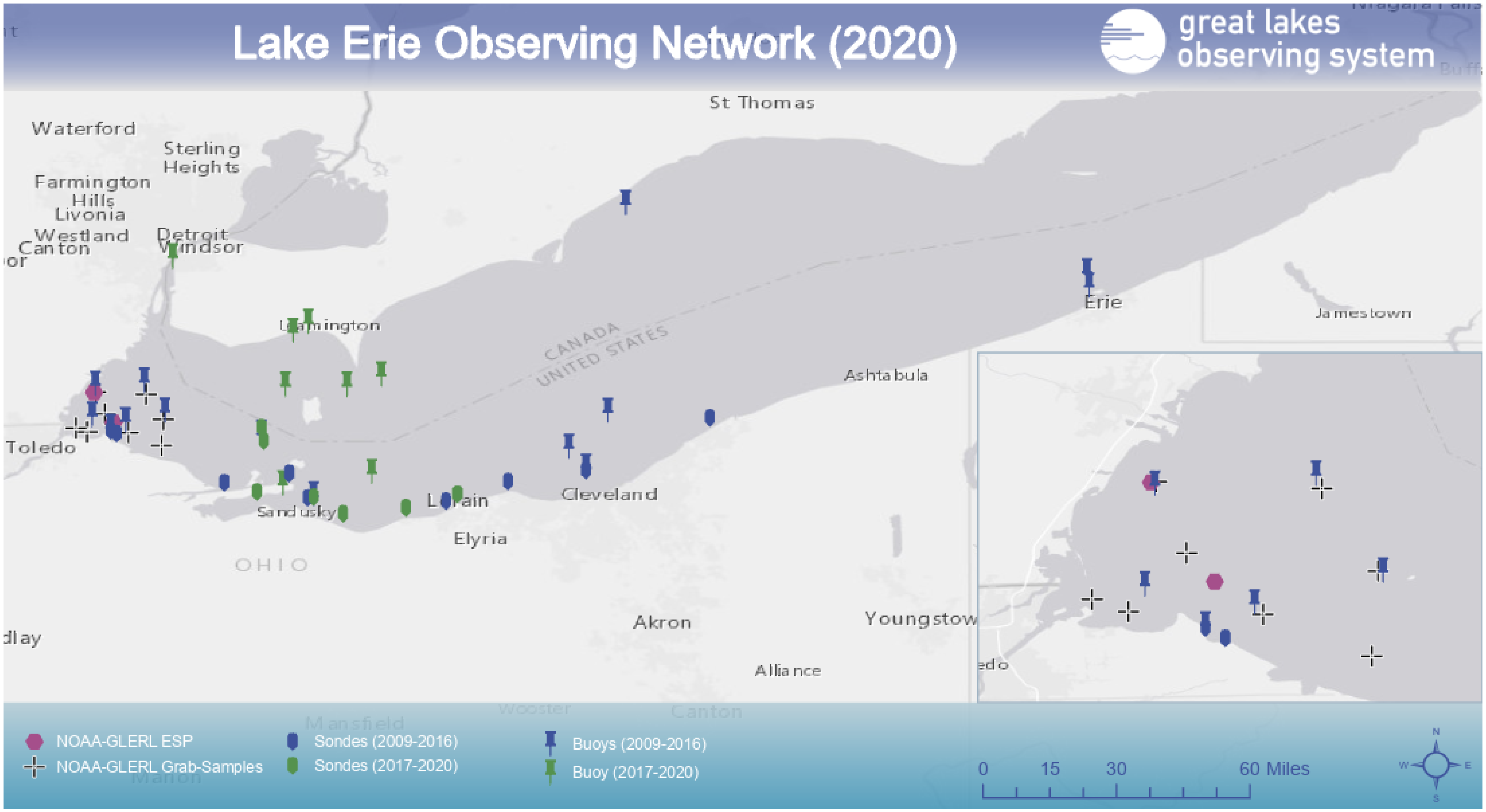
With funds from the EPA Great Lakes Research Initiative (GLRI) and NOAA, GLERL deployed the first 2G ESP in 2017 near the Toledo Water Intake, and a second one at GLERL's WE8 real-time monitoring station (Figure 2.2). The ESPs are deployed on a custom-built mooring assembly that allows the instrument to be deployed subsurface while obtaining near real-time samples from the surface and bottom to assess the toxicity distribution throughout the water column. The annual calibration, maintenance, and testing of the ESPs prior to deployment is conducted by GLERL.

Currently GLERL operates three 2G ESPs in Lake Erie. The third system enables the samplers to be swapped as needed, to allow for the continuous deployment of two 2G ESPs throughout the field season at the two locations indicated above. The addition of these systems is helping Lake Erie HABs observations become increasingly automated with the capability to measure HABs toxicity levels in near real-time. Microcystin data from the deployments conducted by GLERL so far is being transitioned to the GLOS data portal and is intended to be integrated with the HABs EWS in 2021.

Implementation of common O&M protocols for all water treatment plants in the network

The importance of doing the proper maintenance and upkeep for the monitoring equipment to ensure reliability of the data gathered cannot be overemphasized. Following manufacturer recommendations for calibration and maintenance is important, but consideration also needs to be given to the extent of use of the equipment and the environmental conditions they are deployed in, because those could trigger the need for more frequent calibration and maintenance.

From 2017 through 2020, the installation, maintenance, and calibration of the WTPs equipment was done by the same entity, LTI. This ensured that common protocols were used, and calibrations done with the same frequency. Except for UWSS' and Put-in-Bay's monitoring systems that were installed on buoys, all other systems were shore-based, which reduced costs and simplified maintenance. As an outcome of this project, LTI developed a document titled "Guidance for Owning and Maintaining Water Quality Sondes" based on lessons learned throughout the duration of this project. The document is included in Appendix 1. Biofouling was perhaps the most common issue encountered throughout the field season. Some guidelines on how to deal with this problem can be found in Appendix 1.



Lake Erie Observing Network (2020)
great lakes observing system
St Thomas
Waterford
Sterling Heights
Farmington Hills
Livonia
Westland
Canton
Detroit
Windsor
Leamington
Erie
Jamestown
Ashtabula
Toledo
Cleveland
Lorain
Sandusky
Elyria
OHIO
Akron
Youngstown
Alliance
CANADA
UNITED STATES
NOAA-GLERL ESP
NOAA-GLERL Grab-Samples
Sondes (2009-2016)
Sondes (2017-2020)
Buoys (2009-2016)
Buoy (2017-2020)
0 15 30 60 Miles

.

*Table STYLEREF 1 \s 2. SEQ Table * ARABIC \s 1 3: Water Treatment Plants associated with the Lake Erie HABs Observing Network and the population of the cities they serve.*

Locality	Population Served	Public Water Supply System
City of Cleveland	1,308,955	Cleveland Water
City of Toledo	480,000	Toledo Public Utilities
City of Sandusky	81,403	Sandusky Department of Public Works
Union Water Supply System, ON	75,000	Union Water Supply System, ON
City of Mentor	74,500	Aqua Ohio
City of Elyria	68,000	Elyria Public Utilities Department
City of Lorain	64,152	Lorain Utilities Department
Ashtabula	39,838	Aqua Ohio
City of Oregon	33,518	Oregon Water Treatment
Ottawa County	28,019	Ottawa Counnty Regional Water District
City of Avon Lake	23,659	Avon Lake Regional Water
City of Vermilion	10,569	Vermilion Utilities Department
City of Huron	6,893	Huron Water Division
Caroll Township	3,700	Caroll Water & Sewer District
Village of Marblehead	903	Marblehead Water Department
Village of Put-In-Bay	762	Put-in-Bay Utilities Department
Middle Bass	400	Lake Erie Utilities Company
Total Population	**2,300,271**	

Table 2.4: Lake Erie harmful algal bloom monitoring network equipment upgrades made as part of the IOOS Ocean Technology Transition project.

Owner / Operator	Platform type	Station ID	Start Year	Equipment			
				Datalogger		Modem	
				Original	Upgrades	Original	Upgrades
City of Oregon, OH	Sonde	LEORGN	2015	Campbell Scientific CR1000	Campbell Scientific CR6	Sierra Raven XT (2G)	Sierra Raven RV50 (4G)
Ottawa County Regional Water Treatment Plant, OH	Sonde	LEOC	2015	Campbell Scientific CR1000	Campbell Scientific CR6	NA	NA
City of Avon Lake, OH	Sonde	LEAVON	2015	Campbell Scientific CR1000	Campbell Scientific CR6	Sierra Raven XT (2G)	Sierra Raven RV50 (4G)
Aqua Ohio - City of Ashtabula, OH	Sonde	LEASH	2015	Campbell Scientific CR1000	Campbell Scientific CR1000X	Sierra Raven RV50 (4G)	NA
Aqua Ohio - City of Mentor, OH	Sonde	LEMENTOR	2015	Campbell Scientific CR1000	Campbell Scientific CR1000X	Sierra Raven RV50 (4G)	NA
Village of Marblehead, OH	Sonde	LEMRBHD	2015	Campbell Scientific CR1000	Campbell Scientific CR6	Sierra Raven XT (2G)	Sierra Raven RV50 (4G)
City of Elyria, OH	Sonde	LEELYRIA	2015	Campbell Scientific CR1000	Campbell Scientific CR6	Sierra Raven XT (2G)	Sierra Raven RV50 (4G)
City of Toledo, OH	Moored buoy	45165	2009	Campbell Scientific CR1000	NA	Sierra Raven RV50 (4G)	NA
City of Toledo, OH	Moored buoy	TOLCRIB	2015	Campbell Scientific CR1000	NA	Sierra Raven RV50 (4G)	NA
City of Toledo, OH	Sonde	TOLLSPS	2015	Campbell Scientific CR1000	NA	Sierra Raven XT (2G)	Sierra Raven RV50 (4G)
City of Lorain, OH	Sonde	LELORAIN	2018	Campbell Scientific CR1000	Campbell Scientific CR6	Sierra Raven RV50 (4G)	NA
City of Sandusky, OH	Sonde	LEBIWW	2018	Campbell Scientific CR6	Campbell Scientific CR6	Sierra Raven RV50 (4G)	NA
City of Huron, OH	Sonde	LEHURON	2019	Campbell Scientific CR6	NA	Sierra Raven RV50 (4G)	NA
City of Vermilion, OH	Sonde	LEVERM	2019	Campbell Scientific CR6	NA	Sierra Raven RV50 (4G)	NA
Carrol Township, OH	Sonde	LECARR	2019	Campbell Scientific CR6	NA	Sierra Raven RV50 (4G)	NA
Village of Put-In-Bay, OH	Sonde	SBIPIB	2019	Campbell Scientific CR6	NA	Sierra Raven RV50 (4G)	NA
Middle Bass, OH*	Sonde		2020	NexSens X2	NA		NA

2.3 Annual cycle of the WTPs observation network

Throughout the OTT project, the installation, maintenance, and calibration of the WTPs monitoring equipment was done by LTI. As part of the OTT project, GLOS purchased an extra YSI EXO2 sonde with additional sensors to increase resilience of the system. This allowed for the extra components to be swapped as needed in regularly scheduled maintenance visits, so that the network could be operational all throughout the HABs season without interruptions.

A summary of the WTPs monitoring assets maintenance schedule can be found in Appendix 1. The process followed during the OTT project is outlined:

A. Spring calibration and deployment of sondes and buoys
    i. Reconditioning and calibration of sondes was conducted in spring during a calibration event sponsored by GLOS and conducted by Fondriest Environmental Inc.
        1. Participating sonde owners within the Lake Erie HABs EWS were the utilities shown in Tables 2.1 and 2.2, as well as GLERL, University of Toledo, Bowling Green State University, and Ohio State University – Stone Lab.
        2. Utilities with qualified technical personnel in staff did not attend and conducted their own assessments.
    ii. Sonde deployments and Put-in-Bay buoy preparation and deployment were also planned and conducted at this time.
B. Summer maintenance
    i. Approximately monthly visits were performed during the field season, starting as soon as the sondes were deployed, to clean and re-calibrate the sensors. Additional details can be found in Appendix 1.
    ii. Coordinated visits with the purpose of conducting repairs, communications troubleshooting, equipment inspection, tune-ups, or upgrades, were also conducted as needed.
C. Fall maintenance
    i. Buoys and sondes that needed to be removed at the end of the field season were collected by LTI for clean-up, maintenance, and storage.
D. Winter calibration, repairs, and planning
    i. A winter calibration event, like the spring event, was organized. The event was sponsored by GLOS and conducted by Fondriest Environmental, Inc.
    ii. LTI provided technical support for troubleshooting.

