Visualization for River Data

First Author
University of Illinois-Chicago
Noah Angeles
Chicago, IL 60607, USA
nangel6@uic.edu

Second Author
University of Illinois-Chicago
Jamie Auza
Chicago, IL 60607, USA
auza2@uic.edu

Third Author
University of Illinois-Chicago
Jacky Duong
Chicago, IL 60607, USA
jduong2@uic.edu

Abstract
The goal of our visualization is to help river researchers in their struggle to preserve and reconstruct rivers around the world. We are able to identify the different pairs of pools and riffles in many different rivers, as well as the asymmetry between these pairs. This information will help in the reconstruction of fish spawning areas which is the ultimate goal of analyzing this domain.

Introduction
Our visualization focuses on two main problems that Erin Bray and her team wanted to focus on. The first main problem we wanted to solve is identifying pools and riffles. In our project, we explore different ways to identify these peaks in the various rivers that we have been provided. We used a trendline to clear out some undesirable noise in the data, and we also identified all other max and minimums inside the data to visualize points that may be peaks that were not included by our trendline algorithm. The second problem we wanted to focus on was asymmetry. To identify asymmetry in the pool and riffle pair, we used bar graphs to visualize and compare the lengths of each rising and falling slope. The goal is to be able to compare the lengths of the slope and use that as a way to measure asymmetry.

Related Works

The entirety of our project has been inspired by resources that Erin Bray has provided for us. The power point that she gave us already had most of the visualization that she and her team of researchers wanted us to focus on. Even though the visualization has pretty much been given to us, our job was to make it more clear in terms of what data we need to highlight as well as show different aspects of the data that is interesting to the researchers.

From our meeting with Erin Bray, the main reason we are analyzing this particular data domain is for reconstruction and preservation of rivers. Specifically, the researchers will use this data to recreate rivers to better suit fish spawning. The area where fish tend to spawn is where the pool ends and the riffles begin. We are locating exactly where these points are so that the researchers can easily identify where these fish are likely to spawn. Asymmetry of the riffles and pools is for the reconstruction of these points due to natural weathering.

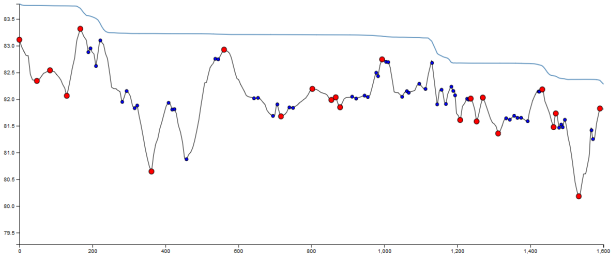
Analysis Tasks

The first of our tasks is to be able to identify riffles and pools within the river. We also want to separate these riffles and pools from the tiny peaks that are not actually a pool or riffle. For this part of our visualization, we were provided with the height of the ground under the water levels. We mainly used the height of the ground to identify the different peaks and then check if these peaks were actually what we wanted. The idea for using a trendline was to determine the highest or lowest point in a certain segment before it rises or falls below the trendline. This is how we cancelled out most of the noise that are not considered riffles and pools. This method makes the data a lot cleaner, but it is not always work depending on what the rivers particular trendline looks like. For that reason, we also identify other max and minimums inside the river to see if any of these points can be considered a pool or a riffles.

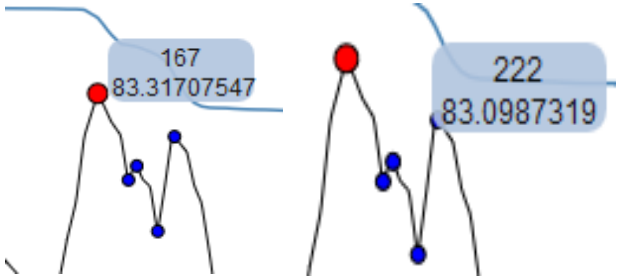
The second tasks we focused on was asymmetry inside the pool and riffle pair. From the article, we concluded that the best way to identify asymmetry was to analyzing the length of the rising and falling slopes. The reasoning we had was that each pool riffle pair consists of a rising and falling slope and we use that knowledge to compare if the two slopes are the same length. If they have similar length, then that would mean they have some symmetry versus other pairs that may have very different lengths. Fortunately, in the data we were given the slope of each point. This is how we were able to determine how far these rise and falls go to ultimately calculate the length.

Visualization Details

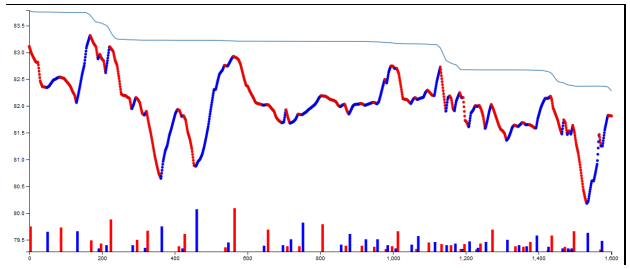
We have three different options for our visualization. The first is being able to pick which river you want to analyze, and there are 14 different rivers in all. The second is a button that shows the maximums and the minimums as well as the pools and riffles identified by our algorithm. The third button replaces whatever graph is being shown into a graph that displays the rising slopes, falling slopes, and lengths.



The graph above shows the max and mins of the river. As you can see, there are two different circles (red and blue) which highlights the peaks of the river. The red circles are the pools and riffles that we have identified using the trendline. These circles are a brighter color and slightly bigger than the other circles because we wanted them to be much more noticeable. These points are what matters most and they should be highlighted. The smaller, less noticeable, blue circles shows the different max and mins that were not included by the algorithm. As you can see, some of these points (red and blue) are very close to each other, so our algorithm is very helpful in cancelling out some of these useless points, but it is not perfect. We just cannot generalize it enough so that it would not include points very close to each other because the data we have for each river varies a lot. Some rivers have way less data than others. The line above the line graph is the water elevation.

