Lab Report

Title: Lab 3 Part 1 Report
Notice: Dr. Bryan Runck
Author: Kyle Olson
Date: 10/29/2024

Project Repository: https://github.com/kolson5581/GIS5571/tree/main/Lab%203
Google Drive Link: N/A
Time Spent: 1 Hours

Abstract
This lab report details the processes and implications of a weighted cost surface analysis that considers different weighting schemes that attempt to account for different ways Dory might balance her preferences. Previously, I created weighted cost surfaces that consider the impact of slope and land classification on Dory's decision making. In order to account for a range of possible ways that Dory might balance these two factors, this model weights her preferences with three separate weighting schemes to calculate the cost to Dory to travel between these locations.

Problem Statement
Part 1 focuses on the problem of determining the best path for Dory to take between her house and her favorite fishing spot, taking in consideration her preferences for travel. The previous lab report focused on the construction of the preferences and cost surfaces, while this one focuses more on the implications of the weighting schemes.

Table 1. Requirements for Part 2

#	Requirement	Defined As	(Spatial) Data	Attribute Data	Dataset	Preparation
1	Elevation data	30m resolution DEM for state of MN	Raster	Elevation	Minnesota Digital Elevation Model - 30 Meter Resolution	Download, resample, convert to slope
2	Land classification data	15m resolution land class data for state of MN	Raster	Land use/land class	Minnesota Land Cover Classification and Impervious Surface Area by Landsat and Lidar: 2013	Download, resample

3	Start and end points	Dory's home and favorite fishing spot	Vector data (points)	none	Created manually	Points placed at locations provided in lab

Input Data
Both datasets used for this problem are found on the Minnesota Geospatial Commons. In order to complete the lab I needed elevation data and I needed land classification data that I could trust would be reasonably accurate. I found a 30m DEM for the entire state of Minnesota which I used for the elevation data. There were a number of Land classification datasets available - I chose this particular set because it had a finer spatial resolution (15m), and I thought it would be better able to capture water, especially in narrow rivers or stream, which Dory wanted to avoid.

Table 2. Input data

#	Title	Purpose in Analysis	Link to Source
1	Minnesota Digital Elevation Model - 30 Meter Resolution	Provide basis for slope calculations.	Minnesota Digital Elevation Model - 30 Meter Resolution
2	Minnesota Land Cover Classification and Impervious Surface Area by Landsat and Lidar: 2013	Provide basis for land class preferences for routing Dory	Minnesota Land Cover Classification and Impervious Surface Area by Landsat and Lidar: 2013

Methods
The methods to develop this lab report are essentially the same as part 2 of Lab 2. The full method section is preserved below, but no changes to the code or methods were required to complete this section.

First, I created my start and end points manually based on the location data provided in the lab assignment. Then I requested the DEM and the land use rasters I would use for my analysis and saved them to file. These rasters covered the whole state, however, so I needed to reduce them to my study area.

To define the study area, I drew buffers of 3 miles around the start and end points, hoping that would provide enough area for Dory to navigate. Once I had the buffers, I determined their extent using Arcpy, and used that to inform the extent I would use for the "Extract by rectangle tool" - limiting the scope of the rasters I would use for the rest of the analysis. I needed to convert the elevation values in my DEM to slope figures, so I did that next. Then I needed to resample one of the rasters so they had matching pixels for the raster algebra. Normally I would resample to match the lower resolution image, however in this case I decided to resample the slope - because it is continuous data, it is safer to interpolate the small differences between neighboring cells. Additionally the land use data is more discrete classes, so resampling might alter the boundaries of land uses in unhelpful ways.

After resampling, I needed to reclassify the data from the raw slope/land use data into scores that would correlate with Dory's willingness to traverse that cell. Table 3 has the reclassified land use and slopes - lower values indicate more willingness to travel across.

