Lab Report

Title: Lab 3 Part 1 Deliverable
Notice: Dr. Bryan Runck
Author: Taylor Andersen-Beaver
Date: 12/1/22

Project Repository: https://github.com/and03449/GIS5571.git
Google Drive Link: N/A
Time Spent: 22 hrs

Abstract
Using a robust group of datasets, analyze the best route for Dory to take when talking from her farm to her fly fishing spot in Whitewater State Park. To properly analyze the best route, this group of datasets must include road networks, water networks, elevation, and land cover classification. Once obtained, using map algebra and a raster calculator, all of the rasters are combined and different weights can be applied to evaluate a cost surface analysis.

Problem Statement
Dory is looking for the best way to get from her farm to her favorite fly-fishing spot in Whitewater State Park. Considering the terrain, roads, traffic, and bodies of water, what are the optimal routes Dory can take to get to her fly-fishing spot (assuming Dory wants multiple options, not just the 'best'). Dory prefers to not walk through any farm fields because they can be muddy in the spring. She also doesn't like crossing water bodies if there isn't a bridge. Other than that, she just wants to take the path that is the most gradual in terms of slope.

Table 1. Requirements Needed for Cost Surface Analysis

#	Requirement	Defined As	(Spatial) Data	Attribute Data	Dataset	Preparation
1	Locations	Dory's farm and her fly-fishing spot (North Picnic area of Whitewater State Park)	Coordinates	N/A	Imported from GitHub	Gather coordinates from deliverable (Dory's farm) and Google (for the North Picnic point)
2	Data Acquisition	Consider the different elements that would impact a route from Dory's farm to the park - roads, water, terrain, elevation, etc	Elevation, road network, bodies of water	Road names, stream order	Multiple from MN Geospatial Commons	Locating data for each of the elements Dory will be facing on her route

3	Buffer	Buffer around the locations to create an area of interest	Distance between and around the points	County names	County boundaries	Full state county boundaries to identify the area of interest
4	Clip	Clip each of the datasets being used to the area of interest established with the buffer	County boundaries	County names	County boundaries	Identify the appropriate counties
5	Erase Streams	Erase portions of the streams that intersect with roads (assuming these intersections are bridges)	Stream polylines, Road polylines	Stream order, Road names	MN Streams, MN Roads	Identify intersections of streams and roads
6	Rasterize	Converting all datasets to rasters to be able to calculate the cost surface	Streams, roads, land cover, elevation	N/A	MN Streams, MN Roads, MN Land Cover, MN DEM	Prepare all datasets to be rasterized
7	Reclassify	Reclassify all rasters to have an equal amount of classes to make the cost surface calculation more accurate	Stream raster, roads raster, land cover raster, elevation raster	N/A	MN Streams, MN Roads, MN Land Cover, MN DEM	All datasets must be rasters
8	Cost Surface	Combine all rasters and calculate the cost surface using different weights	Stream raster, roads raster, land cover raster, elevation raster	N/A	MN Streams, MN Roads, MN Land Cover, MN DEM	All datasets must have same or similar number of classes

Input Data

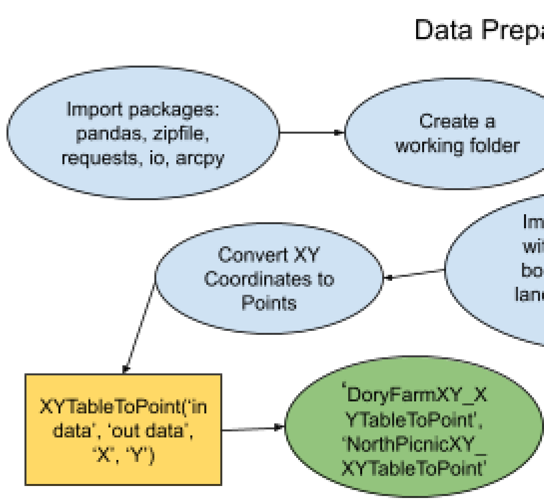
The data I have chosen to use for this lab is MN Road Networks, MN County Boundaries, MN Digital Elevation Model, Stream Routes with Stream Order, and Land Cover Classification. I have chosen these datasets because of the problem statement listed above: what is the optimal route for Dory to take to her favorite fly fishing spot? To determine the optimal route, all conditions need to be considered including roads, rivers, and elevation.