2.4 Threats to the current network

The success of the Lake Erie HABs EWS requires the long-term commitment from all interested parties to support, maintain, and coordinate it. To maximize the system's accuracy and effectiveness, it is important that it includes the best available technology. This requires funding not only to operate and maintain equipment, but to upgrade it. It also requires having trained personnel on staff to handle these tasks and decisions or having the ability to subcontract the work outside the organization.

The operation of the network's water monitoring equipment requires that organizations cover, either individually or as a collective, costs for equipment maintenance, replacement, potential upgrades, and communication charges. To operate as a network also implies that minimum equipment requirements must be met to ensure overall data and information accuracy, communications reliability, and operational simplicity.

The pace at which water monitoring technology evolves can translate rather quickly into new equipment that is less complex to operate, requires less frequent calibrations, is more accurate, and becomes easier to maintain. Similarly, innovation extends to telemetry with the speed of data transmission changing quickly, new developments in internet of things (IoT) technologies, and availability of fee-free wireless sensor networks. Some of these improvements can in some cases result in lower operation and maintenance (O&M) costs but require planning, equipment onboarding, and capital investment in upgrades. Keeping up with new innovative technologies is an area that might require investing in the expertise of external organizations to strategically decide what to invest in and when to do so in a manner that is coordinated across the network.

It is also important to note that each WTP is unique in terms of staffing, operation processes, and the environmental impacts they deal with due to their location along the Lake Erie shoreline. Some WTP locations are more prone to hypoxia, while others might be more prone to HAB impacts. This might translate into different monitoring needs (e.g., type of sensors, monitoring platforms, or platform location), and WTPs' willingness to invest in the monitoring network. Similarly, some utilities have limited technical capacity with a small number of staff, while others can potentially allocate or train personnel to oversee the performance of the monitoring equipment, oversee its maintenance in accordance with manufacturer requirements, and conduct data analysis.

Providing early warning of impending events on an accurate and timely basis is the goal of the Lake Erie HABs EWS. Minimizing potential false positive warnings via continuous optimization and ground-truthing of the system is key. Data compatibility and minimal duplicity with other HAB related data that the WTP might collect to satisfy local/state requirements is also important to ensure long-term use.
Having inconsistent, short-term sources of funding can negatively impact the sustainability of the network. Coordinated efforts to improve, maintain, and sustain the network would be more enticing, especially to smaller WTPs. This could perhaps be accomplished by relying on one or various third parties to coordinate funding, planning for upkeep and onboarding of new instrumentation, as well as expanding the network and demonstrating its value to stakeholders.

2.5 Recommendations for network preservation

The long-term value of the network is dependent on the number of participating WTPs as well as potentially other users, but also on the adequate maintenance of the network to ensure data reliability and high confidence level on the information it provides. To that end, funding with a focus on the network's long-term sustainability is critical to ensure its success.

Recognizing that the impact of the information gathered by the network and provided by the EWS goes beyond WTPs managers is important, since it could also be leveraged to other interested stakeholders in Lake Erie. The network's value needs to be demonstrated to the WTPs customers, state/local water management officials, lake recreational users, commercial fishermen, etc. to ensure their long-term support, and potentially financial support.

The continuous maintenance and enhancement of the network is constrained by lack of funding and limited coordination between users. Ensuring that either trained in-house personnel or a subcontracted organization(s) conducts regular equipment maintenance and calibration according to manufacturer's requirements, implements adequate QA/QC procedures, evaluates data trends, and investigates data anomalies is key. Having consistent O&M and data quality control processes across the network can help ensure consistent performance across monitoring platforms.

Improving coordination and communication between WTPs is also important. Crafting a coordinated response plan as part of the Lake Erie EWS observing network could be of interest instead of leaving this to individual WTP managers. A response plan could document the connection between trend spikes, unusual sensor readings, parameter correlations, etc. and the adequate response and communication protocols among WTPs.

It is important that the Lake Erie EWS can ensure that information provided from key datasets is done in a manner that is simple, comprehensive, and, more importantly, easy to understand and act on. Finding a way to incorporate results from field testing conducted by the WTPs (e.g., ELISA testing for microcystin toxin detection) into the network's EWS as an additional data stream, and a way for the utilities to easily correlate all relevant data could be of interest. Also, WTP managers should be able to easily review and evaluate interannual changes in source water quality as those comparisons could be helpful in data assessment. Having the ability to continue to evolve the information provided by the EWS based on ongoing WTPs managers input would be of great value.

Predictive data analytic tools can potentially translate into more cost-effective plant operations. Additional improvements to the Lake Erie EWS could come from 1) optimizing the monitoring network by deploying assets at key locations and as needed making use of more advanced observing technologies, 2) improving the spatial and temporal resolution of microcystin toxin concentration measurements, 3) optimizing data collection to address issues specific to each utility, and 4) assimilating data into numerical models to produce more accurate forecasts and intake specific forecasts.

A reliable, sustainable, and validated Lake Erie HABs EWS could potentially be extended to other areas of the Great Lakes also impacted by HABs. This regional effort can ultimately be an integral component of NOAA's strategy to implement a National HAB Observing Network to adequately sustain and integrate HAB capabilities and deliver needed operational tools nationwide.

3. The EWS Applied: Data, Access, & Management

3.1. EWS Digital System Overview (Tier 1)

The EWS is underpinned by GLOS' technology platform, called Seagull, that serves to connect streaming sensors, numerical models and other data sources on an interconnected network to end users primarily interested in identifying potential HABs in Lake Erie. This technology platform aims to have data and information that is discoverable, accessible, available and reusable for organizations and individuals interested in predicting HABs. This information platform is responsible for the entire data management lifecycle after data acquisition and transmission, from ingestion into the cyberinfrastructure, quality control, public dissemination, product generation, and long-term storage and archival.

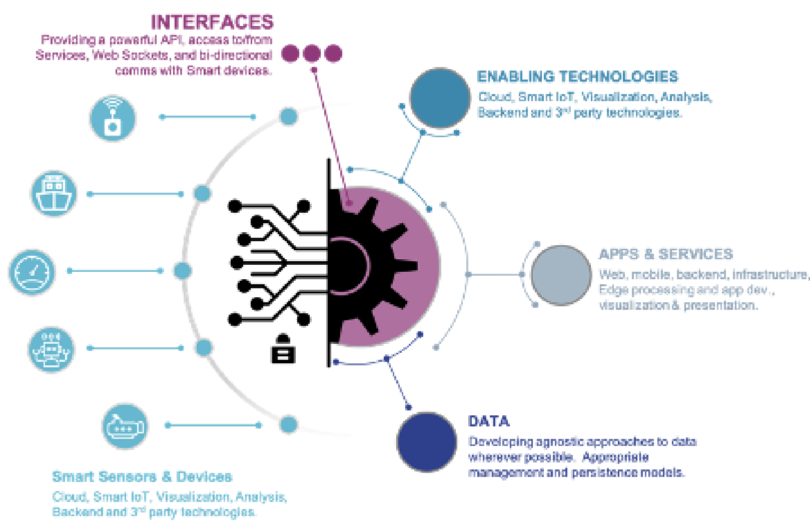


Figure 3.1: Overall system diagram of the Information Technology platform supporting the Early Warning System for Lake Erie harmful algal blooms.

Major components of the information technology platform include a data ingestion pipeline, event management services, application services layer, gateways, monitoring and triage and overall platform services (Figure 3.1). These categories mostly comprise the backend, or behind the scenes, components. Frontend, or user interface, components include maps, dashboards and other information services. The entire system is underpinned by a security model to protect the integrity of the data, the users and the flow of information.

The GLOS IT platform handles streaming, periodic, harvested, or uploaded real-time data, and delayed mode observing data from moored platforms, shore-based stations, and field sampling stations, as well as data from numerical models.

3.2. EWS Infrastructure Overview (Tier 1)

The EWS platform is developed on and using Cloud services. Amazon Web Services (AWS) is the largest current provider of this type of technology and GLOS has adopted AWS for a variety of reasons, including ease of use and interoperability with other Cloud services and the near perfect record of 'up-time', which is critical for systems requiring monitoring and alerting such as the HABS EWS.

Cloud services have been widely adopted by a wide range of verticals, including finance, travel, storage, defense, commerce, and more. Most online services now utilize some form of Cloud services. Cloud services can be thought of as 'building blocks' that developers and administrators can use to build applications, platforms and other services for a wide range of activities. These can include computing power, storage, databases, migration between applications, network and content delivery, management tools, security & encryption, messaging, metadata and identity management (Figure 3.2).

There are multiple resources online regarding cloud computing and services. This document will not delve into descriptive detail but will highlight some of the core AWS Cloud services used in this HABS EWS. Additional details about the components involved in the architecture for this EWS can be found in Appendix 2 (EWS Architecture Components).

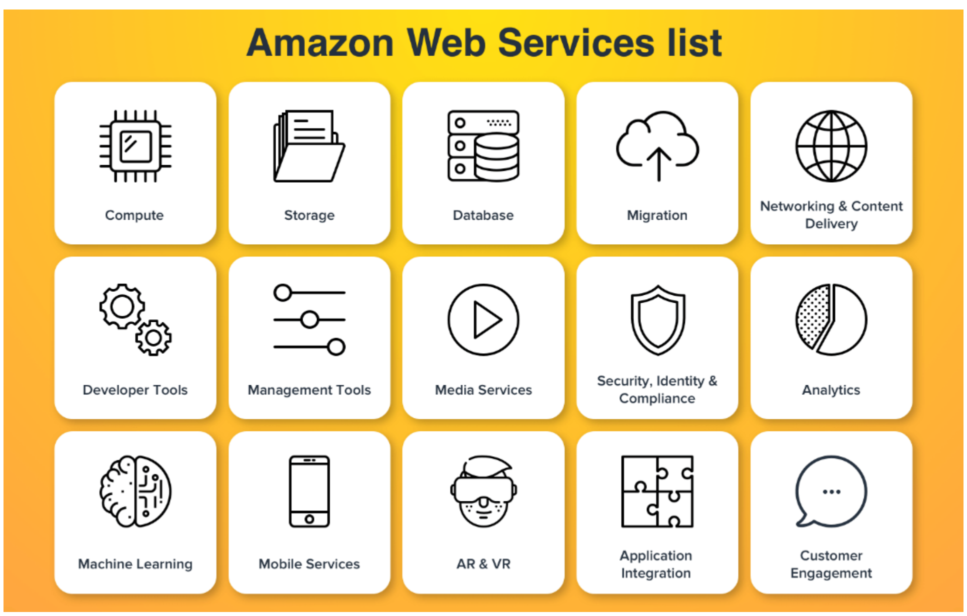


Figure 3.2: An Example of some of the common components and services available within a Cloud environment.

Within these broad categories highlighted above, AWS has individual frameworks, components, and services to assist developers in creating applications, functions, and additional services. At the time of publication, a current list of these services is indicated in Figure 3.3.

Figure 3.3: There are dozens of AWS Services to support a wide range of application and platform development.

The GLOS information technology leverages many of the services highlighted above, including an API gateway, event streaming, data management, data storage, service logic deployment, user pool management and more. As the platform continues to evolve and grow, additional AWS services will be leveraged to support the EWS.

A further description of the AWS services leveraged by GLOS in support of the HAB EWS are described in detail in Appendix 2.

3.3. Data Flow Overview

At a high level, data tunneling for measurements, modeled, or sampled data into the EWS all follow the same pathway. There is ingestion, processing, storage, and conversion to information, triggering of alerts and notifications, visualization, and dissemination.

3.3.1. Data Ingress

The data categories (i.e. real-time, near real-time, and delayed mode) and flow help comprise the core components of the EWS. Without data ingress to the EWS, there is no streaming flow of the data to information lifecycle. The categories of data and flow of these data described above is not finite and can evolve and scale with platform evolution and a changing technical landscape. As observing technology gets smarter and other collection pathways evolve, such as citizen science, so too will the EWS. The structure for our Early Warning System is being built to be adaptable into the future to continue providing the best possible information for our stakeholders.

