Understanding River Plume Salinity through Cross-Filtering

Chris Chung
University of
Washington
Seattle, WA 98195
cc733@uw.edu

Hyun Kim
University of
Washington
Seattle, WA 98195
hyunskim@uw.edu

Lyle Klyne
University of
Washington
Seattle, WA 98195
lylek@uw.edu

Ankit Potdar
University of
Washington
Seattle, WA 98195
apotdar@uw.edu

ABSTRACT

Densely populated coastal areas are often situated at the river mouths, and so gaining an understanding of how river water mixes with ocean water can be key to understanding coastal pollutant, nutrient, and sediment transport in these areas. The tool we are developing aims to aid researchers in understanding how the Fraser River plume (near Vancouver, BC) mixes under different physical forcing conditions. These conditions include measuring against wind speed, wind direction, tidal phase, and tidal magnitude. Our subject matter expert (SME), Sam Kastner, collected ocean salinity data through deployed drifting buoys (Surface Wave Instrument Floats with Tracking, or SWIFTs) during a 10-day field campaign for several hours each day. Through these SWIFT measurements, he observed the impact of these forcing conditions in a variety of data, including buoy position, water salinity and wave height. During the deployments, we observed a large amount of variation in these forcing conditions, which makes understanding the day-to-day variations in plume behavior difficult. Our goal is to develop a visualization of this highly dimensional data, allowing the researcher to further explore, understand, and effectively communicate his research to a variety of audiences.

Author Keywords

Salinity, cross-filtering, plume, data visualization, temperature

ACM Classification Keywords

H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION

An interesting area of exploration for oceanographers, is the formation of river plumes, which occur when ocean water mixes with river water. The sediments transported from the river into the ocean can have large ecological impacts, especially in highly industrialized regions, which may have high levels of pollutants.One example of where this situation may occur, is the Frasier River, located in British Columbia, Canada. It is the longest river in British Columbia, discharging 20 million tons of sediment into the ocean [1]. It provides an excellent baseline to understand how river plumes form, providing data regarding river plume physics. This includes plume circulation dynamics, plume frontal dynamics and mixing, and watershed-river interactions. This information is valuable to researchers, as modelling development of river plumes will help provide decision support for issues related to carbon management, water quality, and ecosystem sustainability.

The transformation and transport of carbon and nutrients in river outflow regions may potentially have a large effect on the surrounding ecosystem. Low alkalinity from freshwater inputs may promote ocean acidification which can impact wildlife such as shellfish [2]. As a result, it is important to understand this phenomena, to determine the extent of river influence. We intend to study how far exactly does a river's impact extend from the continental coastline and beyond and how can salinity be used to measure or predict this.

We worked directly with our SME and sought to create a tool that improves upon his existing workflow, which was time consuming and inflexible. The existing workflow included data cleaning in Matlab, and visualizations generated through hard-coded layers in Google Earth. Our goal was to allow for rapid data exploration of the dataset using a custom tool.

RELATED WORK

Although the area of ocean salinity and river plumes have been studied for many years, the development of visualizations to support these findings has not kept up. One visualization tool was influenced by a paper by Samsel et al [3] caled "Interactive Colormapping."This tool was published in 2016, and presents a workflow, enabled by an interactive colormap tool that allows a scientist to interactively apply complex colormaps to scalar data. Although this tool solely focuses on ocean salinity, and not river plumes, it provided many heuristics that current editing tools fail to provide to allow scientists to easily explore the subtleties of salinity.

One example was the importance of color, which was described as a critical component of a good visualization, especially if data is complex or has nuanced ranges. We compared the use of ColorBrewer [3], the NASA color tool [4], and Gregor Aisch's chroma.js [5]. The use of these tools was very helpful in selecting the optimal color scales, as well as the impact of color spaces.

METHODS

To approach this problem, we followed the user centered design process. This consists of four phases: 1) Plan, 2) Design, 3) Prototype, and 4) Evaluate. The first step we took was the planning phase, where we began conducting interviews with our SME, to determine what his expectations were. We had him sketch his ideal layout, to understand what requirements he was expecting. It was from this first meeting that we decided on implementing cross-filtering in our tool. A pain point raised by the SME with his current workflow, was the inability to explore and compare multiple dimensions simultaneously.