Table 2. All Datasets Needed for Cost Surface Analysis

#	Title	Purpose in Analysis	Link to Source
1	Minnesota Roads	Raw input dataset for routing analysis from MNDOT	https://gisdata.mn.gov/dataset/trans-roads-mndot-tis
2	MN County Boundaries	Having a foundation to clip the data to the study area	https://gisdata.mn.gov/dataset/bdry-counties-in-minnesota
3	MN Digital Elevation Model	Using to analyze elevation on the path from point A to point B	https://gisdata.mn.gov/dataset/elev-30m-digital-elevation-model
4	Stream Routes with Stream Order	Locate streams and rivers including stream order to determine bigger rivers from small streams	https://gisdata.mn.gov/dataset/water-strahler-stream-order
4	NLCD 2019 Land Cover	Land cover classification to classify different land types to help analyze the best path	https://gisdata.mn.gov/dataset/biota-landcover-nlcd-mn-2019

Methods

To begin the analysis, first I need to locate the data needed and prepare it for analysis. Starting with importing packages into an ArcGIS Pro Notebook, create a working folder to store all of the data and import the data into the notebook. Lastly, creating a CSV file with the start point (Dory's Farm) and another CSV with the end point (North Picnic), importing the CSVs into the ArcGIS Pro Notebook and converting the XY Coordinates to Points.



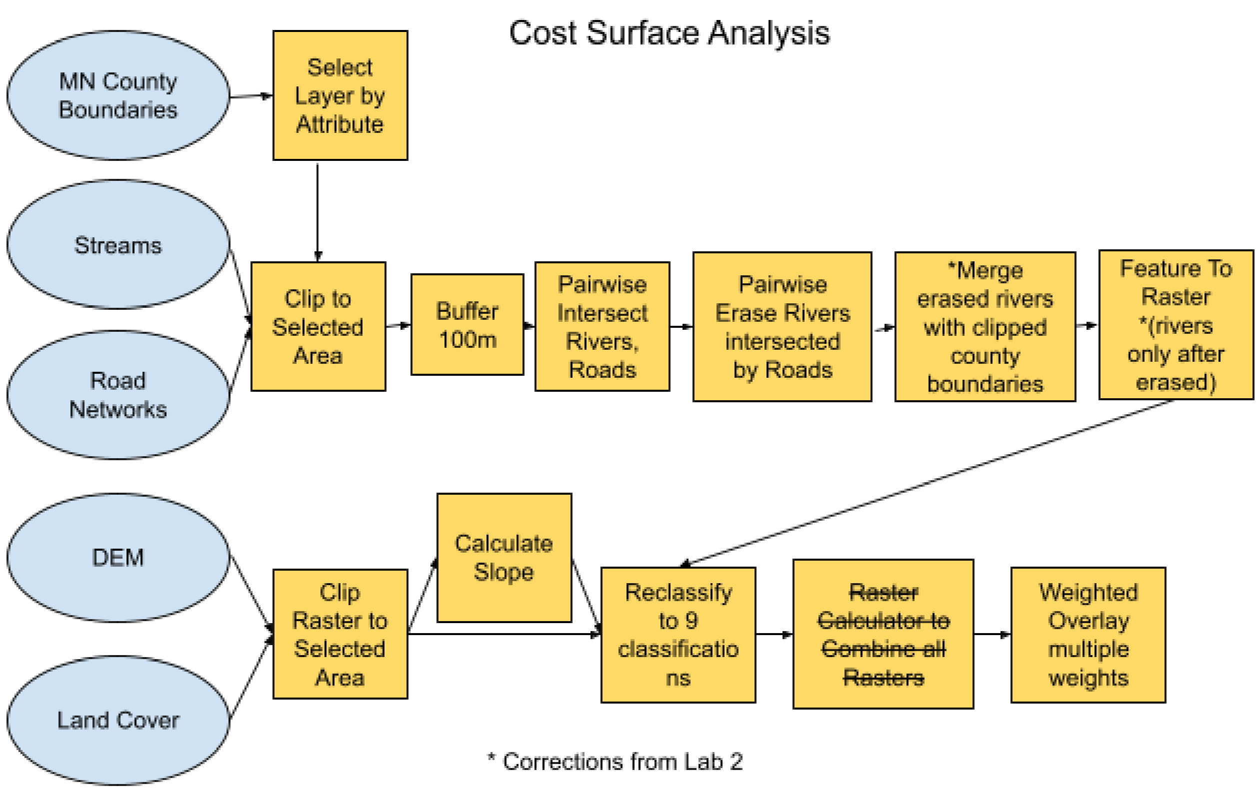
With the data imported and prepared for analysis, the next step is to clip the data to the same extent. The points are in two different counties and 'as-the-crow-flies' crossing over a third county, so I have decided to clip the data to all 3 of those counties: Olmstead, Wabasha, Winona.