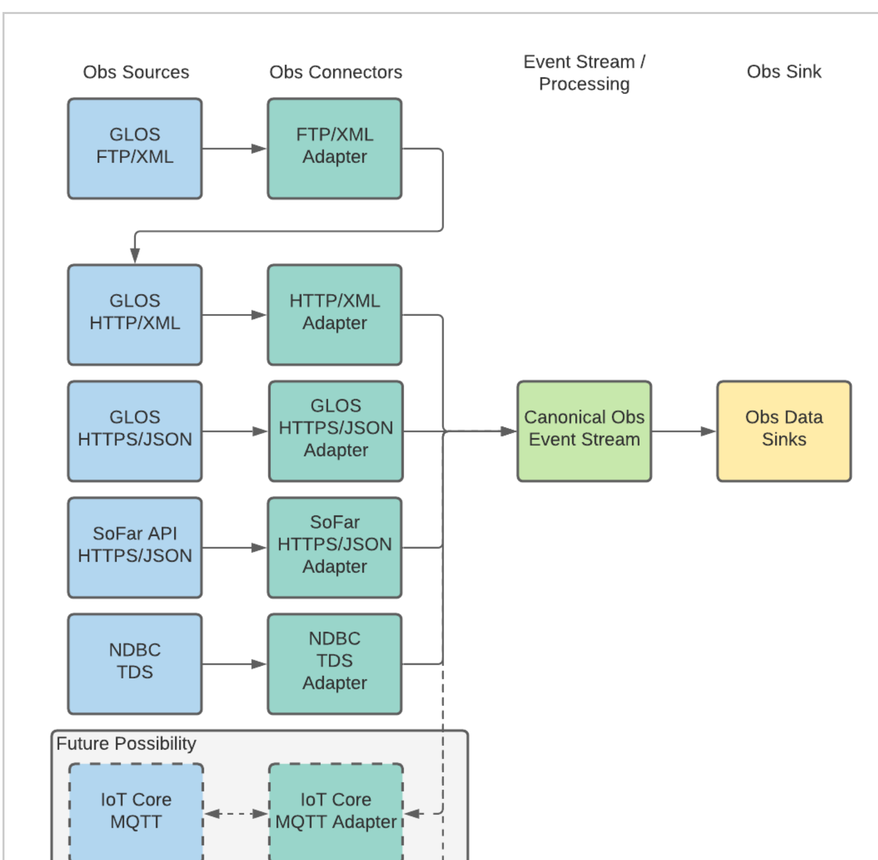
Real- and Near Real-time Data Ingress

Data Transmission/Receipt

The (near) real-time data ingress pipeline into the EWS also includes other observing data updated only daily or weekly, such as ESP or field sampling data. Data Ingress interfaces are composed of several protocols and matching adapters. There are two main ways of submitting observation data to the Seagull/GLOS EWS platform:

1. Data Loggers: Sensors are connected to data loggers that transmit data via a modem to the Seagull system or an external system.
2. Observational Server: The Observation Local Server includes software such as iChart/Sofar/WQDataLive that aggregates data from the data logger and forwards this data to GLOS. This also includes any customer scripts written by customers that send observational data to GLOS. This software typically runs on a desktop computer or server at the data provider's location.

Figure 3.4 describes how the observational data interfaces relate to the adapters and the rest of the observational data ingress pipeline.



Observing and Model Data Ingest Adapter

This component is a composition of multiple adapters that translate information flowing between the observational data sources and the canonical model used by the Seagull platform. Observational data sources utilize a variety of protocols and data formats. The Observation Data Ingest Adapter is responsible for converting the various data streams into a protocol and format that is suitable for integration with the Event Stream. These adapters support bidirectional communication where appropriate.

A few examples of the adapters are:

- GLOS Observational Data HTTP/JSON Application Programming Interface (API)
  Allows observational data to be ingested into the Seagull System. The request is authorized by an API key that is associated with the specific reporting device when the device is provisioned by the data provider in Seagull. See the detailed documentation of the API here.
- Third Party Clouds
  The EWS platform also allows the observational data to be pushed from third party clouds to a GLOS Observational API delivery endpoint.
- NDBC Consumer
  The NDBC consumer interface is a TDS interface that is used by the Observational Data Ingest Adapter to periodically poll buoy data from NDBC and push it into Seagull.
- Future Observational Interfaces
  Future Observational Interfaces may include protocols such as MQTT, AMQP, or FTP, and file formats such as Protocol Buffers, Parquet, or Avro.

Delayed Mode Data Ingress

The EWS Model Data ETL[4] service extracts data from NOAA's Great Lakes Operational Forecast Model System (GLOFS) which is an automated, numerical, model-generated current (nowcast) and future (forecast) prediction of physical water conditions, including water levels, currents, and temperature.

Separate numerical model runs are done for each lake at slightly different times. All lake models use the same hydrodynamic numerical algorithms but the grid characteristics change for each lake. The output from GLOFS includes graphical products and NetCDF files available through a TDS server or public S3 bucket. Seagull periodically polls the S3 bucket looking for newly added NetCDF model files. The files are downloaded via HTTP.

3.3.2. Data Processing

[4] ETL: Extract, Transform, and Load. In computing, extract, transform, load is the general procedure of copying data from one or more sources into a destination system which represents the data differently from the source or in a different context than the source. https://en.wikipedia.org/wiki/Extract%2C_transform%2C_load

The primary system diagram (Figure 3.5) represents the core design of Seagull and the EWS system. A majority of the components affecting the Seagull system are described in the diagrams that follow (Figures 3.6-3.8).

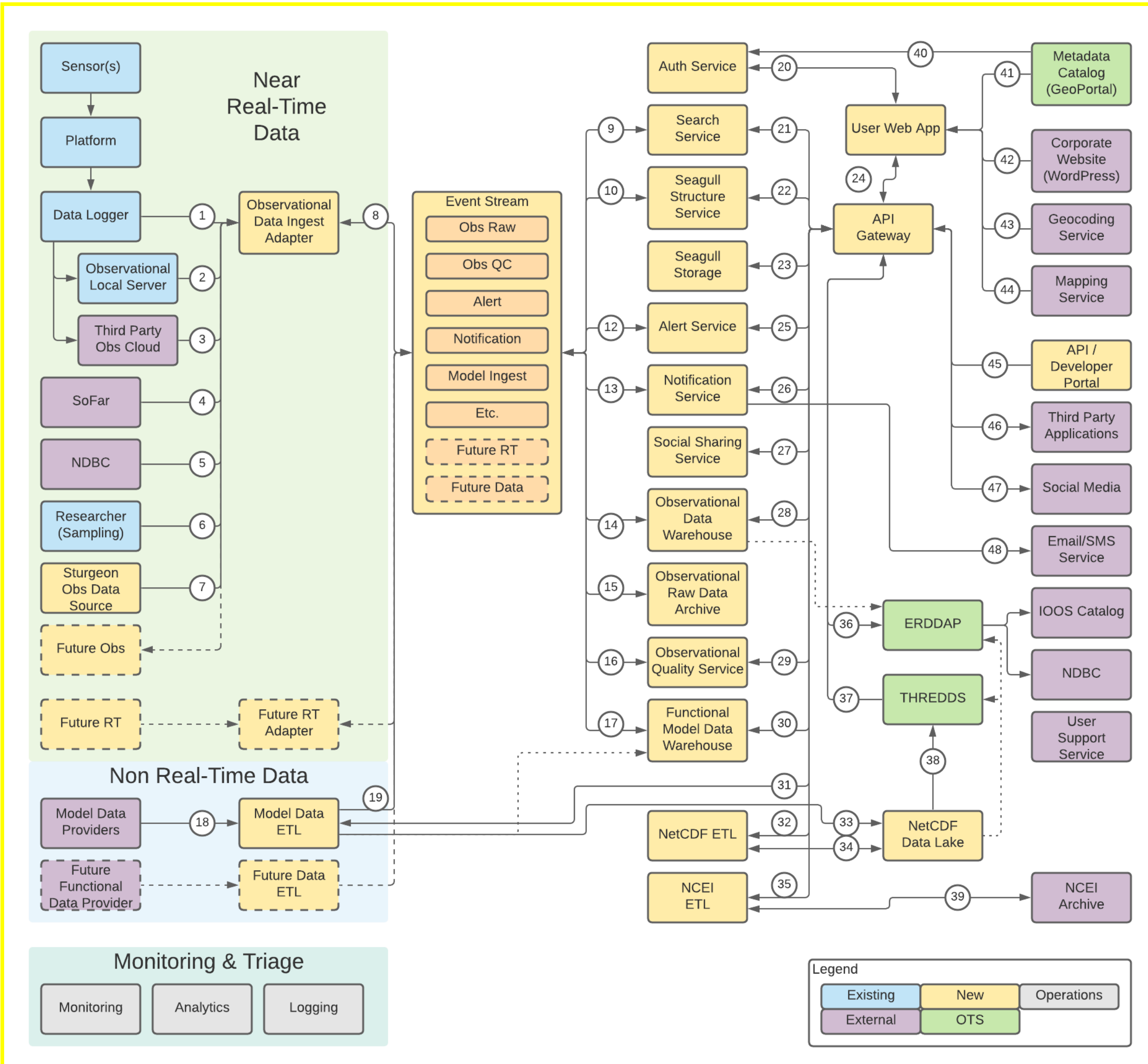


Figure 3.5: An overview system diagram of the core design for Seagull, the EWS parent platform and main GLOS IT platform.

Real- and Near Real-time Observing Data processing

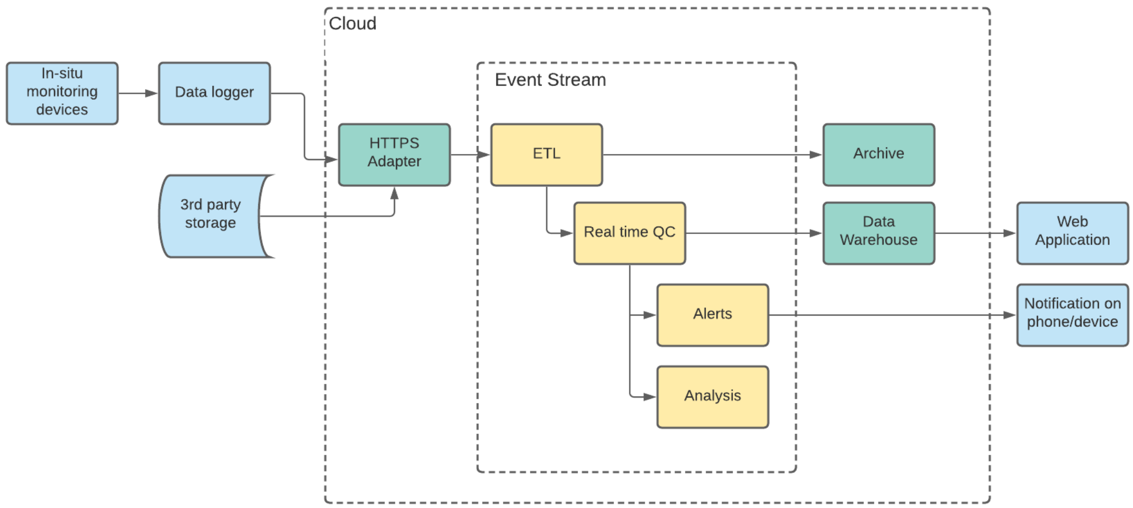


Figure 3.6: The high-level flow of data from platform devices to the cloud to event streams and beyond.

