Chesapeake Bay Water Quality and Living Resource Data Visualization

Abhishek Abhishek, James Guan, and John Wolf

Abstract—The Chesapeake Bay is one of the most studied estuaries in the world. This project leverages a long term water quality data set and habitat requirements for various Chesapeake Bay living resources in an interactive visualization through three-dimenionsl space and time. The corresponding visualization provides a regional view of long-term averages in water quality parameters and habitat requirements of fish and shellfish; coupled with a station-by-station view for water quality data for user-defined time periods. The results show the potential for using interactivity to communicate key place-based living resource stories associated with water quality data observations.

Index Terms— Chesapeake Bay, visualization, water quality, habitat, living resources, monitoring.

1 Introduction and Motivation

The Chesapeake Bay is one of the most studied ecosystems in the world. The Chesapeake Bay Program (Chesapeake Bay Program 2016a), which oversees the Chesapeake restoration effort, maintains over 30 years of water quality monitoring data that has been collected throughout 3D space and time. This water quality monitoring data is routinely used in scientific investigations as well as communicating water quality conditions for regulatory purposes.

However, the water quality data is currently accessed from a website with minimal visualization capabilities. It is up to the user to download data and then bring it into software of their choice to process and visualize. In addition, the website does not currently provide the option to interpret conditions relative to fish and shellfish habitat requirements. These habitat requirements relate to various water quality parameters such as dissolved oxygen (DO), salinity, and chlorophyll a (Funderburk et al, 1991).

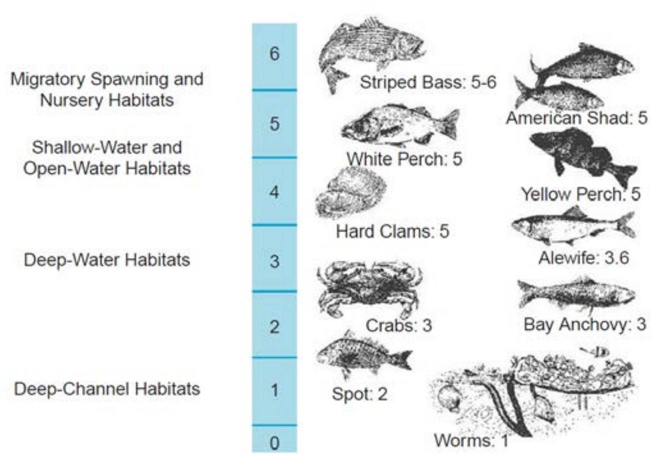

Figure 1. Example of dissolved oxygen requirements of select Chesapeake Bay living resources. (Chesapeake Bay Program 2016b).

2 Goals and Research Questions

The overall goal of this project is to provide a visual representation of water quality data at user-specified locations and to show how those water quality measurements support living resources in the Bay. For this proof of concept, water quality data from the Chesapeake Bay monitoring program's sample locations for the year 2015 were utilized.

The approach leverages the visualization mantra of "overview first, zoom and filter, and details on demand" (Shneiderman 1996). This project serves as a proof of concept for visualizing the water quality data in a manner that is directly relatable to water column and bottom habitat requirements. The application provides the capability to explore water quality data through a geographic interface that enables the user to select a geographic location and time period of interest. That information is then capable of being related to habitat requirements of various Chesapeake Bay living resources.

3 Literature Review/Related Work

Attempts at three-dimensional visualization of water quality data are not new. Forgang et al (1996) developed a visualization system for the comparison of simulated and measured water quality (SCIRT). This was followed by a more robust tool that visualized output from the Chesapeake Bay Program's three-dimensional eutrophication model (Stein et al 2000). These efforts, however, focused on visualizing 3D model output as opposed to monitoring data.

Estuary monitoring data is typically four dimensional (3D space and time), but is collected at discrete sample location (as opposed to water quality model output that is frequently continuous). The depth and temporal dimensions of monitoring data present challenges for 2D or 2.5D space.

Shisko et al (2010) presented IGODS, a "complete software solution to support studies of ocean, coastal and inland waters". IGODS allows for the visualization of water quality parameters along depth profiles at multiple locations. Since it focuses on monitoring data as opposed to modeling data, the Shisko paper provided much of the inspiration for the current project, which customizes the IGODS approach for the unique ecological characteristics of the Chesapeake Bay.

