Mobile device application for monitoring cyanobacteria harmful algal blooms using Sentinel-3 satellite Ocean and Land Colour Instruments.

Blake A. Schaeffer[*a], Sean W. Bailey[b], Robyn N. Conmy[c], Michael Galvin[d], Amber R. Ignatius[e], John M. Johnston[d], Darryl J. Keith[f], Ross S. Lunetta[a], Rajbir Parmar[d], Richard P. Stumpf[g], Erin A. Urquhart[h], P. Jeremy Werdell[b], Kurt Wolfe[d]

[a]National Exposure Research Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, Research Triangle Park, NC 27711, USA

[b]Ocean Ecology Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD, USA

[c]National Risk Management Research Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, Cincinnati, OH, 45268, USA

[d]National Exposure Research Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, Athens, GA, 30605, USA

[e]University of North Georgia, Institute for Environmental and Spatial Analysis, Oakwood, GA, 30566, USA

[f]National Health and Environmental Effects Research Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, Narragansett, RI, 02882, USA

[g]National Oceanic and Atmospheric Administration, National Centers for Coastal Ocean Science, Silver Spring, MD, USA

[h]Oak Ridge Institute for Science and Engineering (ORISE), National Exposure Research Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, Athens, GA, 30605, USA

*Corresponding author: schaeffer.blake@epa.gov, 919-541-5571

Highlights

- Cyanobacterial blooms (CyanoHAB) are a human and ecological health concern.
- CyanoHABs are infrequently monitored to provide needed management information.
- The CyAN mobile app provides passive access to satellite data and basic analysis.
- Quantitative comparison against 2017 state advisories demonstrated app capability.
- CyAN app delivers spatial and temporal information on cyanobacteria concentration.

Software availability

Software name: Cyanobacterial Assessment Network (CyAN)
Developers: U.S. Environmental Protection Agency, Office of Research and Development
Contact information: Mike Galvin, Galvin.mike@epa.gov
Alternative contact: John M. Johnston, johnston,johnm@epa.gov
Hardware required: Smartphone
Software required: Android 4.1 and above
Program language: JAVA and Python
Availability and cost: Currently available to US Departments of Health and Departments of Environment state agencies upon request. No cost.

Abstract

Cyanobacterial harmful algal blooms (cyanoHAB) cause human and ecological health problems in lakes worldwide. The timely distribution of satellite-derived cyanoHAB data is necessary for adaptive water quality management and for targeted deployment of water quality monitoring resources. Software platforms that permit timely, useful, and cost-effective delivery of information from satellites are required to help managers respond to cyanoHABs. The Cyanobacteria Assessment Network (CyAN) mobile device application (app) uses data from the European Space Agency Copernicus Sentinel-3 satellite Ocean and Land Colour Instrument (OLCI) in near real-time to make initial water quality assessments and quickly alert managers to potential problems and emerging threats related to cyanobacteria. App functionality and satellite data were validated with 25 state health advisories issued in 2017. The CyAN app provides water quality managers with a user-friendly platform that reduces the complexities associated with accessing satellite data to allow fast, efficient, initial assessments across lakes.

Keywords: satellite, water quality, mobile device application, harmful algal bloom, cyanobacteria

1. Introduction

1.1 Human and ecological health problems from cyanobacterial harmful algal blooms

Water quality is a critical consideration in determining water resource availability for human consumption, aquatic life, and recreation (U.S. EPA, 2013). Harmful algal blooms are environmental events that occur when algal or cyanobacterial populations impact water quality and result in negative environmental or health consequences (Smayda, 1997). Cyanobacterial harmful algal blooms (cyanoHAB) occur worldwide and have been documented across the United States (Loftin et al., 2016). U.S. states frequently issue health advisories or close recreational areas due to potential risks from cyanoHAB exposure (Graham et al., 2009). CyanoHABs may produce toxins and cause nuisance odors, hypoxia, unappealing surface scums, undesirable finished drinking water, increased drinking water treatment costs, and economic and infrastructure costs such as loss of revenue from recreation and businesses that rely on appealing or potable water (Dodds et al., 2009; Steffensen, 2008). New tools are needed to facilitate the development of reliable and cost-effective monitoring programs at lake, watershed, state, regional, and national scales.

1.2 Management need

Many U.S. states experience challenges in developing cyanoHAB monitoring programs because of the need to cover large geographic areas with insufficient resources. A single management agency often oversees numerous isolated lakes scattered across a large landscape. Water quality monitoring efforts also are constrained by the availability of personnel and limited financial resources, presenting managers with a serious challenge in prioritizing water quality sampling regimes to best monitor dispersed waterbodies. Water quality managers need access to timely and consistent data to protect the designated and beneficial uses of water. Satellites can provide such data.

Although satellite data has been available for a number of years, historically, few management decisions have been based on satellite information because data dissemination to water quality managers has been limited to either photographs or file-formatted products that require specialized training to process images and interpret data. Managers would substantially benefit from software that reduces the barrier of accessing satellite data to facilitate better public and environmental health protection of water bodies. Timely and effective distribution of satellite-derived data is necessary to provide warnings within days and seasonal assessments in the same calendar year. Managers responding to the immediate impacts of cyanoHABs need timely, useful, and cost-effective delivery of information from the satellite data (Schaeffer et al., 2013).

Satellite remote sensing technology provides an ability to assess cyanoHAB abundance for spatially resolvable inland recreational and public water supplies using algal pigments as surrogates for HAB and cyanoHAB abundance at sufficient spatial-temporal resolution to detect changes. In particular, the European Space Agency Copernicus program provides a new series of Sentinel-3 Ocean and Land Colour Instruments (Berger et al., 2012; Donlon et al., 2012) at 300 meter (m) pixel resolution suited for detecting cyanoHAB abundance changes (frequency, extent, magnitude, and duration) in larger lakes (Clark et al., 2017; Urquhart et al., 2017).

1.3 Current software status

A number of ecological forecast software tools already exist related to harmful algal blooms and chlorophyll-a that serve as a proxy for eutrophication status and phytoplankton biomass. Examples include forecasting with combined sonde mooring data and an artificial neural network (Coad et al., 2014); varying the chlorophyll-a to phytoplankton biomass ratio in a CE-QUAL-W2 water quality model (Sadeghian et al., 2018); fuzzy logic models for algal bloom forecasts (Kim et al., 2014); applying Markov chain Monte Carlo bayesian modelling to understand nutrient and zooplankton controls on cyanobacteria (Malve et al., 2007); integrated models to capture land use, nutrient budgets, meteorological and hydrological data to forecast cyanobacteria concentrations to inform adaptive management practices (Norton et al., 2012); and a Windows-based Software EcoTaihu model integrates water quality measures and satellite data to predict cyanobacteria in lake Taihu (Zhang et al., 2013). However, none of these existing software platforms provide the ability to visualize satellite derived water quality data directly or view time series of satellite data to inform decision making.

Two primary scientific software packages traditionally are used for satellite data processing and analysis related to water quality satellite missions. The U.S. software package is the National Aeronautics and Space Administration (NASA) Sea-viewing Wide Field-of-view Sensor (SeaWiFS) Data Analysis System (SeaDAS) (Baith et al., 2001), an open-source and free software package. SeaDAS is a comprehensive software package for the processing, display, analysis, and quality control of a wide array of satellite data. While the primary focus of SeaDAS has historically been ocean color data, it is applicable to other satellite-based earth science data analyses, such as inland and coastal water quality data. The European software package is the SeNtinel Application Platform (SNAP) available from the European Space Agency. SNAP also is an open-source platform and focuses on the exploitation of earth observation data.

Both software packages are desktop computer-based and require some scientific knowledge in the field of ocean color remote sensing as well as sufficient computer hardware to handle the satellite images and processing capabilities. In addition, computer code language expertise, typically in JAVA or Python, is beneficial to batch process large numbers of satellite files. These software packages produce derived water quality products such as cyanoHAB abundance. However, the software to date is not developed for repeated, intuitive, and rapid assessment of inland waters for cyanoHAB monitoring by a diverse variety of water quality managers and stakeholders. Therefore, alternative software solutions are necessary to reduce the data access limitations and to reduce required management programmatic support (Schaeffer et al., 2013) for satellite-derived data on cyanoHABs.