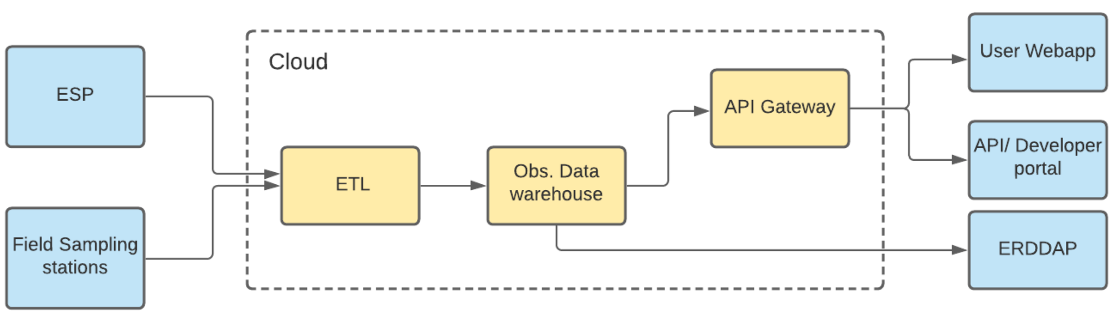
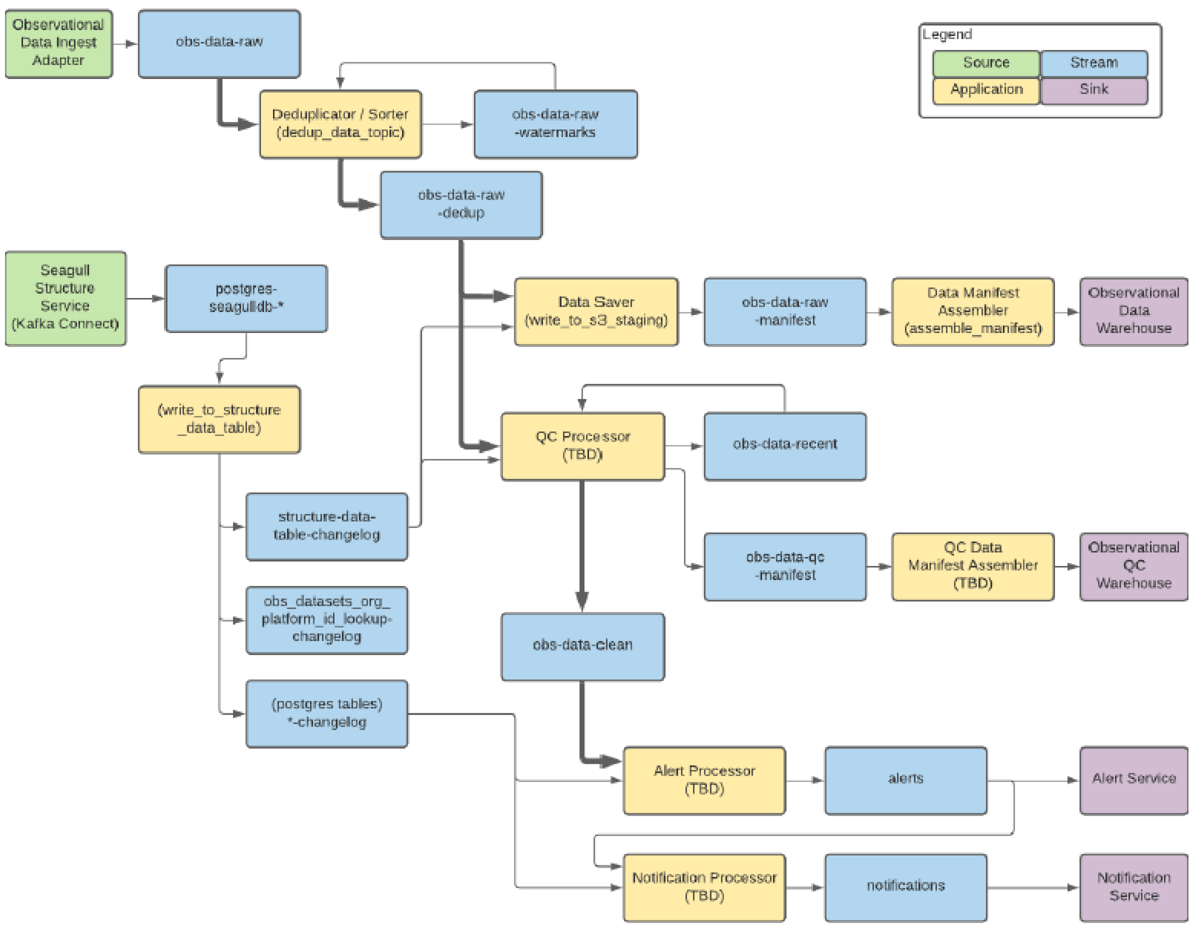


Figure 3.7: Typical flow of grab sample data in the Lake Erie early warning system operated by GLOS.

Event Streams Overview

The GLOS EWS takes significant advantage of "event stream management" as part of the AWS cloud offering. These are described in detailed in Appendix 2. An event-driven architecture uses events to trigger and communicate between decoupled services and is common in modern applications built with microservices such as the GLOS EWS. An event is a change in state, or an update, like ingestion of new data, or a data value exceeding a predefined threshold. Events can either carry the state (the data value, the source, the category of event) or events can be identifiers (a notification that a data value was received).



Event-driven architectures have three key components: event producers, event routers, and event consumers. A producer publishes an event to the router, which filters and pushes the events to consumers. Producer services and consumer services are decoupled, which allows them to be scaled, updated, and deployed independently.

The EWS's Event Stream components' primary responsibilities are the ingestion and processing of observational data. Processing of observational data by the event stream includes the following:

- Associating the observational data with matching metadata;
- Validating the ingested data against GLOS metadata requirements;
- Validating the ingested data against metadata;
- Evaluating QARTOD[5] quality checks;
- Evaluating relevant real-time alert conditions;
- Generating resulting processed data streams;
- Generating resulting event streams. These event streams can be used as other inputs into embedded, linked, or connected systems.

The Event Stream may also be used to ingest and process functional model data when necessary. Ingested model data is processed in the following ways:

- Evaluating relevant real-time alert conditions, and
- Generating resulting event streams.

The Event Stream consumes configuration and changes that take place in the Back-end Structure Service, Metadata Service, Alert Service, and Notification Service.

The Event Stream produces messages to be consumed by the Notification Service, Observational Data Warehouse, Functional Model Data Warehouse, and Observational Data Archive. While the Event Stream itself does not perform the processing, the components consuming from the Event Stream perform the processing and publish their results back to the Event Stream.

EWS Observational Quality Control Processor

This service has the responsibility of evaluating the quality of observational data supported by the EWS. The incoming observational data are evaluated against QARTOD quality checks. The various quality checks are described below in Table 3.1. This service enriches the stream of incoming observational data with the quality check results and publishes the enriched stream of data back to the Event Stream.

Table 3.1: Quality assurance/quality control of real-time oceanographic data (QARTOD) test names and their description.

QARTOD Test	Description
Flat Line	A check to ensure that the incoming data is varying, and is otherwise not endlessly repeating any particular value.
Gross Range	Ensures that the data point does not exceed various minimum or maximum values.

[5] QARTOD refers to the Quality Assurance/Quality Control of Real Time Oceanographic Data. QARTOD tests are developed by the IOOS Program Office for several commonly measured parameters.

Rate of Change	Tests the data point against the maximum allowable rate of change.
Spike	Checks if a data point exceeds a threshold relative to adjacent data point.
Primary	An overall rating for the value

Table 3.2: QARTOD quality check result flags are summarized in the following table.

Value	**Meaning**	**Description**
1	Pass / Good	Data has passed critical real-time quality control tests and are deemed adequate for use as preliminary data.
2	Not evaluated, not available, or unknown	Data has not been QC-tested or the information on quality is not available.
3	Questionable / Suspect	Data are considered to be either suspect or of high interest to data providers and users. They are flagged as suspects to draw further attention to them by operators.
4	Fail / Bad	Data are considered to have failed one or more critical real-time QC checks. If they are disseminated at all, it should be readily apparent that they are not of acceptable quality.
9	Missing Data	Used as a placeholder when data is missing.

EWS Notification Service

The EWS Notification Service manages notification configuration for the system. Notification management includes the following:

- Who should be notified?
- What should they be notified about?
- How should they be notified?
- How often should they be notified?
- Notification lifecycle tracking

EWS Observational Alerting Service

The primary goal of the Alert Service is to manage the real-time and delayed mode alert configurations throughout the system. The alert service includes an API for managing the alerts. It also pushes real-time alert configuration into the Event Stream so the event stream can process alerts. In addition, the service is responsible for evaluating non-real-time alert conditions and pushing the results back into the Event Stream. The alert conditions are evaluated by the Notification Service for determining who should be notified. Table 3.3 describes the types of alerts that the system is capable of evaluating.

Table 3.3: The Early Warning System has various types of alerts available to different system users. The alerts, their accessibility, and their descriptions are included here.

Alert	**Who**	**Description**
Threshold	Anyone	A parameter value exceeds a user defined threshold (e.g. water temperature is above 70F)
Rate of change	Anyone	A parameter values rate of change exceeds an absolute maximum value. (e.g. air pressure has dropped more than 100 mbar in the last hour)
Obs Data stream: start	Owner, GLOS Observing (Obs) Team (unless data is private)	An alert for the first successful transmission of data following the deployment of an observational device.
Obs Data stream: quality concerns	Owner, GLOS ObsTeam (unless data is private)	A data stream has fallen below quality standards.
Obs Data stream: loss of communication	Owner, GLOS Obs Team (unless data is private)	No data has been received from a device for more than a specified amount of time.
Obs Data stream: Missing Data	Owner, GLOS Obs Team (unless data is private)	Data is being received from the device but particular parameters are missing from the data stream.
Obs Data stream: Excess Data	Owner, GLOS Obs Team (unless data is private)	Data is being received from the device but there is more data being received than the device has been configured to collect.
Obs Data stream: Invalid Data	Owner, GLOS Obs Team (unless data is private)	Data is being received from the device but the message cannot be parsed or the data is invalid.
Moored Buoy: Geofence Violation	Owner, GLOS Obs Team (unless data is private)	Moored Buoy has drifted outside of the allowable geographical tolerance for the deployment site.

Model Data processing

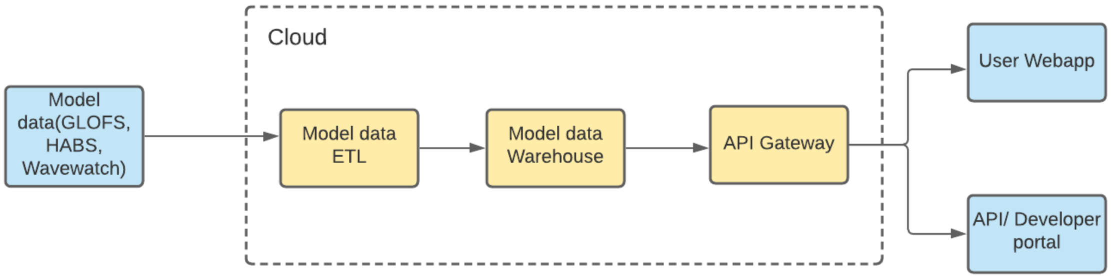


Figure 3.9: Typical flow of model data in the Lake Erie EWS operated by GLOS.

The model data ETL process pulls data from the model data providers, transforms the data as necessary, and loads the data into multiple locations (Figure 3.9).

The service can publish relevant model data to the Event Stream and send model data to the NetCDF Data Lake, from which the data can be accessed via ERDDAP for API, application or programmatic integration with the EWS. Additionally, the process can load data into the Model Data Warehouse. This warehouse is used to record and provide access to model data that is ingested by the model data ETL. The model data in this warehouse is structured in such a way to provide immediate, programmatic access to the model data. The warehouse includes an API for querying data, which allows for integration with the EWS.

3.3.3. Data Storage and Archival

The Observational Data Warehouse is responsible for storing all ingested observational data. It also provides the ability to query all observational data. Access to the warehouse is provided via an API. The warehouse consumes observational data messages from the Event Stream and preserves them. A portion of the warehouse data is available immediately and another portion is available with a short delay.

Information that must be presented immediately includes the latest observations and the last 7 days of observations.

Information that can be presented with a short delay includes:

- Aggregations (hourly, daily, etc.),
- Extended time spans, and
- Any complex queries.

All the observational data is also periodically archived at the National Centers for Environmental Information (NCEI).

The Observational Raw Data Archive is a long-term storage archive for recording the raw ingested observational data. It includes different levels of accessibility for managing storage costs. It is populated by consuming messages from the Event Stream. It also includes the ability to forward raw incoming messages to the API Gateway so they can be viewed in a web user interface.

The Functional Model Data Warehouse is used to record and provide access to model data that is ingested by the Model Data ETL. The model data in this warehouse is structured in such a way to provide immediate, programmatic access to the model data. The warehouse includes an API for querying data.

The NetCDF data lake contains all ingested observational data and all ingested functional model data formatted as NetCDF files. Not currently included in the data lake are datasets uploaded by users. The data lake is the definitive data source for all services that consume NetCDF file data such as ERDDAP.

3.3.4. Data Dissemination

Web Apps

Administrative Web App

The future development of an administrative web application will allow an authorized user access to information that helps facilitate the support and growth of the Seagull Platform. Envisioned users of the admin web application include the GLOS Observing team and the technical support team. The web application consists of dashboards and user interface (UI) elements to present the big picture necessary to support and maintain the platform.