Visualizing living resource habitat in estuaries and river system hasn't progressed as far as water quality. An interesting application that combines fish species requirements with a virtual reality game engine is presented by Chen et al (2014). The ability to simulate and visualize actual fish ecology through virtual reality holds much progress in the future integration of water quality and fish habitat visualization efforts.

Leveraging Google Maps as an interactive locator tool has been used in a number of environmental applications. One example that is closely aligned with the intent of this project is the HydroQual application developed by Accorsi et al (2014).

The integration of water resource visualization systems and geographic information systems has also been a focus of recent research and development. Kim and Park (2012) demonstrate how environmental changes can be visualized through a GIS-based visual analytics tool. Their application of this tool for location-based oceanographic data demonstrates the values of various visual encoding approaches and user-specified filters.

4 Design Methods

Requirements and sample data have been gathered by Chesapeake Bay Program data managers. The authors developed a Google Maps front end with water quality sample locations identified by visual glyphs on top of various map layers displayed on the Google base map. The map layer choices were generated in the geographic information system ArcGIS software package and exported to keyhole mark-up language (KML) for integration with Google Maps. The JavaScript D3 library was used to provide interactivity with the symbols on the base map. Once a sample location is selected, interactive charts provide detailed information on water quality parameters (DO, salinity, water temperature and chlorophyll a) at that site for a time period defined by the user. The site-specific charts (scatter plots and line charts) were also coded using the D3 JavaScript library.

4.1 Data Abstraction

For the proof of concept, 2015 data was downloaded in csv format from the Chesapeake Bay Data Hub. Data abstraction is as follows:

Item → Station ID
Attribute → Date (temporal)

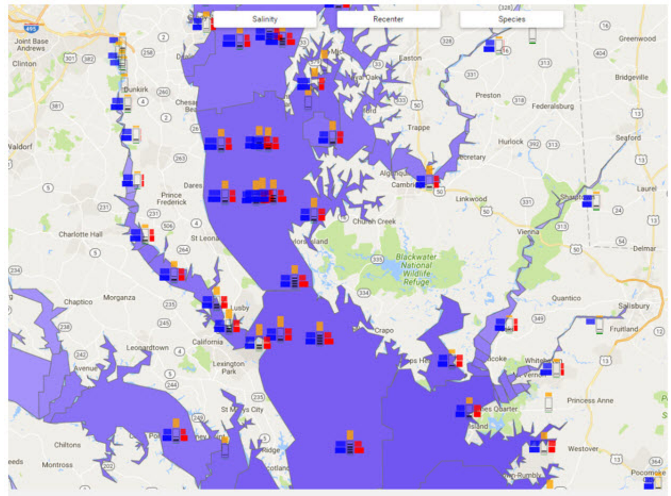

Figure 2. Regional map of Chesapeake Bay with water quality monitoring station glyphs.

Attribute → Layer (categorical)
Attribute → Parameter (categorical)
Attribute → Value (quantitative)
Attribute → Species (categorical)
Spatial → Latitude
Spatial → Longitude
Spatial → Depth

4.2 Task Abstraction

Task abstraction focuses on the actions and targets of a given visualization (Munzner 2014). With respect to Actions, this project focuses on the **analysis action** of consumption of information. The user is primarily interested in the discovery of information through scientific inquiry. To some extent the visualization may also benefit the user in the presentation of information.

The **search action** focuses on whether the identity and target of the visualization is known in advance. This project has elements of both known and unknown locations and targets. For example, the user may know in advance what water quality station they are interested in and then proceeds directly to that station's information. Alternatively, the user may leverage the overview panel to help identify the location (water quality station) of interest. Similarly, the user may know the target parameters to evaluate, or they may be more interested in exploring or browsing various targets under different hypothetical scenarios.

The **query action** focuses on the identification, comparison or summarization of data. Again, this visualization has elements of all three. The visualization facilitates the identification of information at a single user-defined target, including patterns in water quality parameters at a pre-selected location. The visualization enables comparisons among multiple water quality parameters at both regional and station-specific scales.

The targets of the visualization include both spatial and non-spatial variables. The scatter plots enable the user to view trends in water quality data by depth during specific time intervals. The presentation includes multiple attributes (water quality and living resource considerations) in three-dimensional space.

4.3 Design Considerations

As previously indicated, the authors chose to follow the "overview first, zoom and filter, and details on demand" approach. Initially, the overview map provided a regional view of the water quality monitoring network with the point locations of monitoring stations identified by small circles with color saturation reflecting long-term mean summer salinity values at each station. Based on the results of the prototype presentation and subsequent project evaluation, additional variables were visually encoded for each water quality monitoring station. This was achieved through the development of a custom glyph so that the user could better visualize differences among stations on the regional map.