1.4 Mobile device application solution

Unlike previous software packages, a mobile device application (app) would reduce the need for scientific expertise in ocean color interpretation and hardware requirements associated with the use of satellite data. An app would provide intuitive ability through a graphical user interface (GUI) to scan water bodies for changes in cyanoHAB abundance without expertise in computer programming or computer languages. Georeferenced data would allow managers to monitor their particular water bodies of interest without having to filter through numerous satellite images of water bodies not associated with their region. Managers could set query thresholds to identify if cyanoHAB abundance exceeds a certain limit. In addition, by using advanced alert systems, an app would allow passive reception of data instead of active acquisition minimizing the amount of time commitment on behalf of the manager. Managers would benefit from multiple methods of notification through a mobile phone app that could, for example, change the colors of map pins based on previously set threshold levels.

In addition, remote sensing data traditionally are provided as files covering an entire region with data for a particular moment in time. An app would allow managers to select a single location of interest to quickly visualize the quantified cyanoHAB abundance value, and a scaling capability would provide larger ecosystem context. The ability to quickly query a single pixel location and obtain a time-series of information has not been readily available for nontechnical users and is only now being addressed through time-sensitive data formats such as data rods (Grant and Gallaher, 2015; Teng et al., 2016). With a mobile phone app, in addition to obtaining a current cyanoHAB abundance value, users would also obtain a time-series of historical cyanoHAB values for the location of interest, providing temporal context.

A software design allowing for a 70/30 contribution blend between scientist and managers, which has previously been successful in the food industry (Ariely, 2010), is an attempt to maximize ownership of the satellite data. Scientists provide 70% of data handling, including processing, geophysical product development, basic assessment capabilities, and delivery. Managers are responsible for about ~30% of the information associated with their existing efforts such as setting thresholds, identifying site locations, and making decisions to take additional action based on available data.

This effort focuses on the development of a mobile platform data delivery app because mobile data access and mobile devices are more ubiquitous and already outpacing sales of traditional desktop computers. Globally, more than 84% of the population lives in areas with mobile 3G or better data access (ICT, 2016), with 75% smart-phone adoption in North America accessing 4G and eventually 5G (GSM Association, 2016b). Global smart phone and mobile data access is expected to reach 70% of the world population by 2020 (GSM Association, 2016a).

We demonstrate the use of the Cyanobacteria Assessment Network (CyAN) app, which uses information from the European Space Agency Sentinel-3 satellite Ocean and Land Colour Imagers (OLCI) to create a cost-effective, timely, and intuitive satellite data delivery system accessible through Android 4.1 and above smartphones. The SeaDAS software package produces derived water quality products such as cyanoHAB abundance that are the processing backbone of the CyAN app.

2. Materials and Methods

2.1 Copernicus Sentinel-3 satellite Ocean and Land Colour Instruments

Full resolution (300 m at nadir, the point directly below the satellite on the Earth surface) scenes from the European Space Agency's OLCI were obtained for the Contiguous United States (CONUS) starting in 2016. Standard OLCI Level-1B data are archived at the NASA Ocean Color website https://oceandata.sci.gsfc.nasa.gov/. The data were processed using the NASA standard ocean color satellite SeaDAS processing software package version 7.4 (Baith et al.,

2001) and the Shuttle Radar Topography Mission (SRTM) GC land mask (Carroll et al., 2009). Images were processed to Albers Equal Area projection with nearest-neighbor interpolation.

Spectral albedo, $\rho_s(\lambda)$, was generated by removing Rayleigh reflectances and molecular absorption from the top-of-atmosphere signal measured by the satellite. Clouds were masked using a spectral albedo threshold algorithm that accounts for turbid water to eliminate misidentification of pixels with bright reflectances resulting from intense blooms. The ρ_s estimates were used to calculate the Cyanobacteria Index (CI) from the spectral shape algorithm centered on 681 nanometers (nm), CI = -SS(681) (Wynne et al., 2008), and routinely used in Lake Erie (Stumpf et al., 2016).

The OLCI CI output for each image then was converted to cyanoHAB abundance in cells per milliliter (cells mL^{-1}) following Wynne et al. (2010), where cyanoHAB abundance=1.0×10^8*CI. Field validation of the CI algorithm was previously demonstrated (Clark et al., 2017; Lunetta et al., 2015; Tomlinson et al., 2016). Clark et al. (2017), reported correspondence across the spectrum of cyanoHAB abundance ranges spanning 10,000 to >1 million cells/mL (mean absolute percentage error, MAPE = 28.6%, coefficient of determination, R^2 = 0.95). Satellite derived values below 109,000 cells/mL and above 1,000,000 cells/mL had correspondence of above 80% with in situ samples collected within 7±days of a satellite match up. While the CI algorithm had lower correspondence performance between 109,000–1,000,000 cells/ mL, this was expected due to the lack of validation data in this concentration range and the large temporal range for coincident satellite match-ups (Lunetta et al., 2015). The categorization of satellite derived CI values based on threshold levels, as described below, would further reduce the impact on algorithm error and uncertainties. Weekly 7-day composite images were created by retaining the maximum value detected for each pixel within the time period. The use of a 7-day composite minimizes the impacts of cloud cover and maximizes the frequency of available data based on a typical work week to guide management decisions.

2.2 Thresholds

Cell counts and microcystin concentrations are commonly used to evaluate potential health risk, and many states have customized thresholds based on additional information gathered locally (Graham et al., 2009). For example, Oklahoma and Massachusetts developed state-specific guidelines to establish protective levels. Oklahoma issues warnings to lake users if cell counts exceed 100,000 cells mL^{-1} and microcystin concentrations exceed 20 µg L^{-1}.

Massachusetts has established guidelines for issuing an advisory against contact with water when cell counts exceed 70,000 cells mL^{-1} or microcystin concentrations exceed 14 μg L^{-1}. The World Health Organization (WHO) has a three-level guideline approach, which describes concentrations of the ubiquitous photosynthetic pigment chlorophyll-a and cyanobacterial cell abundance (cells mL^{-1}) to determine the level of associated risk to support a warning or closure. WHO provides estimates of microcystin that could correspond to the cell abundance at each guideline level. The U.S. Environmental Protection Agency also has the drinking water health advisory for cyanobacteria microcystins toxin (U.S. EPA, 2015). Satellite observations cannot detect toxins (Stumpf et al., 2016) but can quantify cyanoHAB abundance (Kutser, 2009). CyanoHAB abundance is perhaps better suited for assessing nationwide risks due to limitations related to toxin monitoring (Clark et al., 2017). Therefore, the CyAN app allows for thresholds to be set based on user preferences of cyanoHAB abundance in cells mL^{-1}.

2.3 State advisor validation

State advisory data was accessed January 2018 from California at http://www.mywaterquality.ca.gov/habs/where/freshwater_events.html, Oregon http://www.oregon.gov/oha/ph/HealthyEnvironments/Recreation/HarmfulAlgaeBlooms/Pages/Blue-GreenAlgaeAdvisories.aspx, New York at http://www.dec.ny.gov/chemical/83332.html, Idaho at http://www.deq.idaho.gov/water-quality/surface-water/recreation-health-advisories/, New Jersey at http://www.nj.gov/dep/wms/bfbm/cyanoHABevents.html, Utah at https://deq.utah.gov/Divisions/dwq/health-advisory/harmful-algal-blooms/, and Vermont at https://apps.health.vermont.gov/vttracking/cyanobacteria/2017/. Only Oregon and California included latitude and longitude coordinates of the advisory location, so these same coordinates were used in the demonstration of the mobile application. Random locations were selected within the listed waterbodies for New York, Idaho, New Jersey, Utah and Vermont since no coordinates were provided with the advisory information.

3. Architecture and implementation

The U.S. Environmental Protection Agency (EPA) crowd-sourced the development of the CyAN app for use on the Android 4.1 operating system and above using the JAVA programming language. Architecture was developed to achieve the following objectives:

- Provide passive data delivery system for satellite derived cyanoHAB concentration products to water quality managers.