API / Developer Documentation

This web application is utilized by third party and internal developers as a source of API documentation. Detailed information about the various interfaces is included in the development web application.

ERDDAP

ERDDAP is a standard off-the-shelf data server and web application for providing a consistent way to share scientific data in addition to creating maps and graphs. ERDDAP serves observational and model data. ERDDAP is also a requirement for IOOS certification.

User Web App

The User web application is the primary means of interaction between Seagull and users. The User web application is a mixture of public and private UI elements that allow users to interact with Seagull Features. In addition, the user web application also allows authorized user access to information that helps facilitate the support and growth of the Seagull Platform.

The GLOS EWS takes advantage of Seagull, the new information technology platform designed for the ingestion, metadata registration, quality validation, processing, and distribution of a wide range of Great Lakes marine data. Seagull is leveraged by a responsive, React JavaScript web application for rapid and intuitive query, consumption and download of information. There are several components to the user experience of this web application, but of interest to HABs for viewing, discovering and downloading data and information is the platform viewing component. A user is able to navigate a map, which clearly shows platforms, their key indicators of high priority parameters, and explore more content through an intuitive web application (Seagull). Presented in a dashboard style layout, users are able to rapidly visualize multiple parameters that could be HAB indicators, download the supporting data, generate graphs and charts, and set alerts or favorites on selected platforms that are of particular interest to them.

Metadata Catalog (GeoPortal)

The metadata catalog is the central repository for metadata in the Seagull system. It also integrates with the Auth Service to share the Seagull User Pool. The User Web App uses the Metadata Catalog to provide metadata specific search results. Data contributors, platform owners and other authorized users have the ability to add, edit, update and delete metadata about platforms, data and other relevant information. The GLOS EWS search function has the ability to query metadata across the platform and return search results in a variety of applications, further underscoring the platform approach. The metadata catalog functionality is leveraged from a commercially available, free, open-source project from Esri (www.esri.com) called GeoPortal Server.

4. Future Directions

4.1 Monitoring Network

From the time when the Toledo HABs crisis took place in 2014 until now, significant improvements have been made to help provide water treatment plant managers with forecasts and monitoring information about lake changing conditions. To better understand algal bloom concentration, toxicity, location, and trajectory, different efforts are taking place in the Great Lakes region. The goal is to provide advanced warning to WTP managers so they can prepare and respond in a timely and adequate manner. Efforts are being directed to improve both real-time observations and forecasting efforts but also to enhance the IT platform.

The following highlights some of the work that is being conducted or will be done in the near future to expand and improve the monitoring network in Lake Erie and its watershed:

1. NOAA GLERL and CIGLR deployments of the three 2G ESPs in western Lake Erie will take place in 2021 to allow for automated real-time measurements of microcystin levels. This information will complement satellite observations as well as weekly sampling and airborne hyperspectral images collected by these organizations. Plans are also ongoing to test an uncrewed aerial vehicle that incorporates a hyperspectral camera during the 2022 field season. Improvements to the sensors' suite of buoys and other observing platforms deployed by these organizations is continuously ongoing.
2. Cleveland Water planned to expand their buoy network last year and during the 2021 field season two new buoys have been deployed.
3. With GLOS support the following projects will be implemented in 2021-2022:
    - Lake County Department of Utilities, that owns two water treatment plants in central Lake Erie, will be deploying sondes to continuously monitor raw water quality. These water treatment plants serve approximately 117,000 residents.
    - The City of Defiance WTP in the Maumee River will be deploying equipment to continuously monitor raw water quality and nutrient levels.
    - Heidelberg University will be installing telecommunication equipment to provide real-time data from the Heidelberg Tributary Loading Program (HTLP), and additional nutrient and water quality instrumentation at some of the HTLP stations.
    - Wayne State will be transitioning the Huron-Erie (HE) Drinking Water Monitoring Data Platform to GLOS IT platform. The HE monitoring platform was developed by Wayne State University in collaboration with regional watershed groups, WTPs located along the HE corridor, the Great Lakes Water Alliance, and others to ensure drinking water safety to the more than 4 million residents in the HE watershed.
    - Florida Atlantic University in collaboration with MTU, University of Minnesota and NOAA GLERL will be deploying, testing, and validating a novel autonomous holographic imaging

system (AUTOHOLO) with the goal of improving the Great Lakes region's ability to observe, detect, and manage HABs. The testing of the system will take place during the 2022 field season.

- NOAA GLERL will be testing a new autonomous surface vessel (ASV) during the 2021 field season. Plans to equip the ASV with a Monterey Bay Aquarium Research Institute third generation Environmental Sample Processor are ongoing with the goal to test it during the 2022 field season.

In general, more work is needed to be able to measure cyanobacteria concentration and toxicity levels at the needed spatial and temporal scales. Additionally, expanding the region's ability to collect data during the winter months is of critical importance to be able to get a wholesome understanding of ecosystem changes and their triggers. The coordination of hyperspectral and multi-sensor imaging systems with other more advanced platform and sensor networks can provide continuous surveillance capacity to further enhance Lake Erie's early warning system capabilities.

4.2 Forecasting

Ecological forecasts are an important component of early warning systems as they can provide an indication of ecosystem changes with possible impacts on public health and the local economy. As already indicated in Section 2.1.3, NOAA National Centers for Coastal Ocean Science (NCCOS) provides a HABs forecast, normally from July through October, that focuses on cyanobacteria or blue-green algae blooms in Lake Erie. As of 2021 it includes:

- a seasonal outlook,
- a summary of current bloom location and extent, as well as an indication/alert to the presence of scum and/or toxins,
- satellite images of the observed bloom position and extent, and
- a five-day forecast of bloom location (at surface and at depths) and movement, based on satellite imagery, and modeled currents.

Funded by the NCCOS Coastal Hypoxia Research Program, GLERL has developed the experimental Lake Erie hypoxia forecast system. Work is ongoing to advance it to sustained operations so that it can be used by central basin water treatment plant managers. Plans are in place to integrate numerical model outputs from those two forecasting systems into the EWS.

GLERL and other partners are also working to develop probabilistic toxin and hypoxia forecast models. The goal is to provide the probability of 1) exceeding microcystin toxin public health advisory levels in western Lake Erie, and 2) the occurrence of hypoxic events in central Lake Erie. This type of forecasts can provide quantified levels of risk to water treatment plant managers to trigger a timely response. This can be an important early warning system tool.

Validations of the numerical models with observing data from advanced sensor networks capable of providing information regarding algal bloom concentration, toxicity, location, and trajectory, and similar pertinent information for hypoxic events is important. Collecting and using year-round observations to validate numerical models is also critical as is the assimilation of observing data to improve forecasts accuracy. The use of artificial intelligence models based on field data to simulate and anticipate changes could also contribute to the development of better management approaches.

4.3 IT Platform

The short-term directions for the IT platform include some primary new capabilities in terms of ingesting new platforms, expanding and enhancing metadata, search, account managements, customization, alerts, notifications, data download and more. Additional details can be found at www.glos.org/seagull.

Figure 4.1: Planned short-term roadmap for GLOS' Information Technology platform.

Longer term plans for the platform powering the HABS EWS include easier management of data and platform ingestion, expanded API access, integrated search to include external sources, encryption, group and project management.

Underpinning all of this is scalability, security, and stability of the overall system. Given the criticality of HAB data and its importance for fresh drinking water, ensuring the integrity of the data, data flows, and information that reaches end users is of paramount importance. GLOS will develop a robust encryption protocol for the entirety of the lifecycle of the data in conjunction with the federally approved quality assurance utilities (e.g. QARTOD). This will protect the legitimacy of the data with the assurance that others have not had access to influence the integrity of the data during the lifecycle from acquisition through harvest, transmission, and processing.

The EWS will evolve from 'minimum viable product' in its first release through a public beta, then public launch and beyond (Figure 4.1). During this process, users of the system will increase from early adopters to hundreds of thousands of users distributed around the region. Engaging these users with an eye towards a 'multiplier effect' will further promote the awareness and use of this platform. This will be achieved through encouragement of social shareability of data and information from the platform, frequent measuring of a net promoter score, and direct and indirect marketing strategies to promote the platform. This period of scalability will see at first a gradual increase of users and then dramatically increase as confidence and awareness increase. This period of growth is not limited to human users. Platforms, sensors, smart devices and other datasets will also be added to the system further enhancing its appeal by quantity and quality of data availability for predicting HABs in the EWS. This also improves future analytics for machine learning and artificial intelligence.

The backend of the platform will require constant monitoring for system stabilization optimization efficiencies, load balancing, data storage, management of the codebase and requests from users and other systems. Although not necessarily visible to end users, these growth changes on the backend to ensure stability of the overall functioning of the system are paramount for the HABs EWS. Although AWS provides limited toolsets for monitoring the health of the Cloud services, the GLOS development team will provide additional tools to improve and expand the monitoring capacity of the platform.

Beyond the 2-3 year road map, GLOS will explore emerging technologies such as artificial intelligence, augmented intelligence, machine learning and increasing interoperability between Cloud environments. Integrating partner, vendor, research and 3rd party applications, systems and interfaces into the platform will also enhance and further expand the capabilities of the HABS EWS. This will be facilitated by the creation, cultivation and growth of a developer community who can breathe new ideas and expand capacity into the platform. This developer community would leverage the platform (Seagull) Software Development Kit (SDK) which would enable others to enhance, potentially modify, grow and embed platform functionality across a wide range of environments - all with the benefit of improving and enhancing the overall HABs EWS.

5. Financial Sustainability of the System

5.1. Value and cost of the core observing network

The valuation and costs of the core observing network vary based on the operator costs for procurement, maintenance and deployment of the monitoring equipment. Furthermore, costs will vary based on the monitoring needs of specific locations along the Lake Erie near shore area. For example, the city of Cleveland water system is monitoring water quality impacts of hypoxic events versus the city of Toledo's priority of HABs. The following cost ranges are based on information gathered during this project:

1. Equipment:
    a. YSI EXO2 Sonde, Data logger and modem: $15K to $24K depending on equipment configuration and site
    b. Buoy: $5,000 to $50,000 depending on location and the instrumentation to be mounted
    c. ESP: $400,000 to $400,000
2. Equipment installation:
    a. One-time installation cost: $5K to $10K depending on complexity, SCADA considerations, power requirements and other considerations particular to the location/system. Cost based on LTI estimates for shore-based sondes.
3. Consumables:
    a. Calibration supplies and replacement parts: $250 to $800 per year
    b. Cellular data plan: $250 to $1,000 per year, depending on system and data plan
4. O&M:
    a. In-person visits to clean, calibrate, and perform needed maintenance to shore-based equipment: $5,000 to $10,000 per year depending on proximity to other sites. Cost based on LTI estimates.
    b. In-person visits for buoy deployment/retrieval and to clean, calibrate, and perform needed equipment maintenance: $5,000 to $26,000 per year depending on location, type of equipment, and proximity to other sites. Cost based on LTI contracts and estimates.
    c. EPS maintenance includes frequent cleanings of components, repairs, calibrations and testing: $40,000 to $130,000.
5. Contingency costs for non-routine services and replacements are expected to be about $5,000 per year.

The Lake Erie HABs observing network was not designed and planned in 2009, when the first observing assets were deployed, with the goal of developing an EWS for Lake Erie over time. Instead, assets were deployed over the years by various organizations, in the majority of the cases in an uncoordinated manner, and only when and where funding was available. The network of observing platforms that currently form the system, their types and deployment locations, were not part of an initial plan to strategically and cost-effectively build an optimum HABs EWS. This needs to be taken into consideration when looking at total capital and annual costs for the overall system. An optimized EWS for Lake Erie may look different from what it is in 2021.

The total capital cost of Lake Erie HABs EWS core observation system, in its present state and excluding ESPs, is estimated to range from $340,000 to $1,510,000, based on the costs indicated above. This estimate includes the platforms identified in Section 2.

The annual cost to sustain the system is estimated to range from $594,000 to $2,521,400. These costs have been distributed among a number of government and public entities that have an interest in monitoring the health of Lake Erie. The average annual operational costs for an individual platform will range between $10,500 to $32,800.

To help put those costs into perspective it should be noted that studies have been conducted to assess the cost of HAB to the regional economy. A study[6] conducted in 2015 concluded that the 2011 and 2014 HAB events translated into about $71 million and $65 million respectively in lost economic benefit to the U.S. Lake Erie basin. A more recent study[7] concluded that over a 30-year period HABs-related annual costs in the Canadian Lake Erie basin will be CA$272 million in 2015 prices, if no changes are made.