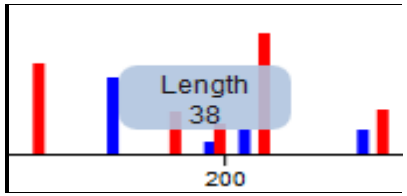


The above figures show that the user is able to hover over the different points within the graph. This is helpful because the researcher involved needs to know these exact points.



The graph above shows our second visualization which deals with the asymmetry task. Comparing to the first graph, this graph is color coded depending on the slope of each point which was given to us. This graph is not a solid line because we wanted to show the slope of each point rather than the line. The blue line is the positive slope and the red line is the negative slope. We used this coloring because of the connotation each color gives off which makes it easier to tell from first glance. Blue is a more positive color and red is more negative. At the bottom along the x axis, you can see bars corresponding to the length of each alternating rising

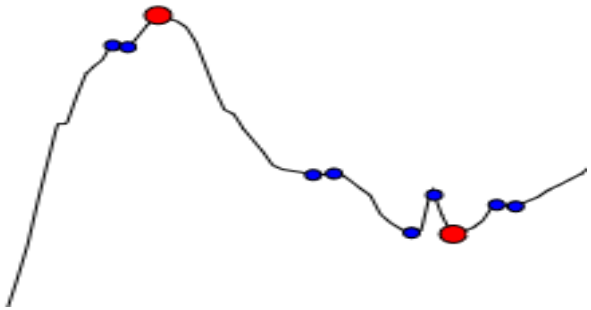
and falling slopes. Again, these are color coded red and blue for falling and rising slopes respectively. The straight bars are supposed to make the asymmetry between each pair very clear opposed to the line graph itself. Each bar is also located where each slope begins, so the user can easily tell its corresponding points in the line graph.



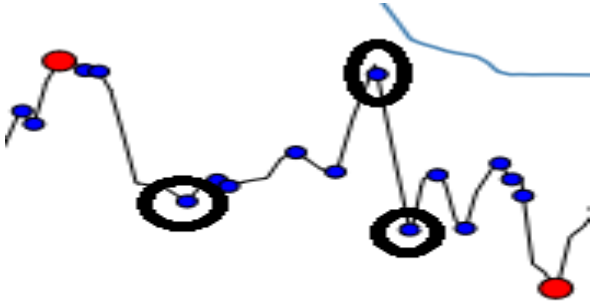
The user is able to hover over each bar to find out the exact length of each slope. This is to ensure of how symmetrical each line is actually (exact versus approximate).

Case Study

To better understand how our visualization solves our tasks we will go through a few theoretical examples. First of all, how do we solve the riffles and pools case with our first visualization?

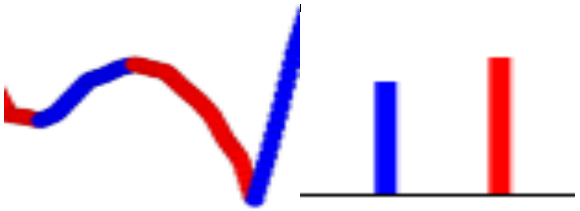


The above picture shows the peak and riffle pair (red circles). As you can see, there are many other points inside this segment. These represent the other peaks without the trendline. Since riffles and pools are basically just peaks, if we didn't have a trendline all these tiny points would have been determined as a riffle or pool which we do not want.

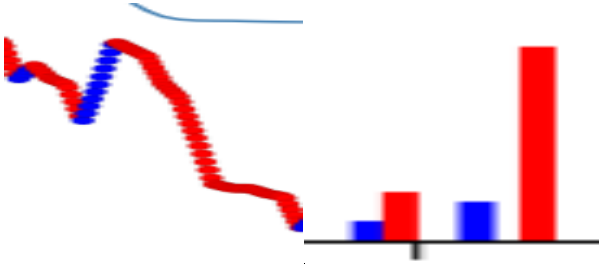


In contrast to the previous picture, the one above has points not highlighted in red, but we would like them to be. The reason these points are not included is because there is already a point much higher that all the others that are above the trendline (first red circle). That is why we decide to include all other max and mins so the user can determine whether points are pools or riffles that were not considered by the algorithm themselves.

The second task is to be able to determine asymmetry between pairs.



The above picture shows a pair that have very little asymmetry. As you can see in the line graph, the pair looks very symmetrical. The bar graph corresponding to the pair is shown on the right which shows that the length of the slopes are nearly identical which confirms their symmetry.



Here is an example that shows very clear asymmetry in two different pairs and their corresponding bar graph. You can tell right away that the falling slopes are much longer than there rising slopes which means they are asymmetrical.

Expert Feedback

The domain expert, Erin Bray, told us that our initial design was very nice and simple which makes it very easy to understand what we were trying to do. She did also say that we could put a title and a legend to our graph to make our intentions clearer. We were already planning to add them, but we just didn't have enough time before sending the prototype to Erin. She also gave us some tips on how to make our asymmetrical analysis of the river segments clearer. She was not too sure of what to make of the slopes at first sight if we didn't give an explanation. She wanted us to make that clear. Basically, the goal was to make everything much clearer so the researcher can easily tell what the intentions were.

She enjoyed how the graphs dynamically change when we change rivers, and the animations was also nice. She did say that identification of pools and riffles lengths are missing which is important. Erin also mentioned that we should include some sort of index for the final symmetry (or asymmetry) of a given pair. She also gave us a way to calculate the length of riffles and pools which we were struggling at, but she believed that we might not have time to implement and debug the new algorithm in time or take the visualization as a whole a step further.