- Provide weekly composites to the managers for combination of timely data distribution and greatest amount of available data (e.g. cloud cover interference).
- Provide a simple, intuitive data display format to water quality managers.
- Provide the ability to monitor multiple locations and quickly view current conditions and seasonal patterns.

The mobile application software is deployed on Red Hat Linux servers. Figure 1 illustrates the communication between system components on the network infrastructure. The data management and administrative servers are backend system components supporting the Android app. The proxy and lightweight directory access protocol (LDAP) servers publically expose the data services provided by the backend and provide upload-access to data imagery.

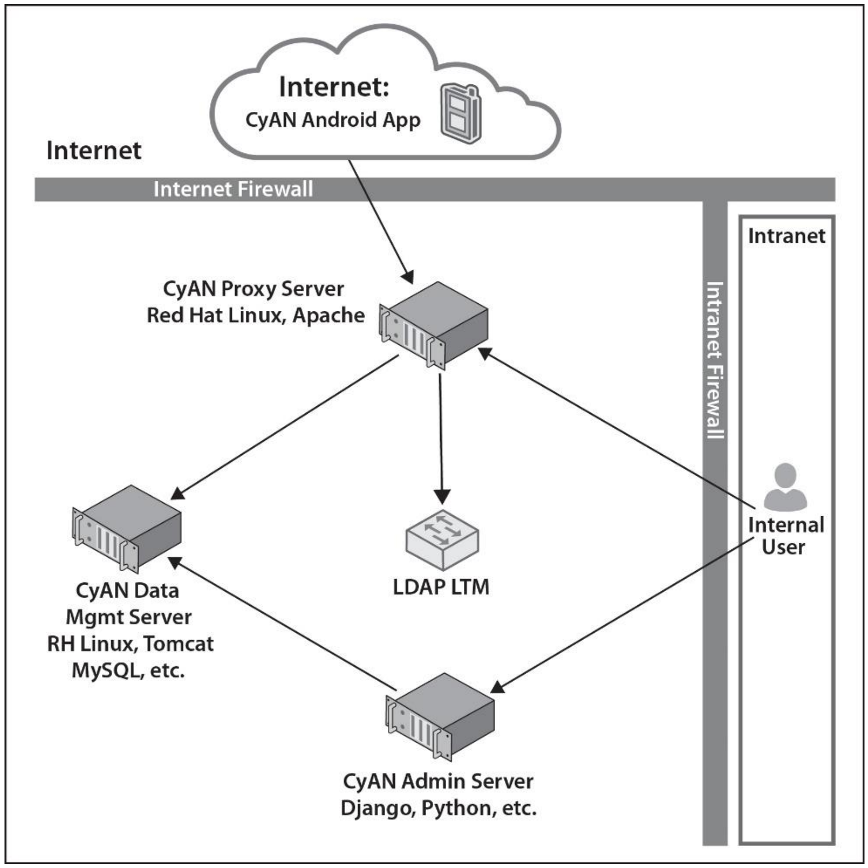

Fig. 1. Schematic of hardware configuration of three servers hosting CyAN mobile app. Proxy, data management, and administrative servers are behind a firewall to the publically accessible internet.

The application logic (Fig. 2) illustrates the communication between the user, Android application, data management server, and administrative website for some common tasks performed. The CyAN administrative server (admin website) is composed of an Apache HTTP web and Tomcat application server. The server is a security-configured, Linux-based operating system running a Python-written administrative web application. This hosting system supports a Django framework web application that mediates and services connections between the uploaded satellite data (cyanoHAB abundance extracted from geoTIF files) and a relational database (MySQL). The Django framework provides the structure for implementing a Model View Controller designed web application (admin tool) that serves as an interface for data management functions, imagery upload, triggering of backend processing for data extraction, data validation, data standardization, and database population.

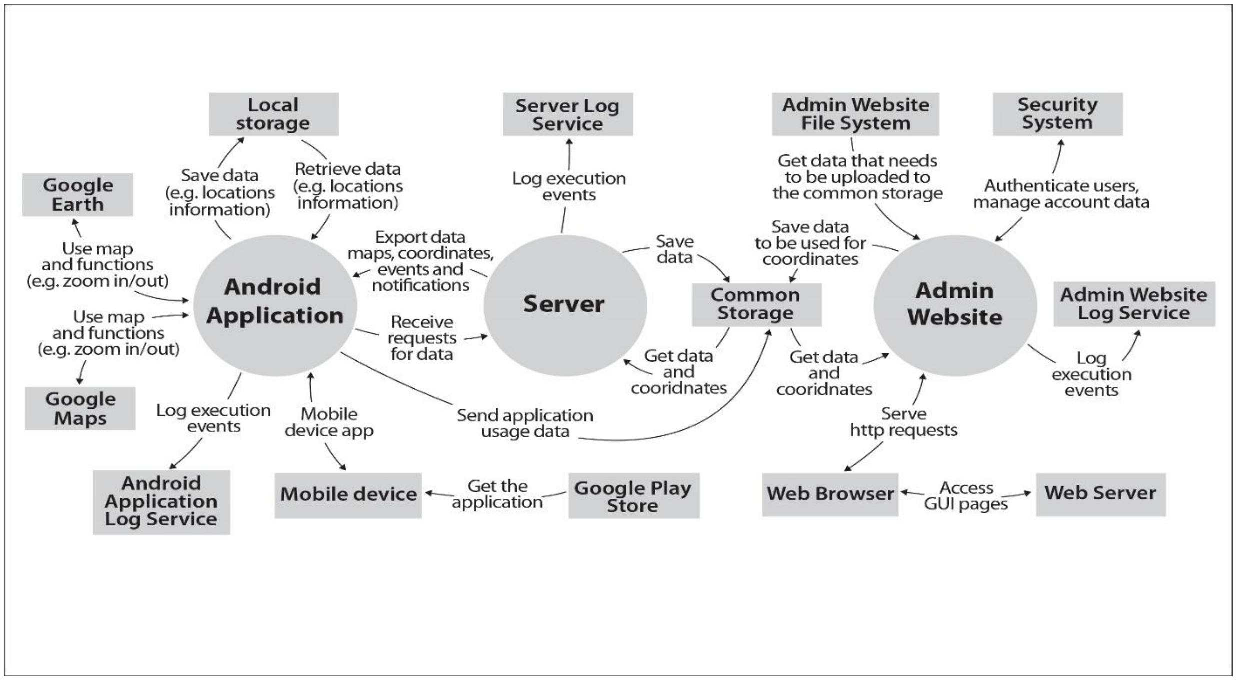

Fig. 2. Logical architecture of information and services between the administrative website (where data is uploaded by scientists), server (that allows import and export of the data and locational information), and Android app (installed on the water quality managers' mobile device).

The data management server is an Apache/Tomcat-frontend Java application that services representational state transfer (REST) application programming interface (API) data calls from the app and responds to the administrative server admin tool's requests for data processing. The data management server is a security configured Linux-based operating system running the Java application's processing backend and hosting the relational MySQL database for storage of satellite data. Processing code employs dependency injection (Inversion of Control [IoC] containers) via the Spring framework to enhance code re-use and component unit testing mechanisms. The MySQL database houses the cyanoHAB abundance information by location. Locations represent 300 m × 300 m raster grid centroid latitude and longitude coordinates. The MySQL location data table was optimized for spatial queries by using the MyISAM database engine and indexing latitude and longitude columns; enhancing response times for data and image API calls by the app. Processing of uploaded imagery to the admin tool web application triggers raster data validation, standardization and extraction, and subsequent storage into the relational database.

The app is an Android component specifically targeted to Android v4.1 (Jelly Bean) through v4.4 (KitKat) that initially tests well through v7 (Nougat). The Android operating system was selected because it provides an open-source platform (Jonoski et al., 2013) and because >75% of the market share uses Android software (IDC, 2017). Among the source dependencies utilized by the app are Google Play for Google Maps and Google Earth base map display, Apache's HTTP Client for managing HTTPS REST calls and responses, Jackson.core that helps parse JSON-formatted HTTP requests and responses, and achartengine for the creation and control of plots. The API data and image requests are REST-style web services using the HTTPS protocol.