By selecting the 3 counties and clipping the county boundaries, I can then use the clip as a template to clip all of the other datasets whether with a pairwise clip or a raster clip.

Once clipped, the datasets follow two different paths: vector and raster. The vector data, roads and streams, are first buffered to 100 meters to make them 'thicker' to eventually be converted into a raster. Next with the vectors, is finding intersecting points and erasing any places where rivers intersect with roads which will essentially erase any rivers that are overlapping the bridges that cross them making the bridges available to Dory's path. Once intersected and erased, the datasets can be converted into rasters and then reclassified.

For the raster data, once clipped, the land cover data can go directly to reclassification. The digital elevation model needs to be sloped before reclassified. The reclassification, for both vectors to rasters or just rasters in general need to be as similar as possible in order to best create a cost surface raster (*Creating a Cost Surface Raster—ArcGIS Pro | Documentation*, n.d.). For this analysis I have chosen to use 9 classifications since that seemed to be the average amount of classifications per raster.

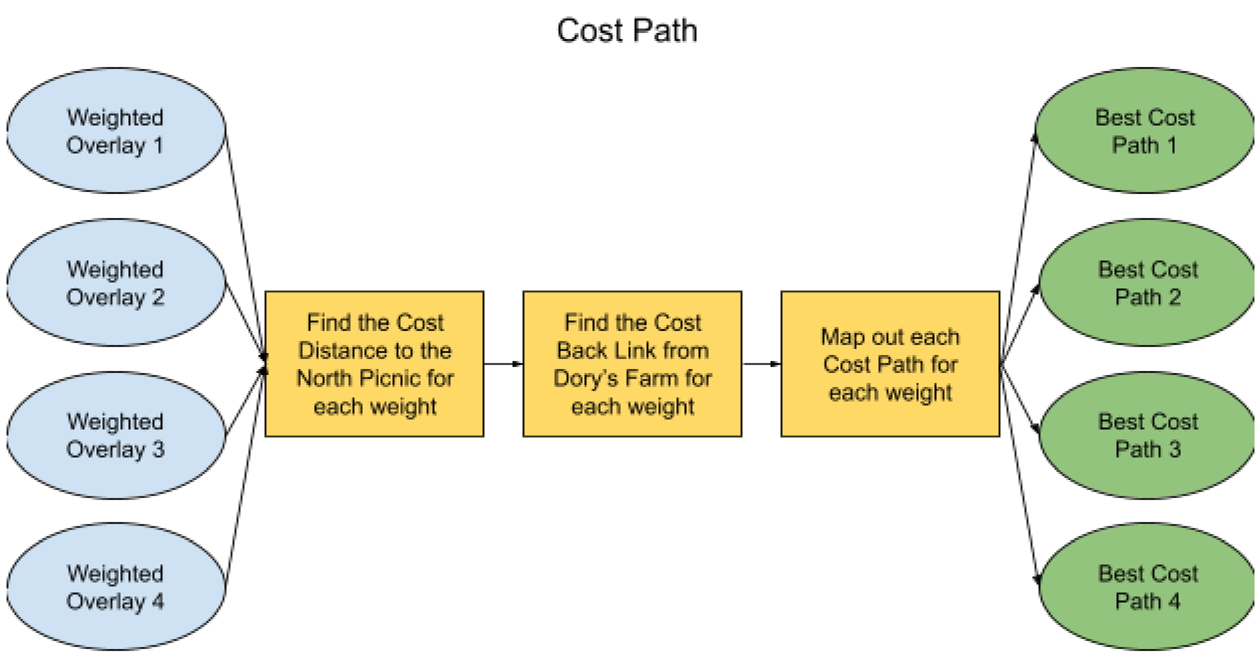
Once reclassified, each of the rasters can be combined into one using the raster calculator which can then be used to find the cost surface analysis (*Raster Calculator (Spatial Analyst)—ArcGIS Pro | Documentation*, n.d.). The last step in the process is to set weights to each of the rasters using the weighted overlay by adding each raster and giving them each a weight calculating to 100% (*Weighted Overlay (Spatial Analyst)—ArcGIS Pro | Documentation*, n.d.).