From the insights gained from this meeting, we were able to get an initial direction for the overall layout of our tool, and the density of information we would like to provide our user. We were able to identify high priority dimensions valuable to our SME. For example, showing a scatterplot of surface salinity against surface temperature is important for validating the dataset and there should be a linear relationship between these dimensions. Our SME would always check this chart before proceeding with further exploration of data. Next, we began designing our interaction model, by storyboarding the interaction through sketching. This consisted of drawing general layout design on paper, as well as possible points of interaction. (see Figure 1)
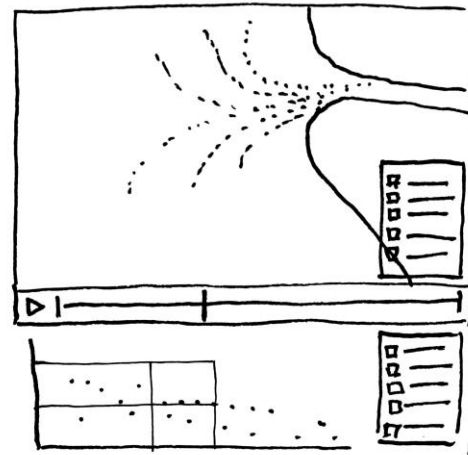

Figure 1: Paper Sketch of Layout

We then began prototyping with Tableau, using the data we received from the SME, to determine the features and visualizations that best showed correlations or trends (see Figure 2). The SME cleaned the data through Matlab before sending it to us, eliminating dirty data recorded from the buoys such as NaN or outliers from when the buoy malfunctioned. From Tableau we were able to identify several correlations between certain dimensions, such as a linear relationship between water temperature and salinity. We also conducted a short user test of our prototype in Tableau with the SME, to determine if it aligned with his user goals. From this we were able to determine that maps and scatterplots would be the main focus of our tool.
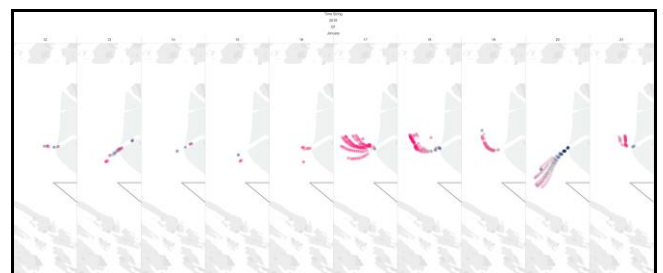

Figure 2: Visualizing Data in Tableau

After this, we began the coding phase, building our visualization tool using D3. Most of the implementation work was done collaboratively in a group setting. We worked on choosing the perceptually correct color palate for our tool. Once we made sufficient progress to demo the tool, we met with our SME for a usability test to get feedback on the current visualizations (see Figure 3).

Compared to the previous user test which was done primarily to gauge his expectations, the next iteration of tess with a higher fidelity allowed him to provide more specific feedback. We were able to then incorporate his feedback in the final version of our tool. This version was then further tested with 4 oceanographers. The results of the test are discussed in the discussion section.
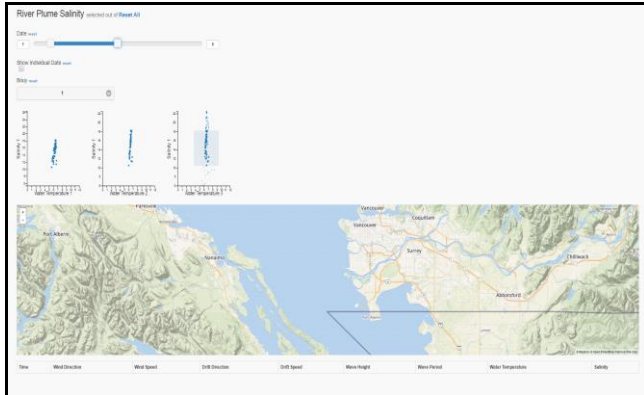

Figure 1: Initial Prototype of Tool

RESULTS

The final version of the tool included the following features:

- Improved workflow through cross-filtering (see Figure 4)
- Show 4 key scatterplots for ocean salinity data
- Explore each day by small multiples (see Figure 5)
- Perceptually correct color for visualizations

On the top level, you can find four scatterplots, which compare surface salinity with different dimensions such as Water Temperature, Salinity at 1.2m, Bouy Drift Speed, and Wave Height.

You can select a subset of the data by clicking and dragging the cursor to form a box, which will then cross-filter the data across all visualizations. Wind direction was also visually encoded for the fourth scatterplot, with respect to true north (0=North, 180=South) We have also provided controls on the middle section, which allow users to select a specific range of dates, data from specific buoys, and the option to view the map data which is located on the bottom level, over small multiples, with each individual map representing data across a single day.