4. Case studies and discussion

4.1 Case 1 – Base map

The CyAN app allows multiple locations of interest, such as resolvable recreational sites or public surface drinking water intakes (Clark et al., 2017), to be marked with color-coded pins (Fig. 3a). Clark et al (2017) previously developed a method for quantifying resolvable lakes given a satellite sensor spatial resolution at nadir. Any water pixel adjacent to the land mask should be used with caution due to the potential for mixed land-water pixels and land adjacency effects. National Hydrography Dataset Plus version 2.0 (NHDPlusV2) lake polygons (McKay et al., 2012) with at least 3 valid water pixels were included as a resolvable waterbody in this study (Table S1). The pin color codes correspond to threshold levels selected by the end-user from a swipe tab using sliding bars (Fig. 3b). The Geographic Coordinates tab permits end-users to input specific latitude and longitude coordinates for locations such as monitoring stations, recreational locations, and public surface drinking water intakes. Users may mark locations on a base map, remove locations, view current cyanoHAB data and recent changes in abundance for marked locations, select locations for comparison, review and set category thresholds, and view and clear notifications. The Notification tab permits administrators to send messages to all end-users, such as notices for software updates, new data uploads, acknowledgement of software bugs, and information on events of national significance.

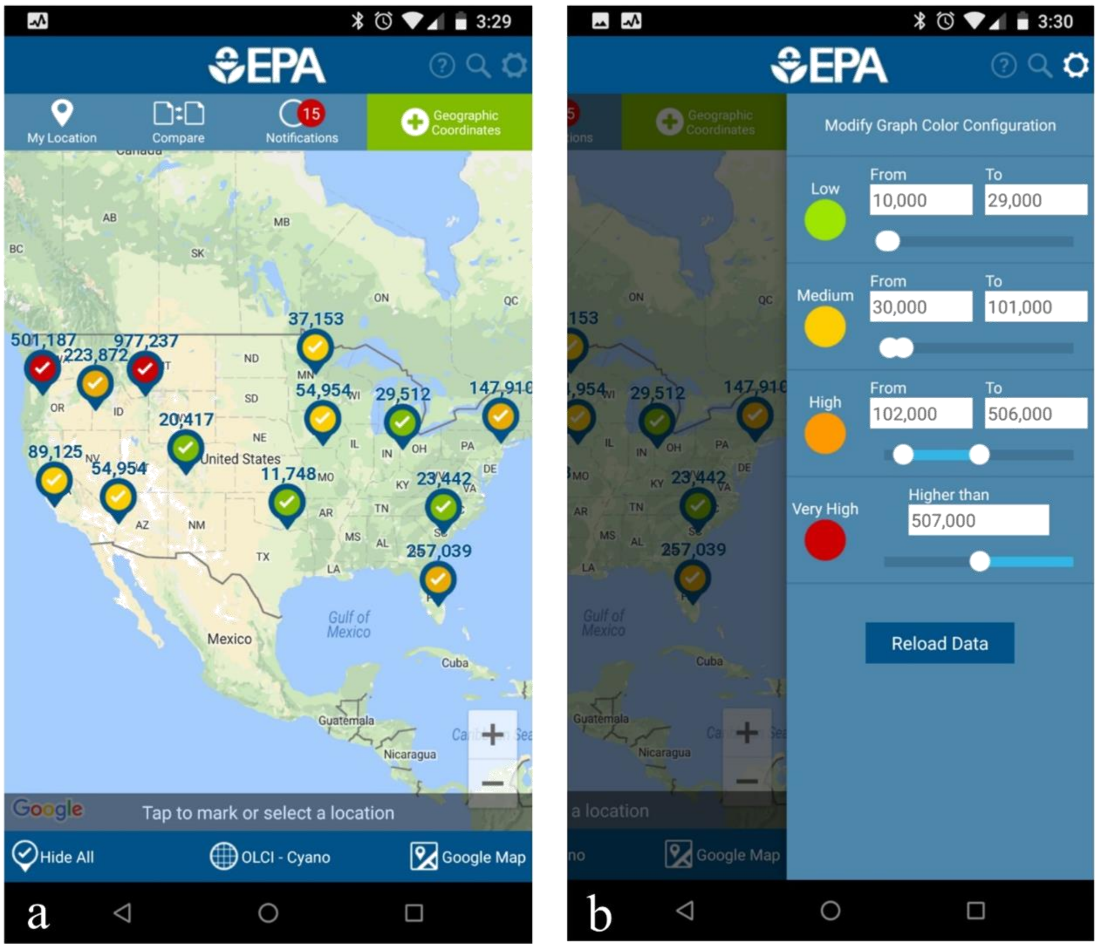

Fig. 3. (a) Main splash page of CyAN app for dropping pin locations and navigating to the My Location, Compare, Notification, and Geographic Coordinates tabs. (b) Side swipe the tab, or select the cogwheel at the top right, to alter the pin color thresholds based on user criteria.

4.2 Case 2 - Locations

Locations are stored in a list for quick and easy comparison to visualize the current cyanoHAB abundance value and the change from the previously reported time step (Fig 4a). The user can visualize latitude and longitude coordinates of locations, location names, cyanoHAB values, and recent changes. Selecting a listed location allows the user to view the marker pin on the map, the satellite data origin, and imagery thumbnails. The end-user first encounters the satellite images after selecting My Location and clicking on the desired location within the list (Fig 4b).

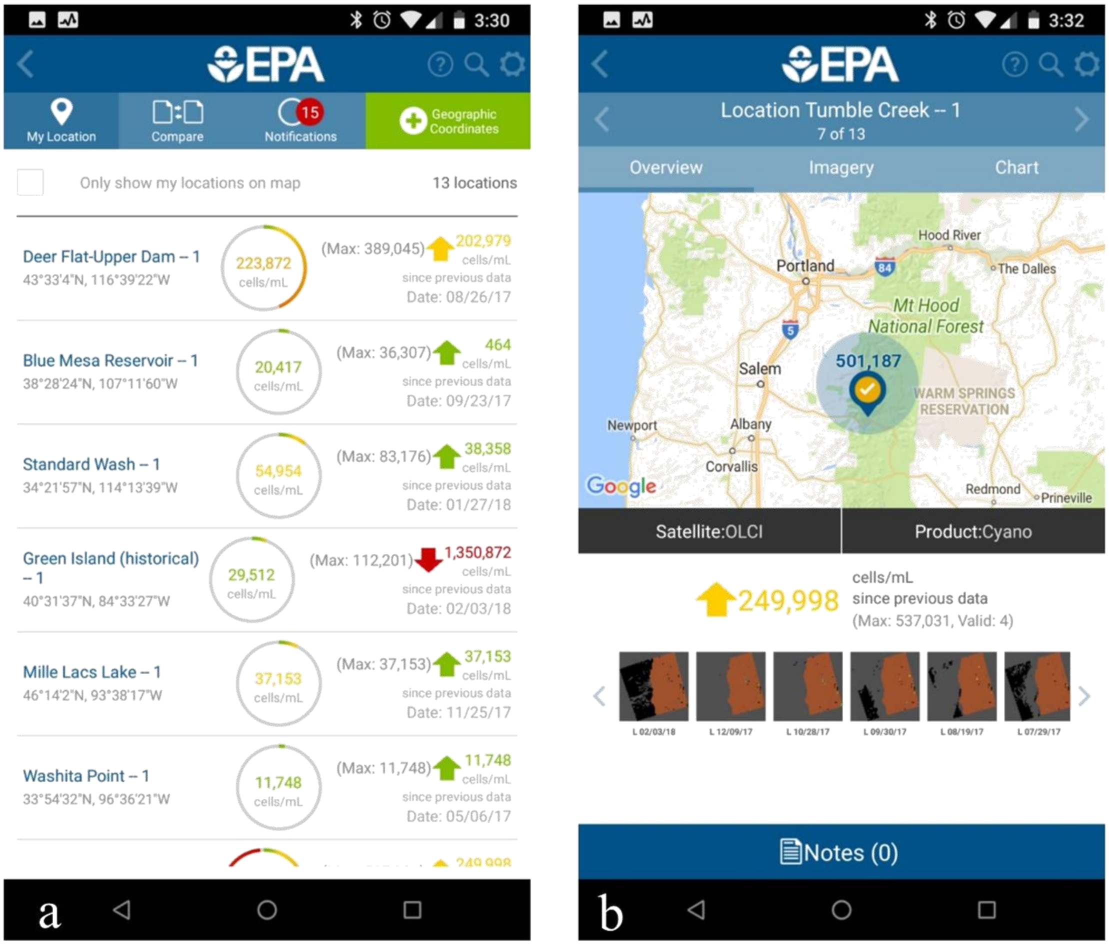

Fig. 4. (a) My Location tab of CyAN app, with storage of all set locations from the user. (b) Selection of a location in the My Location tab allows the user to visualize the thumbnail archive of Sentinel-3 satellite imagery.