5.2. EWS funding options

The three funding models are: public, private and public/private hybrid. All models have their unique set of challenges. The most feasible and reliable is the status quo, the public model. Transitioning from one model to the next will take time, effort and the desire from all partners in the EWS to coordinate.

Public Model: A utility, county, state and federal government ‘shared’ model with a consistent and dedicated capital and operational expenditures. Funding mechanisms for such a model could include federal and/or state appropriations, Lake Erie wide special improvement district and other state approaches. In this case each state should be willing to commit to an annual spend but may approach raising those revenues differently.

Most Lake Erie stakeholders indicated that the local, state, and federal governments should be responsible for monitoring and resolving watershed water quality issues, as well as those industries/organizations deemed responsible for the ecosystem degradation. This was noted in the assessment conducted by the Ohio State University, included in Appendix 2. Funding mechanisms for such a model could include federal and/or state appropriations, Lake Erie wide special improvement district and other state approaches. In this case each state should be willing to commit to an annual spend but may approach raising those revenues differently.

---

[6] Bingham M, Sinha SK, Lupi R. Economic Benefits of Reducing Harmful Algal Blooms in Lake Erie. Environmental Consulting & Technology, Inc. Report, 66 pp, October 2015.
[7] Smith RB, Bass B, Sawyer D, Depew D, Watson SB. Estimating the economic costs of algal blooms in the Canadian Lake Erie Basin. Harmful Algae. 2019 Jul; 87:101624.

Among the main challenges to develop a publicly sustained EWS for Lake Erie are 1) the development of an optimal governance and operational framework, considering that Lake Erie is shared by two countries several U.S. states and one Canadian province, and 2) the long-term consistent allocation of funds that is independent of administration changes and capable of supporting the operation, maintenance, and growth of the system.

Private Model: All aspects of the EWS from monitoring equipment to the data communications are managed by private companies and access to data will be purchased by users. The current ad hoc structure and governance/ownership makes a private funded model not financially feasible at this time given investment and risk. In spite of current market, industry expects 10 percent growth annually with the sector over the coming 5 to 10 years as the need for near real-time monitoring increases (local demand and regulatory), local budgets response utilities' need for digital tools in monitoring source water, and costs of equipment decrease as new and lower costs products combined with analytics increasingly come to market.

Public/Private Hybrid Model: Similar to the private sector model, a hybrid model that combines public/private management. Private sector and some utilities are open to a shared model. In fact, a small percentage of water treatment plants in Lake Erie make use of private or public-private models. Private investment is mostly hindered by a loosely defined data user base and their willingness to pay, which makes this a risky business model. Leveraging a public-private model that supports the complete EWS and its expansion instead of one monitoring element of the system might stand a better chance at being successful.

5.3. Acting Smart about Sustainable Funding

The ad hoc fashion in which the existing EWS was formed is the greatest challenge to developing a sustainable funding model to support it. In the foreseeable future, the status quo of the public model will remain in place. To move away from the ad hoc nature of the system, EWS partners, including federal, state, and local governments, utilities, GLOS, GLERL, among others should begin discussing sustainable funding.

Listed are some sustainable funding ideas within the public model for future consideration:

- Setting up a public endowment to support the maintenance and operation of EWS core observing system. Given that one of the most important challenges is the ability to consistently fund the EWS, an alternative could be a one-time creation of an endowment fund. This income-generating investment vehicle, if designed properly, could sustain the long-term operation, maintenance, and growth of the EWS. An organization would need to be designated to oversee the system, its funding allocations, and operations.
- Directly connect state (and possibly provincial) clean water requirements for water utilities to monitoring source water with GLOS. State clean water requirements should include GLOS' robust IT platform as the standard method to share water quality data among utilities and state environmental agencies.

- Tie the core observing network to U.S. federal infrastructure initiatives. Although poorly funded in the past, the water quality monitoring platforms should be considered an essential part of the drinking water and wastewater systems. Infrastructure is expected to receive an influx of federal government funding support in the efforts to boost the post-pandemic economy.

Appendix 1: Guidance for Owning and Maintaining Water Quality Sondes

Prepared by: LimnoTech, under contract to the Great Lakes Observing System, February 2021

Overview

This document is meant as a guide for individuals and organizations that own and maintain multi-parameter water quality sondes for use in an operational setting (primarily drinking water plants).

Equipment

The sonde equipment used on Lake Erie by most operators is the YSI EXO platform of sensors and associated equipment. The main components of the system include a sonde body, individual sensors, biofouling management system, data/power cable, datalogger, and telemetry system (modem and antenna). Though YSI is referenced here, there are a number of manufacturers that make similar equipment including Eureka and In-Situ. For the purposes of this guide it is assumed that you have already selected a system that will meet your needs including the parameters relevant for your specific operational system. Most systems in Lake Erie are monitoring temperature, conductivity, pH, turbidity, chlorophyll, phycocyanin (cyanobacteria pigment), and dissolved oxygen.

Reference manuals for equipment installed in a typical sonde system are linked below

1. YSI EXO
2. Campbell Scientific CR6 datalogger
3. Sierra Wireless 4G LTE modem

Please refer to these equipment resources to learn more details about this equipment and its recommended care, warranty, and troubleshooting guides. The equipment listed above was primarily provided to water treatment plants as a service, meaning trained and paid external technicians (LimnoTech) were responsible for troubleshooting problems, finding a remedy, upgrading firmware, and communicating any known issues related to equipment. The original installer of the equipment should be consulted for any upgrades or changes to settings on any of the devices as there may be interdependencies of software versions and other compatibilities that need to be checked before upgrades are performed.

Installation

When installing a sonde in a water plant you should consider cellular reception, location of a 120V outlet, physical protection from unauthorized people and the environment (e.g., ice, rodents, lightning, birds), and a location for representativeness of readings (sensor location in lake or water plant system—is it seeing what you want it to see?).

The installation process begins with the site selection, mounting prep dependent on the site, programming and testing a datalogger and sonde for data transmission, and then the installation. Depending on the site selection a sonde can be installed by:

- Suspending in low flow water with a rope.
- Suspending in higher flow with a rope and weight attached to the bottom.
- Installing in a perforated 4” PVC pipe with or without a through-bolt and suspending a sonde by a rope.

Sondes can also be installed on surface buoys to monitor water quality at any depth from the surface to the lake bottom. Buoys require a mooring line to anchors and must be recovered prior to winter. Buoys with surface sondes will foul much faster than water plant sondes and need routine maintenance.

Upon installation of existing sondes, plant engineers and representatives from LimnoTech determined the best location to install the sonde. This included identifying access points to source water, raw water, and other opportunities to install equipment with minimal disruption to plant activities, serviceability, and location to closest power outlet and good cellular signal (Verizon, AT&T, or T-Mobile are common carriers). The installation location will factor into the serviceability of the instrument and verification that the equipment is correctly sampling incoming water. Not covered in this guide is any plant process that may affect the quality of data coming from the instrument which includes items like the following

1. Intake well downtime/out of service: Is there more than 1 “Wet-well”? Is the sonde moved to the “active” wet-well?
2. Pump operation schedule: Does the intake operate 24/7? Or is there a known duty cycle?
3. Flood/drawdown: Is your wet-well prone to unusual changes in water level?
4. Debris: Does your wet-well accumulate a certain type of debris (sediment/dreissenid shells) that require closer attention?

Operators should be familiar with the conditions that might affect the flow of water past the instrument and known service schedules and should communicate those with internal data users and look into a power switch (or other data trigger) that would stop data transmission of sensor data during periods of known downtime.

Fouling

Depending on the environment a sonde is installed, different types of biofouling may occur that can influence sensor readings.

- biofouling
- sediment fouling
- permanganate staining

Biological buildup (algae, mussels) and sediment fouling can be removed but using a soft bristle brush along the sonde buoy and the sensor bodies. All sensor heads should only be cleaned with Kimtech wipes or soft paper towels. To keep a sonde from becoming excessively fouled, a sonde should be cleaned once every two to four weeks, depending on site conditions. Cleaning frequency is dependent on the environment the sonde is installed in.

Depending on the installation location and the amount of permanganate applied to the raw water, staining will occur on all parts of the sonde. This is mainly superficial but will affect measurements from optical sensors if the wiper brush begins to fray. To remove staining, mix 1/3 DIW, 1/3 hydrogen peroxide, and 1/3 distilled vinegar and apply to staining.

All sondes use a central wiper brush that brushes the sensor heads at a set interval. The brush only keeps the sensor heads clean and over time will begin to fray. Normally a brush should last a minimum of six month and as long as one year of continuous use. Brushes are simple to replace and can be obtained from the manufacturer or certified third-party dealer.

Data viewing and remote control

Data from the sonde can typically be viewed in a variety of ways. Some of these are listed below:

- direct hand-held and computer display: handset or USB signal adaptor can be used for current values, setting or changing operational parameters, and to begin internal logging.
- local SCADA: current values, graphs and recent history, older data, alerts
- web portal on computer: current values, graphs and recent history, older data, alerts
- web portal on phone: current values, graphs and recent history, older data, alerts
- viewing data from others: real-time, satellite, forecasts of HABs/wind/waves/currents/upwelling

Sensor Maintenance Schedule

In general, the sonde sensors are very robust but do require attention and may malfunction unexpectedly. Sensors such as turbidity and blue-green algae(phycocyanin)/chlorophyll will be prone to noise during periods of high wind events or being in the presence of an algal bloom, as they are optically based sensors and may be affected by large particles moving past the sensor windows.

Sensor values will drift over time and cause measurements to become inaccurate. The typical rates of drifting are different from the non-optical and optical sensors listed below:

- Non-optical sensors: pH, ORP, dissolved oxygen, conductivity.
- Optical: turbidity, blue-green algae/chlorophyll, fluorescent dissolved organic matter (fDOM)

Experience has shown that the non-optical sensors will drift at an approximate monthly rate of 10 µS/cm, 0.1 pH, or 1% saturated dissolved oxygen. The optical sensors can go up to six months without any noticeable drifts.

To check or change a sensor’s current calibration, standards that are representative of the environment need to be used along with a YSI handheld device or USB signal adapter and EXO-KOR software installed on a laptop. The frequency at which a calibration check needs to be completed is largely based on the level of data integration needed. Calibrations of non-optical sensors should be performed at a minimum of monthly and optical sensors bi-monthly to obtain the best quality of data. During periods when water quality measurements are required to be of high quality (e.g., during summer algal blooms or low dissolved oxygen events), it is recommended to calibrate all sensors at least every six months.

Sondes have been calibrated using the standards listed below:

- Conductivity: 1 Point @1000 µS/cm)
- pH: 2 Points at 7 and 10 buffer.
- ORP: 1 Point at -200 mV
- Turbidity: 2 Points at 0 FNU and 124 FNU. DIW used for 0 FNU.
- Dissolved Oxygen: 1 Point using saturation %
- Blue-Green algae (Phycocyanin) RFU: 2 Points at 0 mg/L and 0.625 mg/L Rhodamine. DIW used for 0 mg/L
- Chlorophyll RFU: 2 Points at 0 mg/L and 0.625 mg/L Rhodamine. DIW used for 0 mg/L

The sonde is typically the only component that requires periodic maintenance and is prone to failure. The data cables, datalogger, and modems should have a long service life (10+ years) before any replacement is necessary. It is recommended that sondes have daily, weekly, monthly, and annual service intervals. The sections below lay out what should occur at each of these service intervals.

Daily

On a daily basis, plant operators should look for any anomalies in transmitted sensor data. Operators will be able to tell pretty rapidly if a value looks out of range, stuck, has an unusual pattern, or otherwise isn't matching the expected behavior given operation conditions. These checks can be performed by a computer (automated), but often the algorithms can't spot consistent trends in the data that would indicate an "out of spec" performance by the sensor. Any anomalies should be logged and reported to the appropriate users and data managers. Some issues may not affect the ability of the real-time sensor to support the primary mission (alert to rapid changes in water conditions for example).

Weekly

On a weekly basis, an operator or lay technician should pull the instrument out of the water and make a visual inspection of the sonde to look for signs of biofouling, anything that could affect the mechanical operation of the wiper, or anything unusual. These checks could involve the operator wiping off the exterior of the sonde with a rag/brush and putting it back down into the water. A more thorough review of the previous week's data should also be performed to investigate any trends that seem inconsistent with other similar equipment or laboratory measurements at the plant (e.g., turbidity or pH). Any notes or reference checks should be written down on a service sheet or daily log book.