The only differences between the functions I performed in Lab 2 to Lab 3 are in the streams/roads process and the final weighted overlay step. For the streams/roads I realized that once the rivers are erased anywhere that the roads intersect then there is no need to keep the road layer. Unless Dory was only interested in taking the roads or only interested in **not** taking the roads then there is no need for them which is why I only kept the streams layer when moving forward to the overlays.

Another step I realized was missing in the streams/roads process was that once I converted the feature to a raster there was a lot of missing data, essentially anywhere there was not a stream which caused problems in the weighted overlay process. To combat this I decided to merge the streams/roads layer (pre-raster) with the initial clipped county boundary to fill any of the missing data from streams/roads with the boundary data (*Merge Rasters Function—ArcGIS Pro | Documentation*, n.d.). Once rasterized and reclassified, I can classify those county boundary cells as 'NoData' or '0'.

The last change I had to make from Lab 2 to get a better result was to skip the step of combining all of the rasters using a raster calculator before calculating the weighted overlay. Omitting this step did not necessarily 'fix' anything but was redundant and unnecessary. While reviewing Lab 2 I also paid closer attention to the reclassification step for the land cover, streams/roads, and the elevation raster to be sure they all aligned better than before.



To find the potential cost paths using the weighted overlays from the previous steps, it was actually a relatively simple process. The first step is to find the Cost Distance - this is the distance to Point B (the North Picnic area) using the weighted overlay raster (*Cost Distance (Spatial Analyst)—ArcGIS Pro | Documentation*, n.d.). The next step is to calculate the Cost Back Link - this is the link **back** to Point A (Dory's farm) using the same weighted overlay raster (*Cost Back Link (Spatial Analyst)—ArcGIS Pro | Documentation*, n.d.). Lastly all is needed are the outputs from the Cost Distance and the Cost Back Link to find the Cost Path (*Cost Path (Spatial Analyst)—ArcGIS Pro | Documentation*, n.d.). These three steps can easily be completed for each of the weighted overlays to produce the best Cost Path.

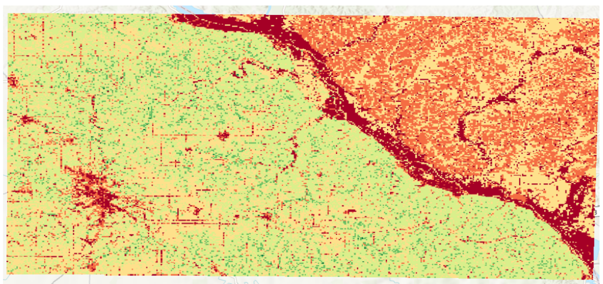
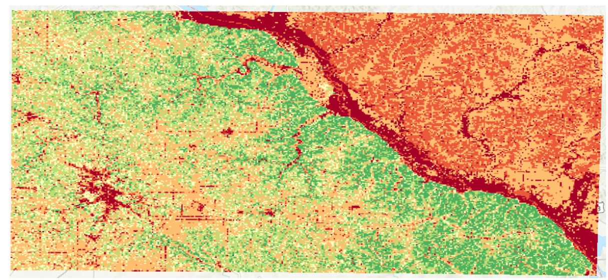
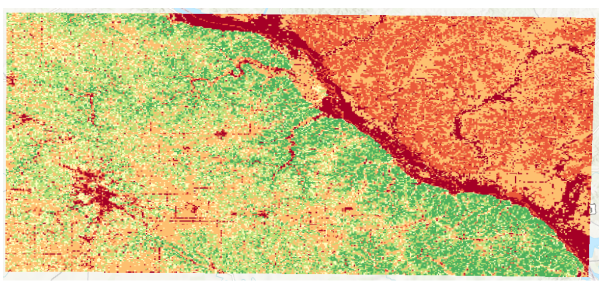
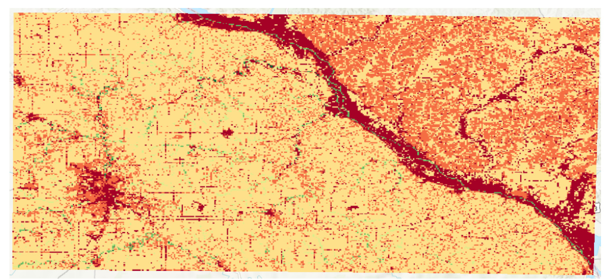
Results

To visualize the difference between the overlays I chose a variety of weights and compared them all below. To start, I started with equal weights of 33% (and one is 34%) to see what an equal, controlled cost surface looked like to be able to see the differences in the other cost surfaces. Each of the rasters are pretty neutral - equal amount of weight on streams, elevation, and land cover.