The end-user may select any satellite file that contains their locations of interest to view the entire satellite tile (Fig. 5a). The CyAN app allows the satellite tile to be downloaded to the mobile device as a PNG file for record-keeping, or for quick viewing in the app for spatial information. The Imagery subtab allows a user to filter the image list by satellite instrument, filter the image list by date, and select image(s) for overlay. Users can overlay images on the base map, adjust the opacity of overlays, and pan or zoom within image(s). The Chart subtab permits users to view a time-series plot of selected locations with supported time frames.

4.3 Case 3 - Comparison

The Comparison tab of the CyAN app allows end-users to temporally view changes in cyanoHAB abundance across multiple locations of interest for an annual rolling time period (Fig. 5b). This analysis requires selecting two or more locations for comparison as described under Case #1. The user is provided with a list of selected locations, names, latitude and longitude coordinates, current cyanoHAB abundance, and recent changes. The Compare Statistics tab includes the areal maximum value of the pixel location and pixel count (3 × 3-pixel maximum cyanobacterial abundance) and abundance delta value. The Blooming Chart subtab provides a time-series plot of cyanobacterial abundance for selected locations.

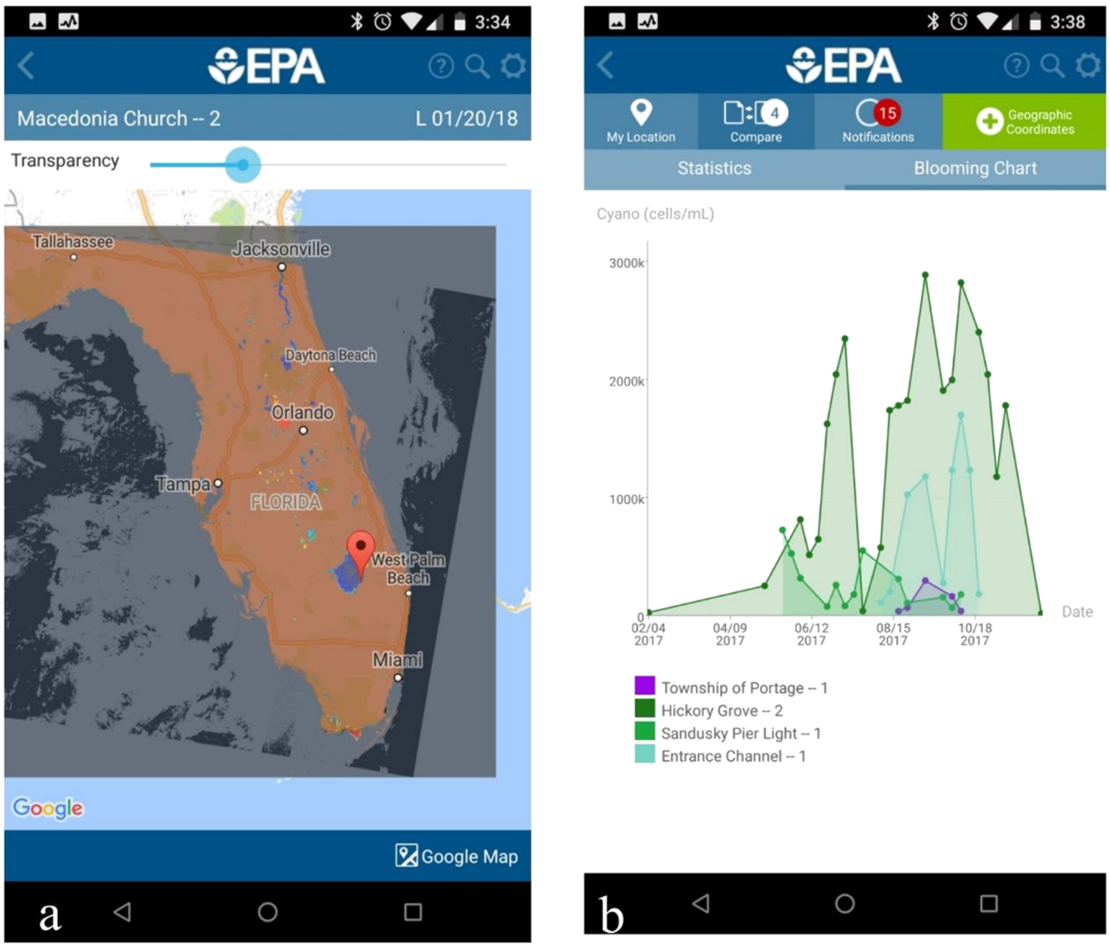

Fig. 5. (a) Selecting a satellite thumbnail image allows the user to visualize the complete satellite tile for the location of interest for spatial patterns. (b) Example of selecting the Compare tab allows the user to visualize temporal comparisons amongst different locations. Each line graph represents a single pixel that contains the geographic coordinate.

5.0 Results and Discussion

App requirements were initially defined from interviews of water quality managers to understand basic end-user limitations and needs (Schaeffer et al., 2013). Starting with these user requirements, the Agile method was implemented during architectural conceptualization and design. Ideation was used for wireframes, prototype testing, and final software development. Software was developed through a series of crowd-sourced competitions. Coders were provided a list of software technologies to use, assembly component diagrams, and class and sequence diagrams and requirements. An ongoing agile process is used to modify the app based on user feedback. Jira and Confluence project management software along with a branched GitHub software repository, daily scrums, biweekly sprint reviews, and frequent releases based on user feedback were used to guide further development and enhancement. Resource utilization is monitored to gather usage statistics and gauge application performance. Average operational statistics for REST API call metrics were GET_IMAGE at 250.45 milliseconds (ms), GET_LOCATION_DATA at 208.49 ms, GET_LOCATION_IMAGES at 3,597.29 ms for up to 12 images, GET_NOTIFICATIONS at 1,698.67 ms, and POST_APP_DATA at 1,361.85 ms.

App data are quality checked separately using GIS software. The mobile application meets National Institute of Standards and Technology (NIST) production environment standards 800-53 Revision 4 security controls and assessment procedures for Federal Information Systems and Organizations (https://nvd.nist.gov/800-53/Rev4/control/SI-2), and Information Directive Policy CIO 2150.4 to provide security for information and information systems (https://www.epa.gov/sites/production/files/2017-06/documents/information-security-policy.pdf) within the EPA National Computer Center including quarterly system and component patching. The app has been beta tested since June 2017 and currently has users in 12 EPA offices, US Army Corps of Engineers, and approximately 16 state environmental and health departments. The app was also used to deliver satellite data on Lake Okeechobee from June through October 2017 where multiple stakeholders desired access to the same satellite data. A sample of beta tester generalized app comments are provided in Table S2.

Satellite data files were converted from netCDF to GeoTIFF to compress file sizes and tiled into equal area sections containing 2,000 columns and rows of 300 × 300 m pixels covering a 600 × 600 km distance. These tiles are based off existing Landsat scene tiles so they may be nested (Fig. 6). Backend average processing time is 1.3 minutes per tile, for 37 tiles across the U.S., totaling 47.54 minutes, which included image validation, data extraction, database population and ancillary imagery generation. Average tile file sizes ranged from 82 KB to 422 KB depending on the number of US lakes and water pixels within each tile (Fig. 7). Tiles used 2,500MB with 11 months of OLCI weekly data for 37 tiles, and the image directory used 8.0 GB or 7% of filesystem. The interval time between a satellite acquisition and the app user accessing the data is typically three days. For example, NASA processed daily images Sunday through Saturday, for each satellite acquisition day, between January 1, 2017 through January 7, 2017. A 7-day weekly composite was created and posted by Tuesday the following week, in the example case by January 10, 2017. The data were uploaded to the app the following day and delivered to the app.