Monthly

On a monthly basis, a trained technician should perform a more thorough cleaning of the sonde body and evaluation of sonde performance against known standards. Because many of the sensors on the sonde are optical, it may only be necessary to check the "zero" value for optical parameters. The more thorough cleaning should include a full scrubbing of all exterior components of the sonde, removing the sonde guard, clean each sensor probe (without removing each probe from the body), and a wiping of the optical windows with Kimtech wipes. Clean water should be used to rinse the sonde body and all components. Other parameters such as pH, conductivity, and ORP can be checked against one standard (that is re-used monthly for this purpose) and used to quickly check for sensors that have drifted too far from their calibrated value. If any parameter is out of calibration range (typically there's a 10% threshold for out of range tolerance) then a full re-calibration is necessary. The sonde manual walks through the correct procedure to initiate a re-calibration routine. Additional equipment, software, and supplies, may be necessary to perform the re-calibration. It is only recommended to re-calibrate sensors that are out of range, rather than always going through the calibration routine every month. User error and inconsistency in calibration methods can lead to inconsistency in the baseline values and may affect the ability to detect long term changes in baseline conditions. Additionally for optical based sensors it is often a dirty optical window that is causing the sensor to read an out of range value. Consider cleaning the optical window again and ensuring no air bubbles are present before initiating a re-calibration process for optical sensors.

Annual

On an annual basis, a factory trained technician should perform a complete teardown of the sonde body and sensors and perform a full evaluation of the sensors. Several components are recommended to be replaced annually by the factory and the factory can perform these higher level maintenance items all at once. The factory-trained technician can replace seals, pH consumable modules, DO caps, wiper seals, brushes, and other O&M items. The user should either participate in a regional collection and intercalibration of similar sondes, or have this service performed by shipping the sonde to the manufacturer for service..

Parts and equipment needed for setting up, calibrating, and maintaining sonde:

- Hand-held sonde controller or laptop
- Calibration solutions
- Calibration beakers and sonde holder (clamp, stand)
- Cleaning solutions
- Soft bristle brush
- Fiber-free wipes (e.g. Kimtech)
- Replacement brushes for wipers

Additional maintenance of sondes, data loggers, communications systems, or SCADA interfaces should only be performed by trained technicians. Some initial troubleshooting can be performed remotely or over the phone, but many operational issues and communications problems require site visits to properly address issues. Loaner instruments and components may be available to minimize monitoring interruptions during critical periods.

Appendix 2: EWS Architecture Components

A2.1 EWS System Architecture

The GLOS EWS (Figure A2.1) leverages Amazon Web Services, for the development and deployment of the code, services and processes related to HABs in Lake Erie and the Great Lakes as a whole. These include many of the core Cloud components described in the graphic above, including an API Gateway, event streaming and management, service logic deployment, and data storage. Further details about the services and specifics of the GLOS EWS in AWS are in the following graphic.

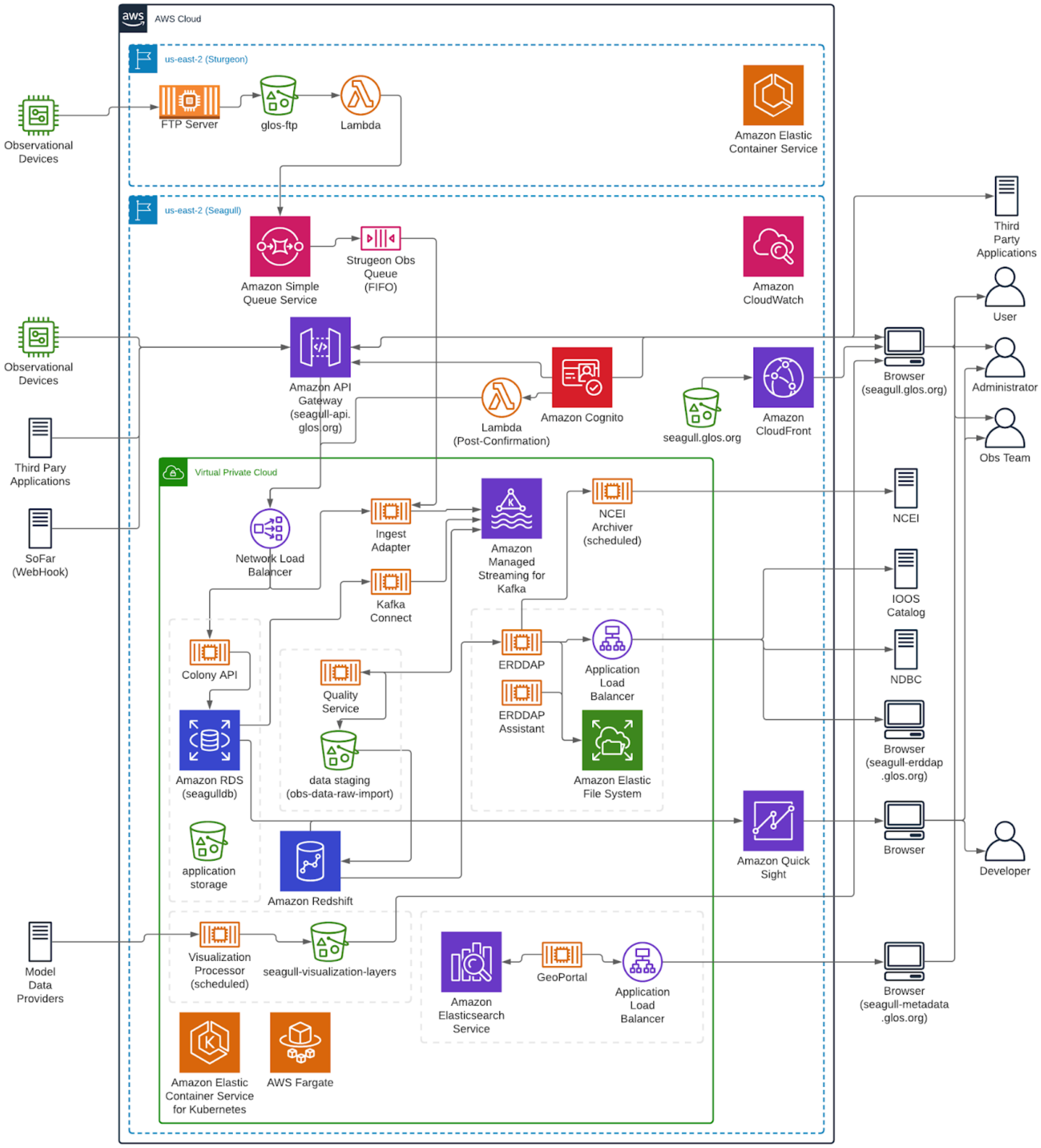


Figure A2.1: An overview of the architecture diagram of the core AWS technologies utilized for Seagull, the EWS parent platform and main GLOS IT platform.

AWS API Gateway

Amazon API Gateway is a fully managed service that makes it easy for GLOS to create, publish, maintain, monitor, and secure APIs at any scale. APIs act as the "front door" for applications to access data, business logic, or functionality from our backend services. Using API Gateway, GLOS can create RESTful

APIs that enable real-time two-way communication applications. API Gateway supports containerized and serverless workloads, as well as web applications.

API Gateway handles all the tasks involved in accepting and processing up to hundreds of thousands of concurrent API calls, including traffic management, CORS support, authorization and access control, throttling, monitoring, and API version management. API Gateway has no minimum fees or startup costs. We will be paying for the API calls we receive and the amount of data transferred out and, with the API Gateway tiered pricing model, we can reduce the cost as our API usage scales.

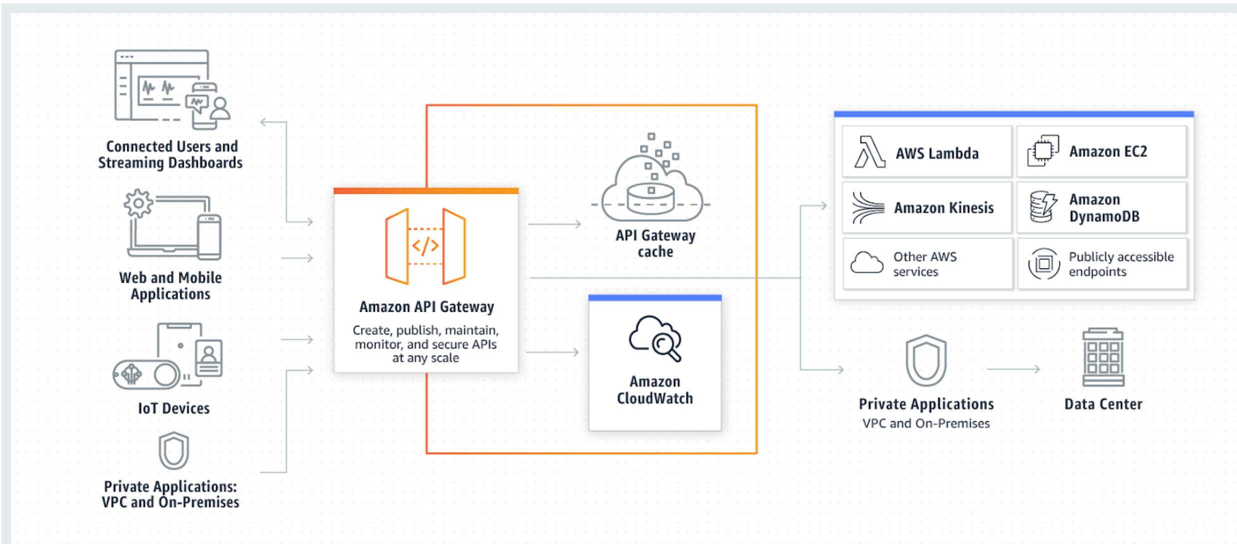


Figure A2. 2: A logical diagram to show how a typical AWS API Gateway works on the cloud.

Event Stream: AWS Managed Streaming for Apache Kafka (MSK)

Managed Streaming for Apache Kafka (MSK) is a fully managed, highly available, secure Apache Kafka service from AWS. Apache Kafka is an open-source distributed event streaming platform. It is especially well suited for handling high-throughput, low-latency, real-time data feeds. It is very common in the cloud computing industry and very well supported by a variety of tools. Traditionally, Apache Kafka clusters are challenging to manage, but those challenges are minimized by allowing AWS to handle them. Pricing for MSK is a combination of perpetual broker instance costs, broker storage fees, and data transfer fees. The broker instances start at $0.0456 per hour. Storage is $0.10 per GB-month. Data transfer fees are the standard AWS fees starting at $0.00 per GB for incoming data and $0.09 per GB for outgoing data, with no charge for internal data transfers.

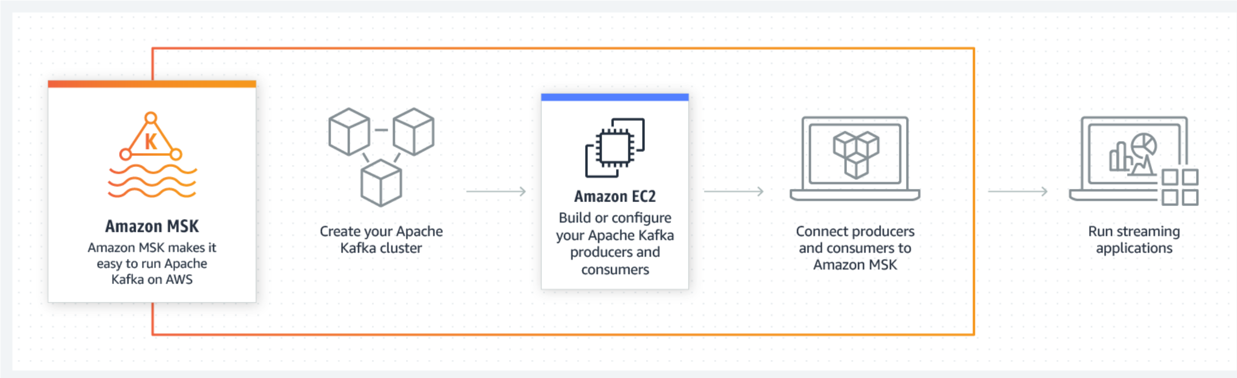


Figure A2.3: A logical diagram that shows how a typical AWS MSK works on the cloud.