There is more of a difference once I added different percentages of weights onto each element of the rasters. For example, in Table 3, the weighted 2 raster is showing much more of the land cover element than the streams because I put a higher weight on the land cover than the streams. This would mean that Dory would want to stick to the land classification and avoid the streams. The opposite can be seen in weighted 4 where the streams have a higher weight than the land cover so Dory would want to follow the streams.

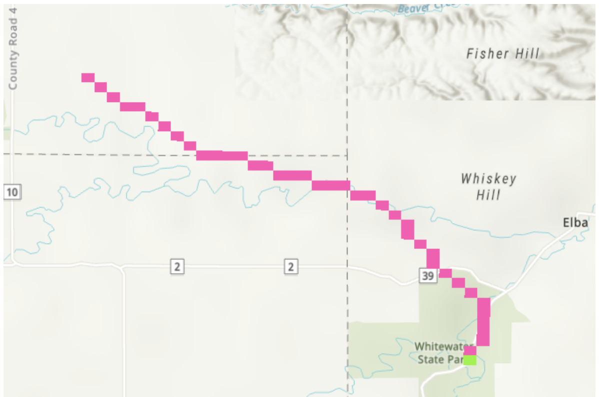
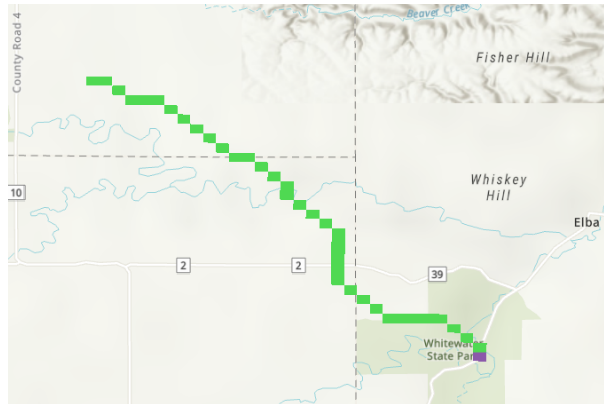
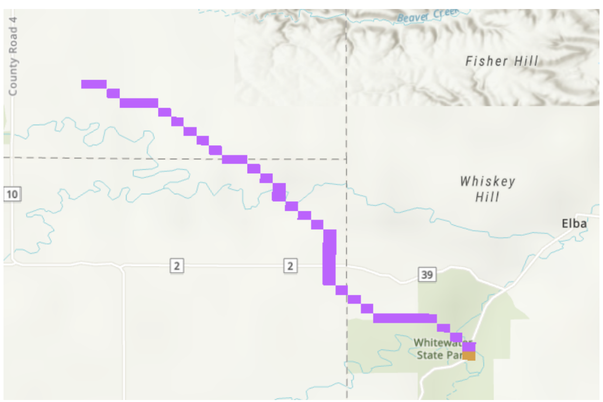
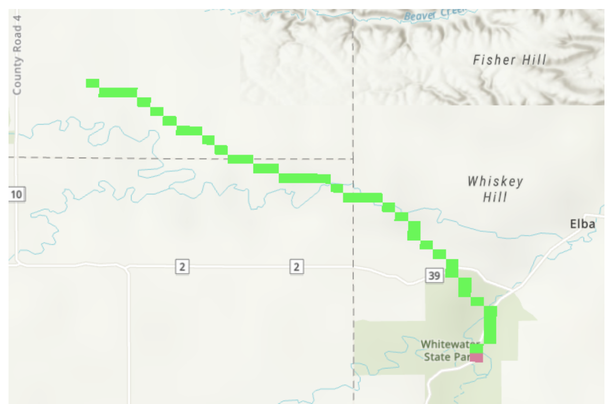
The only weighted overlay that showed little change was weighted 3, which added a little weight to streams, a lot to elevation, and took some weight from land cover. This shows little change from weighted 2 because the land cover and elevation are closely aligned in their output so changing the weights between these two elements does not have much of an effect. The biggest effect is changing the weights of the streams, as seen in weighted 4.