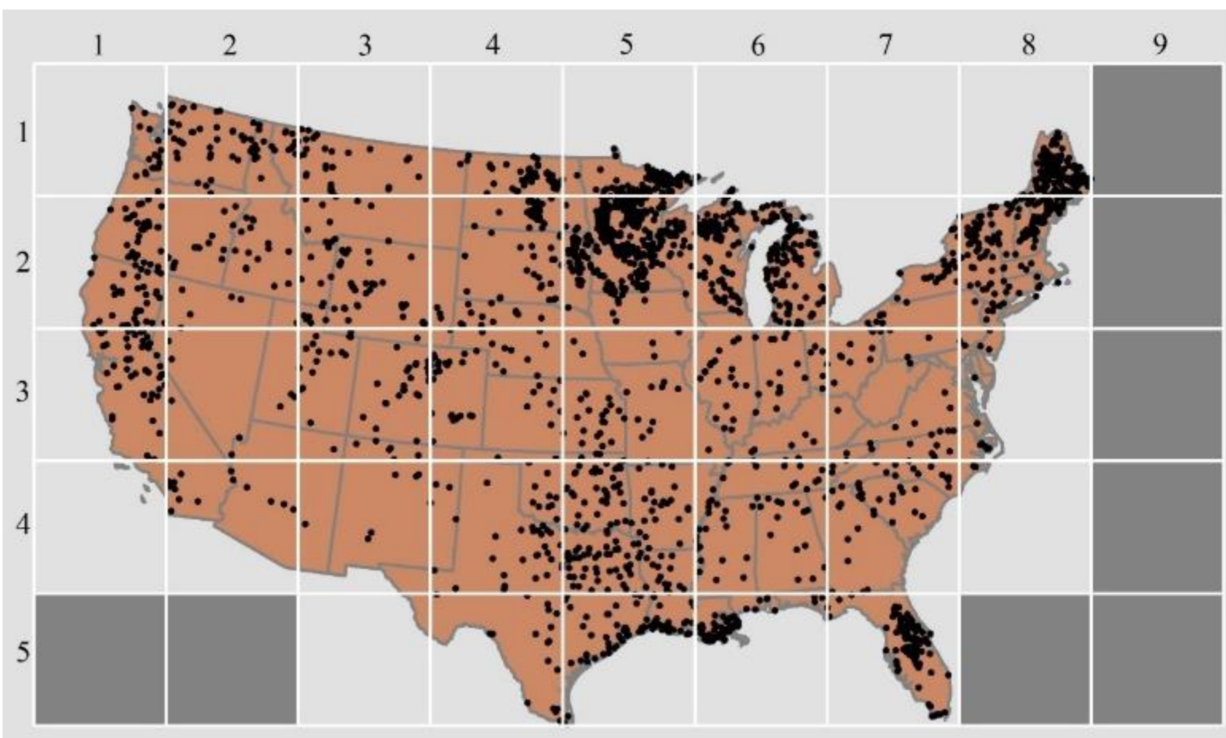

Fig. 6. Map of continental U.S. (CONUS) grid tiles developed for OLCI processing and location of NHDPlusV2 resolvable lakes (Table S1) in each state (black points).

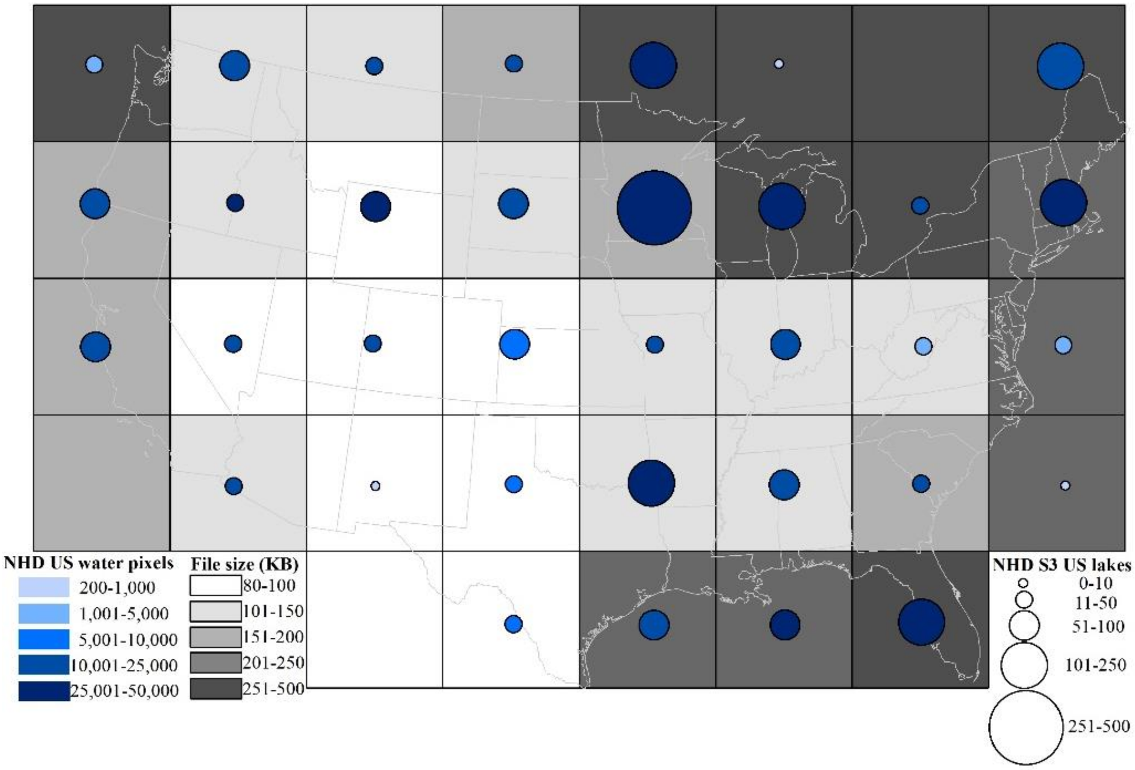

Fig. 7. Map of CONUS grid tiles with number of OLCI water pixels, number of NHDPlusV2 resolvable lakes (Table S1) per tile and file size for each tile.

App functionality and satellite data were validated against 25 state health advisories issued in 2017 across seven states (Table 1). Overall, the GIS extracted values replicated app data and more importantly correctly identified cyanoHAB events during the same periods of time as the reported state health advisories (Fig. 8 and 9; Figs. S1-S12). The date (x-axis) on the app temporal graph used the end date of the 7-day image and therefore the validation exercise confirmed that 16 of the 25 events (64%) would have been identified at least one week earlier with the CyAN app and satellite data. Evidence of early detection was specifically demonstrated in a case study with Utah Lake. Utah Lake routine monthly sampling occurred on June 12, 2017 and did not identify any elevated cyanoHABs. However, the same satellite data set, used by the app, indicated increased cyanoHAB abundance the week ending June 24, 2017 (Fig. S13). The site was revisited for additional sampling the following week to confirm that a cyanoHAB had developed in Provo Bay and a warning advisory was issued. The Utah Lake advisory specifically used the satellite imagery in developing the public advisory stating, "The bloom was first identified in Provo Bay via satellite imagery" (https://documents.deq.utah.gov/communication-office/press-releases/2017-06-29-Bloom-Provo-

Bay.pdf). To clarify the CyAN app was not used in this specific case, but the Sentinel-3 OLCI satellite data were used, and the event was replicated with the app retrospectively as a demonstration. This case study demonstrated how the satellite data and the app assist water quality managers who need to identify toxic blooms or taste and odor issues by targeting specific locations within lakes at specific times.

Table 1. State and satellite resolvable waterbodies from state cyanoHAB advisor lists. Location information entered into the CyAN app Geographic Coordinates tab and associated site name the app provides from the Google Map API. Associated validation figures are also listed.