The glyph design visually encoded annual average values for seven quantitative parameters at each monitoring stations:

- Total depth
- Mean water temperature (surface)
- Mean water temperature (bottom)
- Mean annual salinity (surface)
- Mean annual salinity (bottom)
- Secchi depth
- Chlorophyll a

With the exception of total depth (which was visually encoded with a series of thin horizontal bars corresponding the magnitude of total depth at that station) each of the variables was encoded with a horizontal or vertical bar – the length of which corresponded to the value of the parameter. The distinction among the categories of water quality parameters was visually encoded through variations in hue as well as pattern and orientation of the bars.

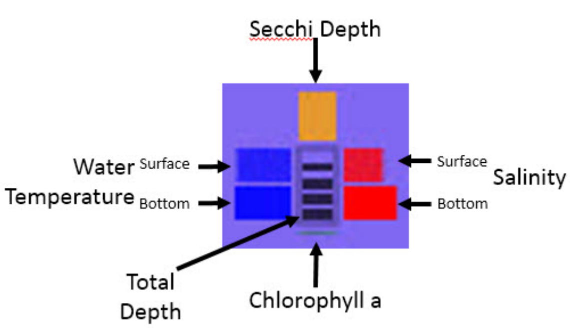

Figure 3. Design of glyphs to encode multiple visual attributes.

The regional map also provides the ability for the user to select species-specific base maps that help to illustrate the expected distribution of fish and shellfish at various life stages. These map layers (which are based on species life cycle maps contained in Funderburk et al 1991) are displayed on top of the Google base map but underneath the water quality station glyphs. As a result, additional insights can be gleaned through the juxtaposition of regional living resource and water quality parameters.

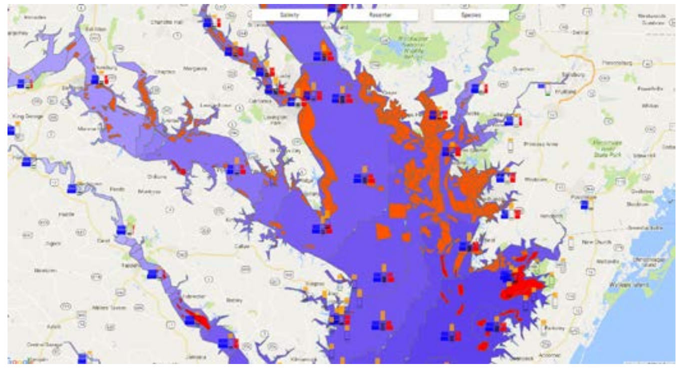

Figure 4. Eastern Oyster habitat in proximity to water quality monitoring stations.

The user can obtain more detailed information by selecting a specific station. When selected, additional filtering and chart options are exposed to the right of the map. Unlike the regional depiction however, the user can now explore the water quality data within specified date ranges and at multiple depths. Date ranges are selected through a calendar interface with the user selecting a beginning and ending date. Generally speaking, water quality measurements at a given station are made once a year for most months, with bimonthly sampling often occurring during the summer. As a result, it is possible for the charts to return "no data" if the date range specified is too narrowly defined.

Once the date range is specified (date filter) the scatter plot updates with values for four quantitative variables - dissolved oxygen, chlorophyll a, salinity and water temperature. These parameters were selected due to their significance for evaluating whether water quality could support various life stages of species common to the Bay. In each case, the y axis represents depth. The resulting chart displays the four water quality parameters with differing x axes data ranges but with the consistent representation of depth on the y axis. Parameters are color coded to correspond to the colors of the x axis labels.

The values displayed on the scatter plot itself show the actual value of a parameter relative to depth through the symbol fill color. However, symbols are also encoded to illustrate whether a specific measurement is acceptable or unacceptable in supporting a particular living resource. The user can select a species and life stage and, based on their selection, the scatter plot values are depicted with either a red or green outline. This outline indicates whether the measurement at that location, depth, and date either supports (green outline) or does not support (red outline) conditions required of the specified species and life stage.