Table 3: Weighted Overlay Rasters with 4 Different Combinations of Weights

Weighted 1 - 33(34)% weight for each raster	Weighted 2 - 10% Streams, 30% Elevation, 60% Land Cover
	
Weighted 3 - 30% Streams, 60% Elevation, 10% Land Cover	Weighted 4 - 60% Streams, 10% Elevation, 30% Land Cover
	

Considering the above weights, I created cost paths for each of the weighted overlays. The weighted 1 overlay is again my controlled overlay where each of the elements has equal weight. The path is pretty direct from Dory's Farm to the North Picnic, not making any special accommodations for streams, elevation, or land cover. Weighted 2 & 3 are the same path - this again is referencing back to what I explained above that the effects of elevation and land cover are closely aligned so changing the weights between the two does not have as much of an effect on the path as the streams do. Weighted 2 & 3 can be seen to veer more South and follow a forested area rather than the streams. Weighted 4 is similar to weighted 1 but does not follow the exact same path. There are slight differences in that weighted 4 follows the streams a little more closely than weighted 1. All 4 of these options are the quickest paths for Dory to take to the North Picnic, depending on which options she prefers.

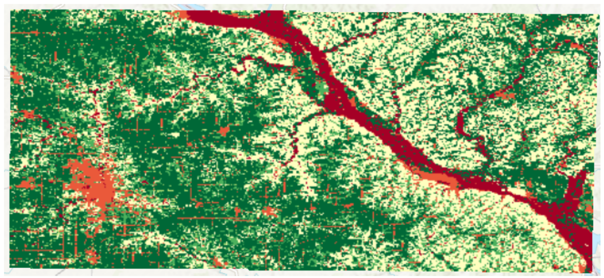
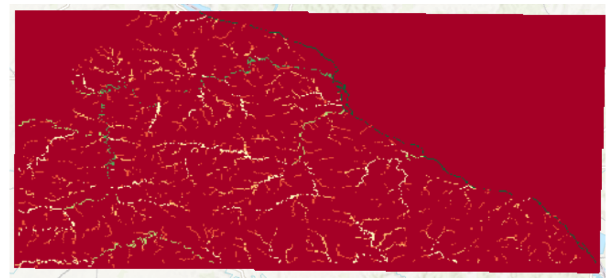
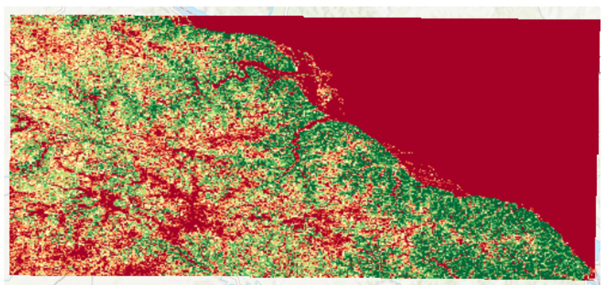
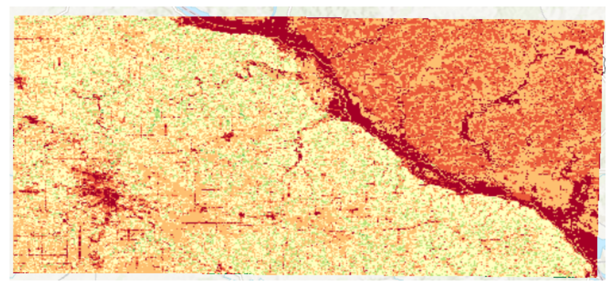
Table 4: Calculated Cost Paths of Weighted Overlays

<table>
<tr><td>Weighted 1 - 33(34)% weight for each raster</td><td>Weighted 2 - 10% Streams, 30% Elevation, 60% Land Cover</td></tr>
<tr><td>
</td><td>
</td></tr>
<tr><td>Weighted 3 - 30% Streams, 60% Elevation, 10% Land Cover</td><td>Weighted 4 - 60% Streams, 10% Elevation, 30% Land Cover</td></tr>
<tr><td>
</td><td>
</td></tr>
</table>