State	State Lake	CyAN App site	Latitude	Longitude	Fig.
OR	Upper Klamath	Upper Klamath Lake	42° 24’ 27.435”	-121° 53’ 30.7314”	8-9
OR	Odell	Odell Lake	43° 34' 26.439"	-121° 59' 41.5968"	8-9
OR	Drews	Drews Reservoir	42° 10' 25.23"	-120° 40' 8.1798"	8-9
OR	Detroit	Tumble Creek	44° 43' 5.8506"	-122° 11' 6.6696"	8-9
CA	Havasu	Standard Wash	34° 21' 56.739"	-114° 13' 38.784"	S1-S2
CA	Black Butte	Black Butte	39° 48' 8.4486"	-122° 21' 34.1634"	S1-S2
CA	San Antonio	Harris Creek	35° 48' 45.8634"	-120° 55' 51.708"	S1-S2
ID	Lake Lowell	Deer Flat-Upper Dam	43° 33' 4"	-116° 39' 22"	S3-S4
ID	Henry's Reservoir	Duck Creek	44° 37' 37"	-111° 24' 44'	S3-S4
NJ	Wanaque Reservoir	Wolf Den Dam	41° 2' 49"	-74° 18' 39"	S5-S6
NY	Allegheny	Pierce Run	42° 2' 23.7408"	-78° 55' 44.4576"	S7-S8
NY	Orange	Orange Lake	41° 32' 57.6378"	-74° 6' 11.0442"	S7-S8
NY	Honeoye	Willow Beach	42° 44' 42.6624"	-77° 30' 46.5294"	S7-S8
NY	Neatahwanta	Lake Neatahwanta	43° 18' 28.9188"	-76° 25' 53.0328"	S7-S8
NY	Chautauqua	Loomises	42° 7' 16.086"	-79° 21' 28.0512"	S7-S8
UT	Deer Creek Reservoir	Deer Creek Reservoir	40° 26' 47"	-111° 29' 14"	S9-S10
UT	Mantua Reservoir	Mantua Reservoir	41° 30' 17"	-111° 56' 8"	S9-S10
UT	Rockport Reservoir	Kent Canyon	40° 46' 46"	-111° 24' 9"	S9-S10
UT	Strawberry Reservoir	Horse Creek	40° 7' 53"	-111° 8' 7"	S9-S10
UT	Utah Lake	Provo Municipal Airport	40° 11' 36"	-111° 43' 60"	S9-S10

VT	Lake Carmi	Lake Carmi	44º 58' 13"	-72º 52' 46"	S11-S12
VT	Lake Memphremagog	Whipple Point	44º 57' 23"	-72º 13' 37"	S11-S12
VT	Mallets Bay	Camp Norfleet	44º 35' 3"	-73º 13' 44"	S11-S12
VT	Mississquoi Bay	Rock River Bay	44º 59' 32"	-73º 5' 53"	S11-S12
VT	St. Albans Bay	Mill River	44º 47' 50"	-73º 8' 54"	S11-S12

Fig. 8. Screen shots of mobile CyAN mobile app time series charts for (a) Upper Klamath Lake, (b) Odell Lake, (c) Drews Reservoir, and (d) Detroit Lake. Each line graph represents a single pixel that contains the geographic coordinate listed in Table 1. Time series screen shots correspond to GIS comparisons in Fig. 9 and Oregon cyanoHAB advisories.

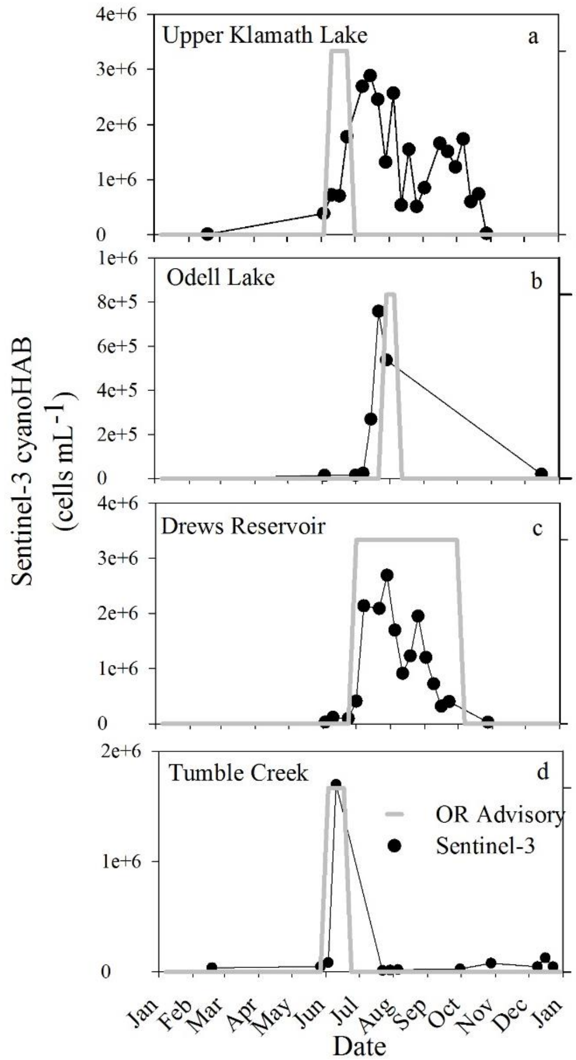

Fig. 9. GIS comparison of satellite derived CI cyanoHAB abundance results with Oregon bloom advisory time series for (a) Upper Klamath Lake, (b) Odell Lake, (c) Drews Reservoir, and (d) Tumble Creek in Detroit Lake.

6.0 Conclusions

Visualization of large scale satellite data was achieved through the design of a software system for a specific data set and limited number of query types (Godfrey et al., 2016). Servers provide the data to the CyAN app accessible through an Android mobile device such as a smartphone, reducing the computational requirements for managers (Khan et al., 2014). The CyAN app represents the first attempt to make satellite-derived cyanoHAB data directly available to water quality managers so that they can make timely decisions based on changes in cyanobacterial abundance. Drinking water utilities, waste-water utilities, recreational water organizations, and scientific ecological research entities may benefit from this software tool. The

CyAN app supports management needs to quantify cyanoHAB occurrence information relevant to the national Safe Drinking Water Act, Clean Water Act, and Harmful Algal Bloom and Hypoxia Research Control Act in the United States.

Unlike previous software packages, the CyAN app lowers the barrier of entry and necessary programmatic support for using satellite data to make water quality management decisions. The CyAN app allows for the democratization of satellite-derived water quality data (Sawicki and Craig, 1996), otherwise difficult to achieve using existing netCDF and HDF file formats that require specialized knowledge for data access and analysis. The CyAN app is the first of its kind to provide a cost-effective delivery system for satellite-derived cyanoHAB data products to water quality managers in a simple display format that is intuitive and applicable across multiple water body locations both for recreational and drinking water sources.

Acknowledgements

This work was partially supported by the NASA Ocean Biology and Biogeochemistry Program/Applied Sciences Program (proposal 14-SMDUNSOL14- 0001), U.S. EPA, and NASA Applied Science Program announcement NNH08ZDA001N (Human Health and Air Quality) under contract NNH09AL531. Satellite data is from Copernicus Sentinel data processed by the European Space Agency. Initial startup funds for this project were provided by the Office of Research and Development Pathfinder Innovation Projects I and II. Hosting costs were supported by the U.S. EPA Office of Science Information Management (OSIM). This article has been reviewed by the National Exposure Research Laboratory and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use by the U.S. Government. The views expressed in this article are those of the authors and do not necessarily reflect the views or policies of the U.S. EPA. NOAA and NASA GSFC provided satellite geoTIF source files. The authors would like to acknowledge the international crowd-sourced community; Rakshit Trivedi, Rashid Sial, and Jennifer Odess at Apirio; Steven Rader at NASA's Center of Excellence for Collaborative Innovation; and Mark S. Murphy and Torrin Hulgren at Innovate! for their assistance and technical support.

References

Ariely, D., 2010. The Upside of Irrationality: The Unexpected Benefits of Defying Logic at Work and at Home. . Harper Collins Publishers, New York, NY.