The map data on the bottom has two possible views, the large map which shows the buoy's position and salinity level measured across all days, and a small multiples view for individual dates.

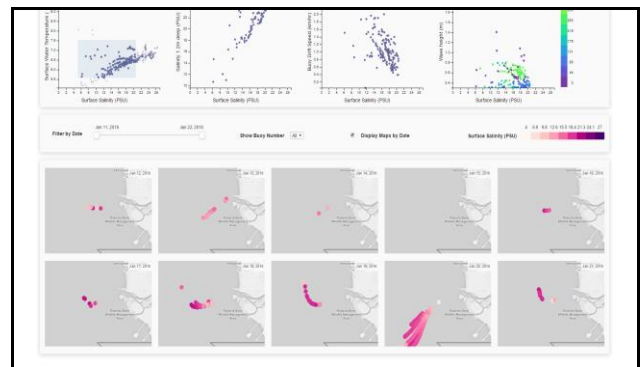**Figure 2: Screenshot of Cross-filter Functionality**

Figure 3: Screenshot of Small Multiples View

To evaluate the effectiveness of our tool, we conducted a usability test with our target audience. This included a pilot test with our SME, who tested our initial prototype. Part of the user test was the evaluation of the different color scales we were considering. For encoding wind direction, we first tested with shapes and color, by double encoding shape and color. Shapes provided a poor experience for finding clusters in the scatterplot. On the other hand, using only colors supported the goal of finding colors well. We wanted to use a discrete rainbow color scheme which was perceptually correct [7]. We choose a library, Color Helix II, which allowed us to quickly explore color palettes in LAB and HSL instead of RGB color representations. We had more challenges deciding on a color scheme for our map, due to the light gray background, which caused lighter and single color scales to be less visible. We received specific feedback from Sam regarding the map color scheme, which was then used for the final iteration of the tool, with the red color scale being chosen for the (see Figure 6).

We used Mapbox mapping library for cartographic visualizations. Since the geographic area of our visualization was relatively small, we used Mercator projections [8]. We were also able to learn more about the general performance of our tool, as well as its current limitations. The map rendering was considerably slow when using Mapbox GL, so after conducting the user test, we switched to Mapbox js. This had several implications. We could not simultaneously interact with small multiple maps as Mapbox js used images instead of WebGL for rendering. However this was a compromise for having usable performance. In the final iteration, we conducted 4 final user tests with a group of oceanographers for additional feedback. One of the main concerns found was the unselect feature for the cross-filtering, which was not always intuitive. 3 out of 4 of the oceanographers found that they sometimes forgot to click within the scatterplot to clear the cross-filter, instead, trying to select other scatterplots, or selecting the empty space outside the scatterplot. Another issue was the inability to select a custom shape, since clusters formed in the data will not always be rectangular in nature. We also found several potential improvements from the feedback during the final user tests, which can be found in the future work section.

ColorTool	Colorscale	User Feedback
ColorBrewer	Red-blue	Best contrast, but displays green hue in color tool, but data does not have enough range to include green, only shows red and blue
Cubehelix II	Rainbow	Can be overwhelming if implemented poorly, but contrasts well when mapped and scaled to our data
Chroma.js	Yellow-green	Used HSL values, but is very hard to see on gray map background, and too similar to the Rainbow scale
ColorBrewer	Red	Single color is less glaring, and has good contrast with the map background

Figure 6: Table of Color Scale Evaluation

DISCUSSION

The audience of the tool, mainly researchers in oceanography and related fields, found the tool to be very helpful and versatile in identifying and evaluating trends in the data. We found through our user tests that our visualization would make a good educational tool for people new to the field, or just learning the data analysis process for collected oceanographic data. One finding was identifying outliers in the data. By utilizing cross-filtering, you can isolate certain subsets of the data on the scatterplots, and then look at the map to determine when the data for that cluster was collected. An example of this was an outlier identified from the first scatterplot from Jan. 21, 2016 to Jan. 22, 2016 (see figure 7). You can see that the cluster is very disjoint from the rest of the collected data, indicating a potential malfunction in the buoy.
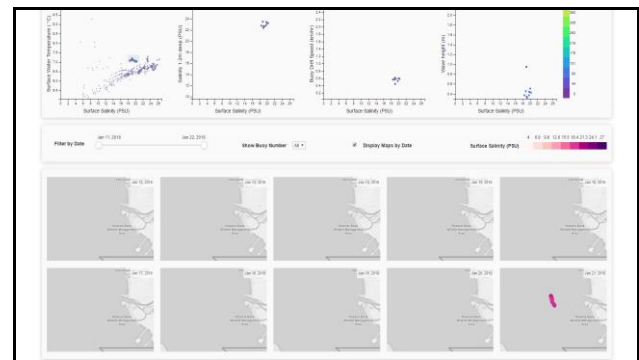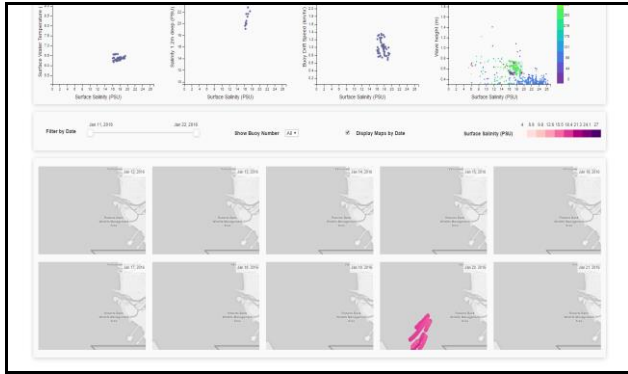