Results Verification

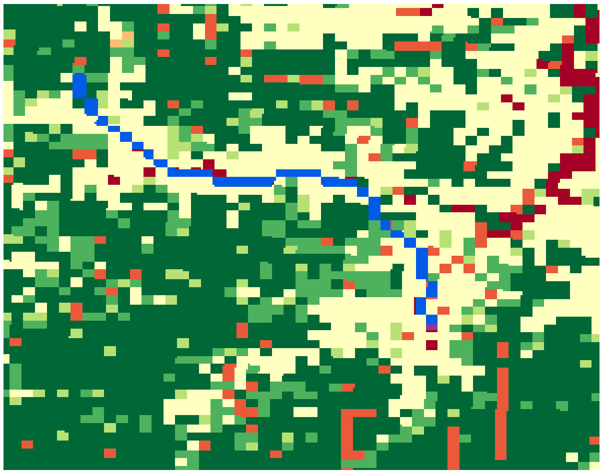
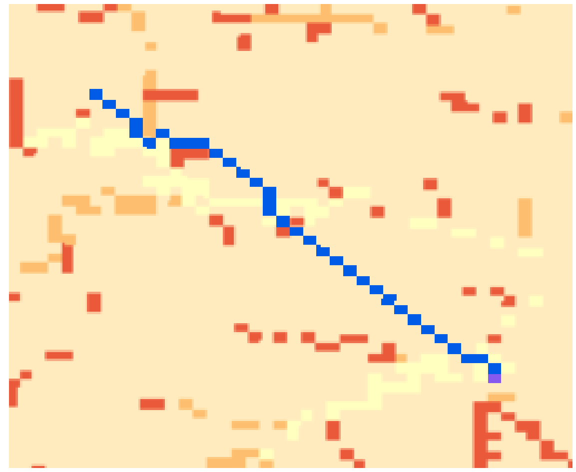
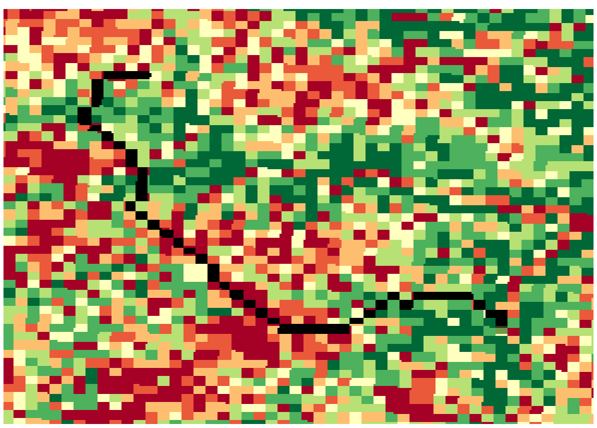
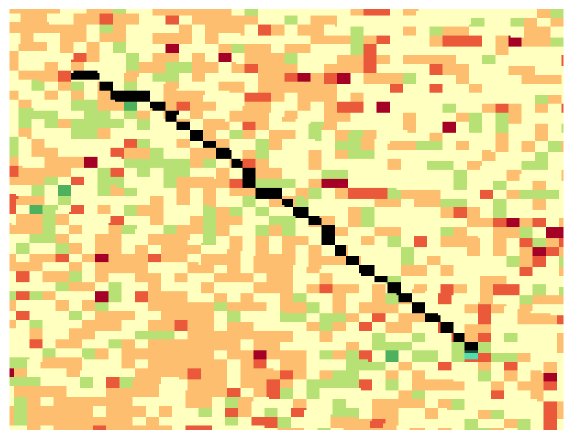
These results can be verified by testing multiple different weights for each raster (streams & roads, elevation, and land cover) and seeing if there is an element that is particularly sensitive to weight. Looking at the ‘Weighted Overlay Rasters’ table above, it is clear that the stream raster is more sensitive to weights than the land cover or elevation because the cost path changes direction when the stream weight changes drastically.

Table 5: Testing Multiple Weights to Verify Results

Land Cover - 100% Weighted	Streams - 100% Weighted
	
Elevation - 100% Weighted	All 3 Weighted Evenly
	

These assumptions can be seen in the cost paths calculated below using the same weights. First is the cost path with the land cover overlay and the path clearly follows the lighter green cells which are the forested area in the land cover data. In the streams example, the path is crossing over streams as minimally as possible and mostly crossing the peach colored cells which are ‘free’ cells not containing any stream data. The elevation example shows the path crossing mostly through the dark red cells which are high elevation areas. And lastly, looking at the example where all three elements have an equal amount of weight, this is just about the most direct path from Dory’s farm to the North Picnic, which makes sense, because Dory is not adding any specific weight to this path in terms of streams, elevation, and land cover.