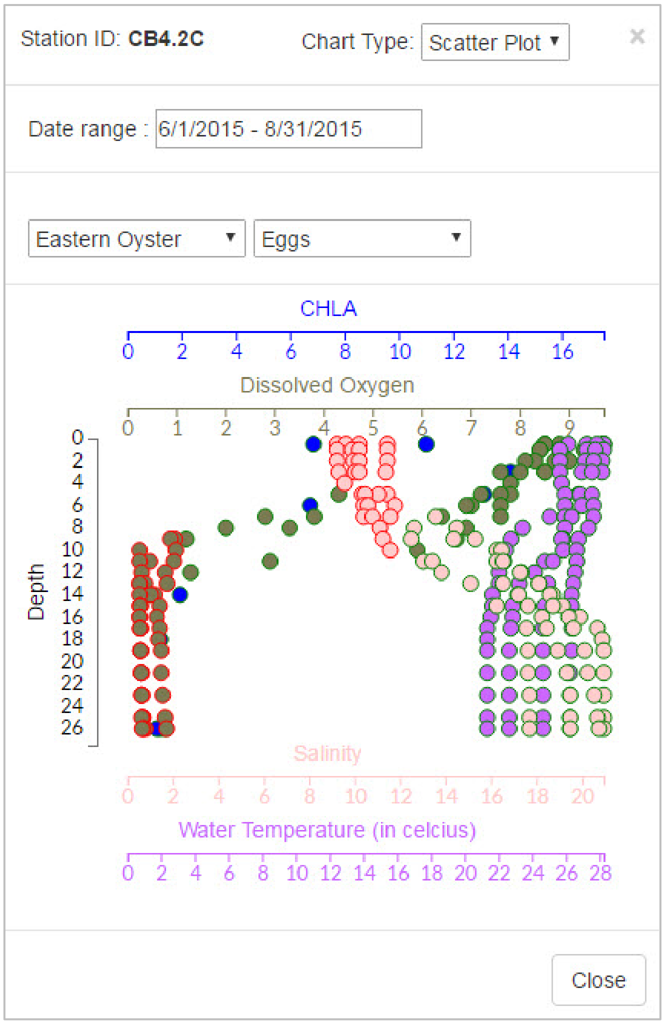

Figure 5 Station-specific portrayal of water quality data at multiple depths given user-selected time range. Color of circle outlines indicate suitability for selected species (red = bad; green = good).

Finally, the user can elect to display the data in the form of a line graph instead of a scatter plot. The line graph is based on the best-fit of the values represented in the scatter plot for a given water quality parameter at multiple depths.

5 Implementation Methods

The visualization was implemented through a combination of proprietary and open source technologies. The regional overview map leverages the Google Maps Application Programming Interface (API) along with map layers generated using the Environmental Systems Research Institute (ESRI) ArcGIS software. The Balsamic prototyping tool was used to develop paper prototypes for interface screen design. Glyphs were prototyped using Microsoft PowerPoint.

The glyphs on the regional overview map were built using the D3 (Data Driven Documents) JavaScript library. The charts depicting conditions at a specific water quality station were developed using JavaScript and D3.

6 Results and Analysis of Results

The visualization was developed iteratively during the course of the semester. The project prototype was presented in class on November 16, 2016. Minor adjustments were made based on feedback obtained after the presentation.

After speaking with the Class Instructor on November 23, additional refinements were implemented. The depth scale on the regional map was replaced to a multi-parameter glyph. The details chart was moved so that it displayed alongside the overview map as opposed to opening in a separate window. This was done in recognition of limitations in a viewer's memory and reducing the number of steps to get information effectively and efficiently.

The resulting visualization provides direct access to more information upon launching the application than the original design. The user needs fewer "clicks" to complete tasks identified in the task abstraction section of this paper.