Figure 7: Screenshot of Outlier Identified

Another insight gained from our tool was identifying unexpected patterns from the data that differ from the expected behavior. One example of this was the formation of a green cluster in our fourth scatterplot, which only happens on one day, Jan. 20, 2016 (see Figure 8). What's interesting is the conditions on this day, which differ from the norm. On this day, the path of the buoy was going south, the wind direction is constant, and the water is very fresh. Usually the path of the buoy goes north, the wind direction shifts, and there is a lot of mixing, as indicated by high salinity levels. This shows a potential trend between physical factors such as wind direction, corresponding to changes in salinity levels.

Figure 8: Screenshot of Cluster Identified

FUTURE WORK

The next step for future iterations of the tool, would be the creation of a data wrangler script to transform the data. This would make the process more efficient, as data would no longer have to be manipulated in Matlab, before being imported. Another potential addition, would be the ability to toggle different dimensions to be mapped on the scatterplots, making the tool more flexible.

There were also potential improvements identified during our final user test that could be included in our future work. This includes the ability to toggle other parameters to be mapped on our scatterplots. The test participants from oceanography felt that other dimensions like nutrients or sediment levels could be useful for comparison. They suggested the ability to have a dropdown next to the axis to change parameters. Another improvement was having the bounding box when a cross-filter was selected to appear across all scatterplots. This would help for context, since isolating subsets of the data is nice, but having the rest of the data out of focus, can be less efficient for data comparison. In the current tool, you have to go back and forth between selecting and unselecting to see trends on other scatterplots. A final improvement that could be made for a future iteration would be adding the ability to cross-filter in the reverse direction. In other words, allow users to select a subsection of the map, and have the cross-filter apply on the scatterplots.

ACKNOWLEDGMENTS

REFERENCES

1. Richard J. Cannings and Sydney G. Cannings. 1996. British Columbia: a natural history, Vancouver: Greystone Books.

2. Joseph Salisbury, Mark Green, Chris Hunt, and Janet Campbell. 2008. Coastal Acidification by Rivers:A Threat to Shellfish? Eos Trans. AGU Eos, Transactions American Geophysical Union 89, 50 (September 2008), 513–513. DOI:http://dx.doi.org/10.1029/2008eo500001

3. Francesca Samsel et al. 2016. Interactive Colormapping. Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '16 (2016). DOI:http://dx.doi.org/10.1145/2851581.2851587

4. M. Harrower and C.A. Brewer. 2011. ColorBrewer.org: An Online Tool for Selecting Colour Schemes for Maps. John Wiley & Sons, Ltd, 261–268. DOI: http://dx.doi.org/10.1002/9780470979587.ch34

5. Anon. Color Tool. Retrieved June 8, 2016 from http://colorusage.arc.nasa.gov/colortool.php

6. Aisch Gregor. Chroma.js. Retrieved June 8, 2016 from http://old.driven-by-data.net/about/chromajs/#/1

7. Niccoli Matteo. 2013. Perceptual rainbow palette – the method. (2013). Retrieved June 8, 2016 from https://mycarta.wordpress.com/2013/02/21/perceptual-rainbow-palette-the-method/

8. B. Jenny. 2012. Adaptive Composite Map Projections. IEEE Trans. Visual. Comput. Graphics IEEE Transactions on Visualization and Computer Graphics 18, 12 (2012), 2575–2582. DOI:http://dx.doi.org/10.1109/tvcg.2012.192