Baith, K., Lindsay, R., Fu, G., McClain, C.R., 2001. SeaDAS: Data analysis system developed for ocean color satellite sensors. EOS, Transactions, American Geophysical Union 82(18) 202.

Berger, M., Moreno, J., Johannessen, J.A., Levelt, P.F., Hanssen, R.F., 2012. ESA's sentinel missions in support of Earth system science. Remote Sensing of Environment 120 84-90.

Carroll, M.L., Townshend, J.R., DiMiceli, C.M., Noojipady, P., Sohlberg, R.A., 2009. A new global raster water mask at 250 m resolution. International Journal of Digital Earth 2(4) 291-308.

Clark, J.M., Schaeffer, B.A., Darling, J.A., Urquhart, E.A., Johnston, J.M., Ignatius, A., Myer, M.H., Loftin, K.A., Werdell, P.J., Stumpf, R.P., 2017. Satellite monitoring of cyanobacterial harmful algal bloom frequency in recreational waters and drinking water sources. Ecological Indicators 80 84-95.

Coad, P., Cathers, B., Ball, J.E., Kadluczka, R., 2014. Proactive management of estuarine algal blooms using an automated monitoring buoy coupled with an artificial neural network. Environmental Modelling and Software. 61 393-409.

Dodds, W.K., Bouska, W.W., Eitzmann, J.L., Pilger, T.J., Pitts, K.L., Riley, A.J., Schloesser, J.T., Thornbrugh, D.J., 2009. Eutrophication of U.S. freshwaters: analysis of potential economic damages. Environmental Science and Technology 43(1) 12-19.

Donlon, C., Berruti, B., Buongiorno, A., Ferreira, M.-H., Féménias, P., Frerick, J., Goryl, P., Klein, U., Laur, H., Mavrocordatos, C., Nieke, J., Rebhan, H., Seitz, B., Stroede, J., Sciarra, R., 2012. The global monitoring for environment and security (GMES) Sentinel-3 mission. Remote Sensing of Environment 120 37-57.

Godfrey, P., Gryz, J., Lasek, P., 2016. Interactive visualization of large data sets. IEEE Transactions of Knowledge and Data Engineering 28(8) 2142-2157.

Graham, J.L., Loftin, K.A., Kamman, N., 2009. Monitoring recreational freshwaters. North American Lake Management Society: Lakeline Summer 18-24.

Grant, G.E., Gallaher, D.W., 2015. A novel technique for time-centric analysis of massive remotely-sensed datasets. Remote Sensing 7 3986-4001.

GSM Association, 2016a. The Mobile Economy 2016. GSMA Head Office: London, United Kingdom.

GSM Association, 2016b. The Mobile Economy: North America 2016. GSMA Head Office: London, United Kingdom.

ICT, 2016. ICT Facts and Figures. International Telecommunication Union: Geneva, Switzerland, p. 8.

IDC, I.D.C., 2017. Smartphone OS, Market Share: http://www.idc.com/promo/smartphone-market-share/os.

Jonoski, A., Almoradie, A., Khan, K., Popescu, I., Andel, S.J.v., 2013. Google Android mobile phone applications for water quality information management. Journal of Hydroinformatics 15(4) 1137-1149.

Khan, A.R., Othman, M., Madani, S.A., Khan, S., 2014. A survey of mobile cloud computing application models. IEEE Communications Surveys and Tutorials 16(1) 393-413.

Kim, Y., Shin, H.S., Plummer, J.D., 2014. A wavelet-based autoregressive fuzzy model for forecasting algal blooms. Environmental Modelling and Software. 62 1-10.

Kutser, T., 2009. Passive optical remote sensing of cyanobacteria and other intense phytoplankton blooms in coastal and inland waters. International Journal of Remote Sensing 30(4401-4425).

Loftin, K.A., Graham, J.L., Hilborn, E.D., Lehmann, S.C., Meyer, M.T., Dietze, J.E., Griffith, C.B., 2016. Cyanotoxins in inland lakes of the United States: occurrence and potential recreational health risks in the EPA National Lakes Assessment 2007. Harmful Algae 56 77-90.

Lunetta, R.S., Schaeffer, B.A., Stumpf, R.P., Keith, D., Jacobs, S.A., Murphy, M.S., 2015. Evaluation of cyanobacteria cell count detection derived from MERIS imagery across the eastern USA. Remote Sensing of Environment 157 24-34.

Malve, O., Laine, M., Haario, H., Kirkkala, T., Sarvala, J., 2007. Bayesian modelling of algal mass occurrencesdusing adaptive MCMC methods with a lake water quality model. Environmental Modelling and Software. 22 966-977.

McKay, L., Bondelid, T., Dewald, T., Johnston, J., Moore, R., Rea, A., 2012. NHDPlus Version2: User Guide.

Norton, L., Elliott, J.A., Maberly, S.C., May, L., 2012. Using models to bridge the gap between land use and algal blooms: An example from the Loweswater catchment, UK. Environmental Modelling and Software. 36 64-75.

Sadeghian, A., Chapra, S.C., Hudson, J., Wheater, H., Lindenschmidt, K.E., 2018. Improving in-lake water quality modeling using variable chlorophyll a/algal biomass ratios. Environmental Modelling and Software. 101 73-85.

Sawicki, D.S., Craig, W.J., 1996. The democratization of data: bridging the gap for community groups. Journal of American Planning Association 62(4) 512-523.

Schaeffer, B.A., Schaeffer, K.G., Keith, D., Lunetta, R.S., Conmy, R., Gould, R.W., 2013. Barriers to adopting satellite remote sensing for water quality management. International Journal of Remote Sensing 34(21) 7534-7544.

Smayda, T.J., 1997. What is a bloom? A commentary. Limnology and Oceanography 42 1132-1136.

Steffensen, D.A., 2008. Economic cost of cyanobacterial blooms, In: Hudnell, H.K. (Ed.), Cyanobacterial Harmful Algal Blooms: State of the Science and Research Needs. Springer: New York, pp. 855-865.

Stumpf, R.P., Davis, T.W., Wynne, T.T., Graham, J.L., Loftin, K.A., Johengen, T.H., Gossiaux, D., Palladino, D., Burtner, A., 2016. Challenges for mapping cyanotoxin patterns from remote sensing of cyanobacteria. Harmful Algae 54 160-173.

Teng, W., Rui, H., Strub, R., Vollmer, B., 2016. Optimal reorganization of nasa earth science data for enhanced accessibility and usability for the hydrology community. Journal of the American Water Resources Association 52(4) 825-835.

Tomlinson, M.C., Stumpf, R.P., Wynne, T.T., Dupuy, D., Burks, R., Hendrickson, J., Fulton, R.S., 2016. Relating chlorophyll from cyanobacteria-dominated inland waters to a MERIS bloom index. Remote Sensing Letters 7(2) 141-149.

U.S. EPA, 2013. The importance of water to the U.S. economy: Synthesis report. Office of Water: Washington, DC, p. 29.

U.S. EPA, 2015. Drinking water health advisory for the cyanobacterial microcystin toxins. Office of Water, EPA-820R15100: Washington, DC, p. 67.

Urquhart, E.A., Schaeffer, B.A., Stumpf, R.P., Loftin, K.A., Werdell, P.J., 2017. A method for monitoring cyanobacterial harmful algal bloom spatial extent using satellite remote sensing data. Harmful Algae 67 144-152.

Wynne, T.T., Stumpf, R.P., Tomlinson, M.C., Dybleb, J., 2010. Characterizing a cyanobacterial bloom in western Lake Erie using satellite imagery and meteorological data. Limnology and Oceanography 55(5) 2025-2036.

Wynne, T.T., Stumpf, R.P., Tomlinson, M.C., Warner, R.A., Tester, P.A., Dyble, J., Fahnensteil, G.L., 2008. Relating spectral shape to cyanobacterial blooms in the Laurentian Great Lakes. International Journal of Remote Sensing 29 3665-3672.

Zhang, H., Hu, W., Gu, K., Li, Q., Zheng, D., Zhai, S., 2013. An improved ecological model and software for short-term algal bloom forecasting. Environmental Modelling and Software. 48 152-162.