Table 6: Calculated Cost Paths of Weighted Overlays for Results Verification

100% Weighted Land Cover Cost Path	100% Weighted Streams Cost Path
	
100% Weighted Elevation Cost Path	Evenly Weighted Cost Path
	

Discussion and Conclusion

To further investigate into what would be an optimal path for Dory to take from her farm to her fly-fishing spot would be to consider more elements that may affect her path. These elements could be time of day, weather patterns, traffic patterns, likelihood of flooding, etc - anything that would affect her travels from her farm to the park. It would also be helpful to have paths with weights of each of these additional elements in the event that if there is a flood or weather then Dory has multiple options on which route to take.

Even still, using the elements I chose initially - roads, streams, elevation and land cover - still give a very good idea of the best routes Dory can take from her farm to her favorite fly-fishing spot. Exploring the different weights and the effects they have on the paths gave me a good understanding of how the weighted overlay function works as well as rasterizing and reclassifying data to be able to calculate many different types of data all together.

With Dory's specific preferences - not walking through farms, only crossing water on bridges, and gradual slope - I believe taking the path where each element is weighted equally will be preferred the most by Dory. It crosses few bridges (as seen in the 'Weighted Streams Cost Path' in Table 6) and mostly stays in forested areas (as seen in 'Weighted Land Cover Cost Path' in Table 6) but, unfortunately, does follow a slightly elevated path (as seen in 'Weighted Elevation Cost Path in Table 6).

I have actually found this lab one of the easier labs as well as one of the most useful. I could absolutely see using this distance optimization analysis in the future over a plethora of different datasets. The methods, though relatively simple, were a good reminder to geoprocessing tools that we have all used before but now we are using them in a more complicated way. The visualizations of the different weights in the cost surface rasters is a good way to compare the different rasters and their influence on the cost surface. I also enjoyed being able to fix the errors I was not able to fix from Lab 2 and get a correct answer for Lab 3. Having more time and being able to revisit the documentation allowed me to double check all of my functions and calculations and correct them.

Because this is such a 'real-life' example, this lab will be a great source to reference back to in future projects. One piece I would be interested in studying a little further would be adding more elements that may have an effect on the route which could be weather, traffic, flooding patterns, etc.

References

Cost Back Link (Spatial Analyst)—ArcGIS Pro | Documentation. (n.d.). Retrieved November 21, 2022, from https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/cost-back-link.htm

Cost Distance—Help | ArcGIS for Desktop. (n.d.). Retrieved November 2, 2022, from https://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-analyst-toolbox/cost-distance.htm

Cost Path (Spatial Analyst)—ArcGIS Pro | Documentation. (n.d.). Retrieved November 21, 2022, from https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/cost-path.htm

Creating a cost surface raster—ArcGIS Pro | Documentation. (n.d.). Retrieved November 2, 2022, from https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/creating-a-cost-surface-raster.htm

Merge Rasters function—ArcGIS Pro | Documentation. (n.d.). Retrieved November 21, 2022, from https://pro.arcgis.com/en/pro-app/latest/help/analysis/raster-functions/merge-raster-function.htm

Raster Calculator (Spatial Analyst)—ArcGIS Pro | Documentation. (n.d.). Retrieved November 2, 2022, from https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/raster-calculator.htm

Reclassify (Spatial Analyst)—ArcGIS Pro | Documentation. (n.d.). Retrieved November 21, 2022, from https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/reclassify.htm

Weighted Overlay (Spatial Analyst)—ArcGIS Pro | Documentation. (n.d.). Retrieved November 2, 2022, from https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/weighted-overlay.htm

Self-score

Category	Description	Points Possible	Score
Structural Elements	All elements of a lab report are included **(2 points each)**: Title, Notice: Dr. Bryan Runck, Author, Project Repository, Date, Abstract, Problem Statement, Input Data w/ tables, Methods w/ Data, Flow Diagrams, Results, Results Verification, Discussion and Conclusion, References in common format, Self-score	28	**28**
Clarity of Content	Each element above is executed at a professional level so that someone can understand the goal, data, methods, results, and their validity and implications in a 5 minute reading at a cursory-level, and in a 30 minute meeting at a deep level **(12 points)**. There is a clear connection from data to results to discussion and conclusion **(12 points)**.	24	**24**
Reproducibility	Results are completely reproducible by someone with basic GIS training. There is no ambiguity in data flow or rationale for data operations. Every step is documented and justified.	28	**28**
Verification	Results are correct in that they have been verified in comparison to some standard. The standard is clearly stated **(10 points)**, the method of comparison is clearly stated **(5 points)**, and the result of verification is clearly stated **(5 points)**.	20	**20**
		100	**100**