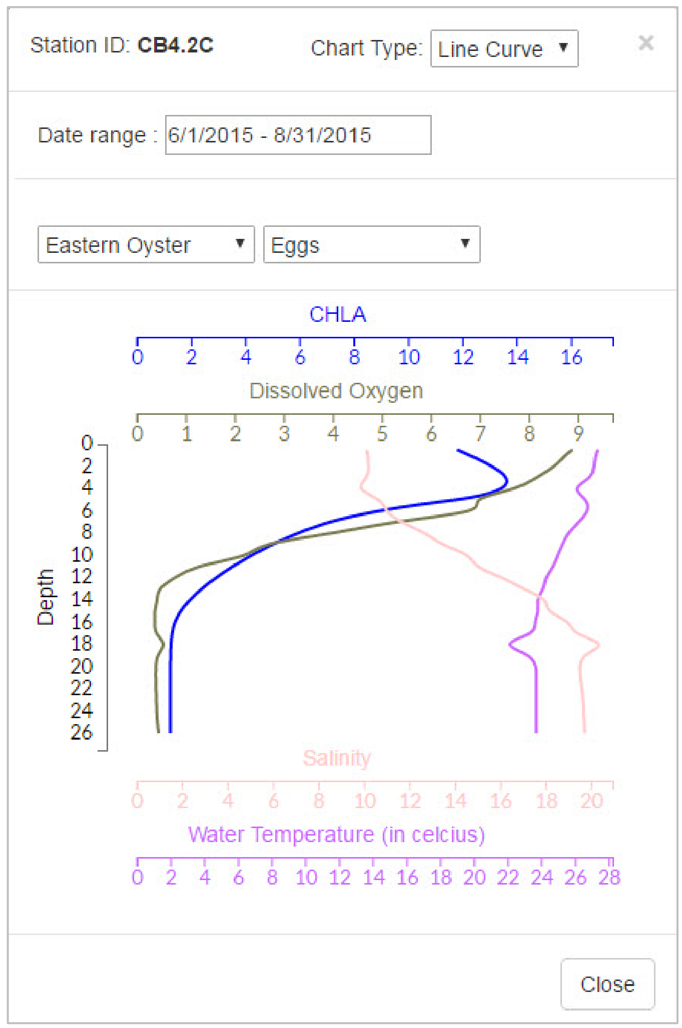

Figure 6. Line chart view of water quality parameters varying by depth.

7 Conclusions

This project involved developing a proof of concept to interactively visualize water quality for the Chesapeake Bay in 3D space and time and to relate that information to living resource habitat criteria. It was the first attempt to integrate that information using a combination of geographic information systems and open source visualization software libraries.

The project proceeded through a deliberate sequence of planning and design steps including data assessment, data abstraction, task abstraction, iterative design, testing, and usability assessment.

It was a useful initial step in the integration of water quality and living resource information, through a dynamic, data-driven application. Ideas for future enhancements include:

- Working with Chesapeake Bay Program data managers and other audiences to further refine requirements and user stories,
- Design improvements to glyphs used in regional overview map,
- Identify limitations and opportunities for a more robust exploration of water quality parameters at both regional and site-specific scales,
- Improvements and/or updates to living resource habitat maps,
- Direct connection to CBP Data Hub API instead of data extracts from data hub.

Acknowledgments

The authors would like to acknowledge Dr. Jian Chen for feedback on the project throughout the duration of the semester.

References

1) P. Accorsi, M. Fabregue, A. Sallaberry, F. Cernesson, N. Lalande, A. Braud, S. Bringay, F. Le Bar, P. Poncelet, and M. Teisseire. HydroQual: Visual Analysis of River Water Quality. IEEE Symposium on Visual Analytics Science and Technology 2014.
2) M. Chen, S. Lai, Y. Hung and J. Yang. A Virtual Education Assistance System in Simulated River Ecosystems with 3D Visualization. Proceedings of the IEEE Visualization Conference 2014.
3) Chesapeake Bay Program Overview. 2016a. Accessed November 4, 2016 from http://www.chesapeakebay.net/about.
4) Chesapeake Bay Program – Dissolved Oxygen. 2016b. Accessed December 15, 2016 from http://www.chesapeakebay.net/discover/bayecosystem/dissolvedoxygen.
5) Forgang, B Hamann, and C. Cerco. Visualization of Water Quality Data for the Chesapeake Bay. Proceedings of the 7th IEEE Visualization Conference (VIS96). 1996.
6) S. Funderburk, J. Mihursky, S. Jordan, and D. Riley (eds). Habitat Requirements of Chesapeake Bay Living Resources. Chesapeake Bay Program Habitat Objectives Workgroup, Living Resources Subcommittee. June 1991.
7) J. Kim and J. Park. Visualizing Marine Environmental Changes to the Saemangeum Coast. IEEE Computer Society. November/December 2012.
8) T. Munzner. Visualization Analysis and Design. CRC Press, 2014.
9) J. Shisko, A. Steele, and M. Mengel. IGODS: An important New Tool for Managing and Visualizing Spatial Data.
10) B. Shneiderman. The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. IEEE 1996.
11) R. Stein, A. Shih, M. Baker, C. Cerco, and M Noel. Scientific Visualization of Water Quality in the Chesapeake Bay. Proceedings of the IEEE Conference on Visualization 2000.