Obs. data storage and management : AWS Redshift

There are two primary approaches considered for handling observation data: Data Warehouse and Data Lake. A data warehouse is a more structured but predictable approach to storing large amounts of data. Data Lakes are usually a less expensive way to store data but not necessarily a less expensive way to query data. Generally data warehouses are built from Online analytical processing (OLAP) databases. Data lakes are usually built from blob storage or distributed file systems and leverage distributed query processing and file scanning tools.

AWS Redshift is a fully managed columnar database from AWS based on PostgreSQL. Compliant with ODBC and JDBC making it very easy to integrate with other applications. Highly scalable and predictable performance. It contains many value-added features such as query caching that help to improve performance. Pricing starts at $0.25 per Hour for a 160GB cluster. Pricing is flexible with lots of options for expansion. Redshift is a traditional OLAP data warehouse. In addition, Redshift also contains support for spatial data types limited to two spatial dimensions.

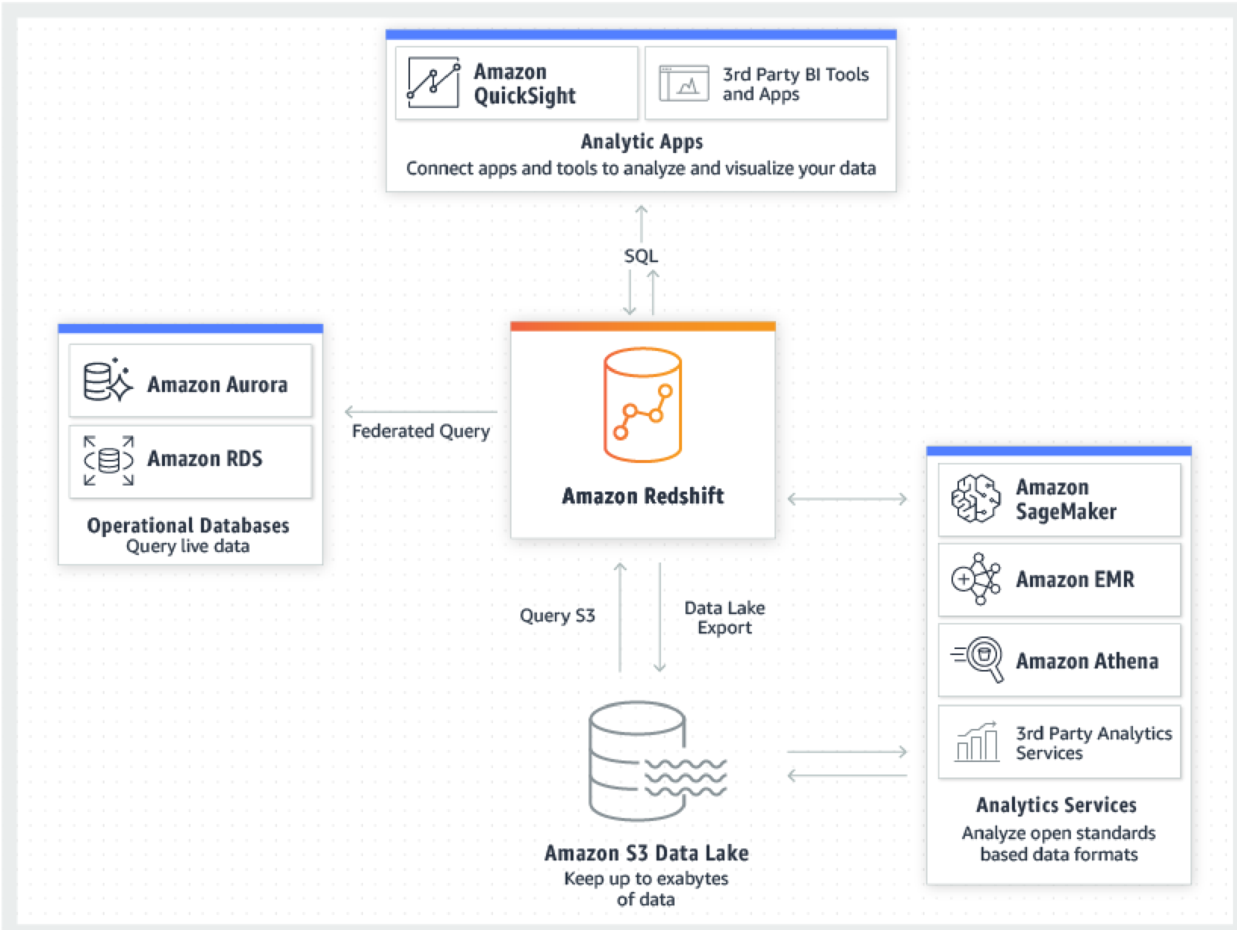


Figure A2.4: A logical diagram that shows how a typical AWS Redshift database works on the cloud.

Service Logic Deployment: AWS Lambda + EKS + Fargate

The EWS architecture component for service logic deployment consists of Lambda, AWS EKS, and AWS Fargate. This mix of technologies offers a very flexible architecture while managing costs and vendor lock-in.

AWS Lambda is a serverless option from AWS that is used to deploy compute resources to AWS managed infrastructure. Lambda is easily scalable to meet the exact demand of the workload. This service natively supports a variety of programming languages including: Java, Go, Node.js, and Python. Pricing is based on the amount of compute time and memory that is consumed when processes a request. The pricing structure is $0.20 per million requests and $0.0000166667 for every GB-second of processing.

Lambda is well suited for low volume or inconsistent workloads that may include spikes in demand. One limitation is the stateless computing architecture that is used. This limitation causes concerns for managing connection pools to databases and introduces complexity for stateful stream processing. If stateful processing is desired, it becomes necessary to store the state external to the lambda function. This introduces latency and additional demand on datastores.

Elastic Kubernetes Service (EKS) is a fully managed Kubernetes service. It is very similar to ECS but it is not exclusive to AWS. Managed Kubernetes services are available in every major cloud vendor. Kubernetes has gradually become the de facto standard in the cloud computing industry for running container workloads. EKS has the option of deploying to EC2 the same way EC2 is an option for ECS.

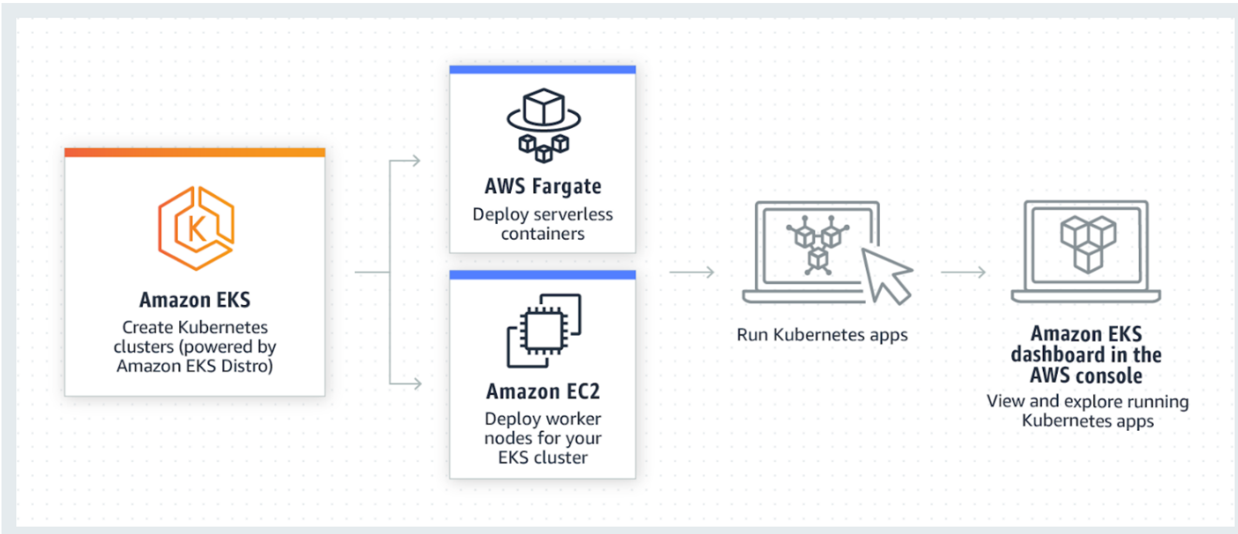


Figure A2.5: A logical diagram that shows how a typical AWS Elastic Kubernetes Service (EKS) is deployed on the cloud.

AWS Fargate is a technology that is used with Amazon EKS to run containers without having to manage servers or clusters of Amazon EC2 instances. With AWS Fargate, we no longer have to provision, configure, or scale clusters of virtual machines to run containers. This removes the need to choose server types, decide when to scale your clusters, or optimize cluster packing.

EWS User Pool Management : AWS Cognito

Amazon Cognito lets us add user sign-up, sign-in, and access control to your web and mobile apps quickly and easily. Amazon Cognito scales to millions of users and supports sign-in with social identity providers, such as Apple, Facebook, Google, and Amazon, and enterprise identity providers via SAML 2.0 and OpenID Connect.

EWS uses AWS Cognito to manage the user pools on Seagull and EWS. The signed up users are intended to have access to the alerting and notifications on the EWS.

Appendix 3: API Gateway Endpoints

GLOS' Seagull is designed to be HTTPS first for better security. Seagull's API is based off of the OpenAPI specification (https://oai.github.io/Documentation/) and is fully documented online (https://seagull-api.glos.org/docs).

In this appendix we list select endpoints of interest in the context of the early warning system.

Seagull native

The API hostname for all of the listed endpoints is https://seagull-api.glos.org.

- Alerts:
    - /api/v1/obs-alert-events - covers alert status, start, end, and source (platform)
    - /api/v1/obs-alert - manages alert events
- Observations
    - /api/v1/obs-dataset-summaries - gets summaries of observations datasets
    - /api/v1/obs-datasets - gets all obs-datasets
    - /api/v1/obs-datasets.geojson - gets GeoJSON formatted data (e.g. map layers)
    - /api/v1/obs-datasets/{obsDatasetId}/metadata - gets metadata for a dataset
    - /api/v2/obs - gets observations for a specified date range
    - /api/v2/obs-latest - gets the latest observations data (last recorded)
- Parameters
    - /api/v1/parameters - gets all available parameters
    - /api/v1/parameter-configurations - gets the metadata for parameters, including display names, units, alert thresholds, and more
- Platforms
    - /api/v1/platforms - gets all available platforms

External to Seagull

- https://noaa-ofs-pds.s3.amazonaws.com/index.html - for the Great Lakes Operational Forecasting model outputs. This URL contains all of NOAA's operational models as well

Appendix 4: GLOSsary of Terms

AMQP:	Advanced message queuing protocol
API:	Application programming interface
AWS:	Amazon web services
CIGLR:	Cooperative Institute for Great Lakes Research
ELISA:	Enzyme-linked immunosorbent assay
EPA:	Environmental Protection Agency
ERDDAP:	The name of a data server
ESP:	Environmental system processor
ETL:	Extract, transform, and load
EWS:	Early Warning System
FTP:	File transfer protocol
GLERL:	Great Lakes Environmental Laboratory
GLOFS:	Great Lakes Operational Forecast System
GLOS:	Great Lakes Observing System
HABs:	Harmful algal blooms
HTTP(S):	Hypertext transfer protocol (secure)
IT:	Information technology
NOAA:	National Oceanic and Atmospheric Administration
IOOS:	Integrated Ocean Observing System
IoT:	Internet of things
JSON:	JavaScript Object Notation; a data-interchange format.
LTI:	LimnoTech, Incorporated.
MBARI:	Monterey Bay Aquarium Research Institute
MQTT:	
NCCOS:	National Centers for Coastal Ocean Science
NDBC:	National Data Buoy Center
NetCDF:	Network common data form
NOAA:	National Oceanographic and Atmospheric Administration
O&M:	Operation and maintenance
OLAP:	Online analytical processing
QA/QC:	Quality assurance/quality control
QARTOD:	Quality assurance quality control of real-time oceanographic data
Query:	Interrogating and retrieval of data based on user parameters
S3:	Simple Storage Service
SDK:	Software development kit
UI:	User interface
UNDRR:	United Nations Office for Disaster Risk Reduction
USGS:	United States Geological Survey
UWSS:	Union Water Supply System

WTP:	Water treatment plant
XML:	Extensible markup language
YSI:	The name of a company that makes environmental sampling equipment