Table 3. Reclassification schemes

Land Use	Land use class	New Score	Slope (degrees)	New Score
Imperious/urban	1-100	1	0-3	1
Wetlands	101-102	4	3-6	2
Open water	103	4	6-10	3
Extraction	104	3	10-20	4
Forest	105-107	3	20-90	5
Managed Grass/Natural Grass	108	2		
Agriculture	109-110	5		

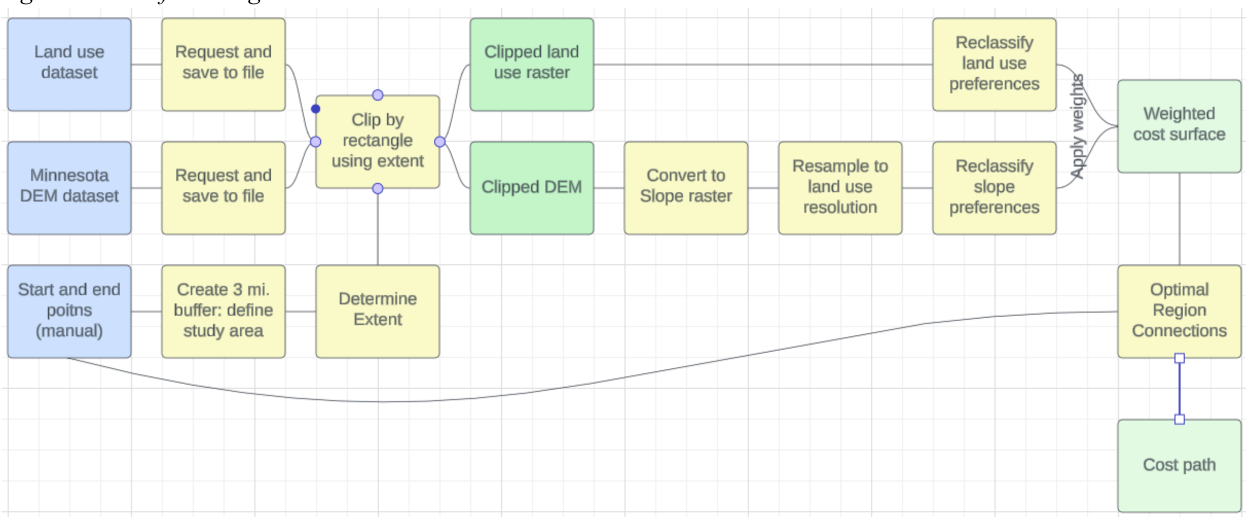
We knew Dory did not want to cross farm fields, so agricultural classes were scored 5. We also knew that water was something she preferred to avoid, so open water and wetlands were scored 4 respectively. Forest and extraction land both seem less desirable to cross than the remaining options so were scored 3. Managed and natural grass and then impervious and urban land are the easiest to cross, so they received scores of 2 and 1 respectively.

For slope, I wanted to really favor low slopes, so 1-3 are all less than 10 degrees of slope, above 10 is a 4, and above 20 was scored a 5. These decisions are somewhat arbitrary, but I believe are defensible based on what we know about Dory's preferences.

After reclassification, I converted the rasters to arcpy raster objects so I could do raster algebra without any other tool. To do the raster algebra, I created a list of weights [.3, .5, .7], and used a for loop to iterate through the weights, applying them to the appropriate raster, and saving each weighted cost surface to file, and adding each file name to a list I would use for the path calculation step.

Lastly, I used those cost surfaces and the starting and ending points in the "Optimal Region Connections" arcpy tool to generate cost paths for each of the weighting scenarios. These paths are added to the map so we can see the path Dory could take, along with a path cost, which allows us to see the total "cost" to Dory for taking that path.

Figure 1. Data flow diagram.



Results

Figure 2 shows the three cost paths generated with the three weighting schemes considered in this lab mapped out. As you can see, the two cost paths which were weighted more towards land classification have more similar routes. Enabling the land class raster (Figure 3) makes it clear that these routes are following roads more closely - which makes sense as these cells were given lower costs and Dory cares more about land class in these scenarios.

Figure 2. Cost paths mapped

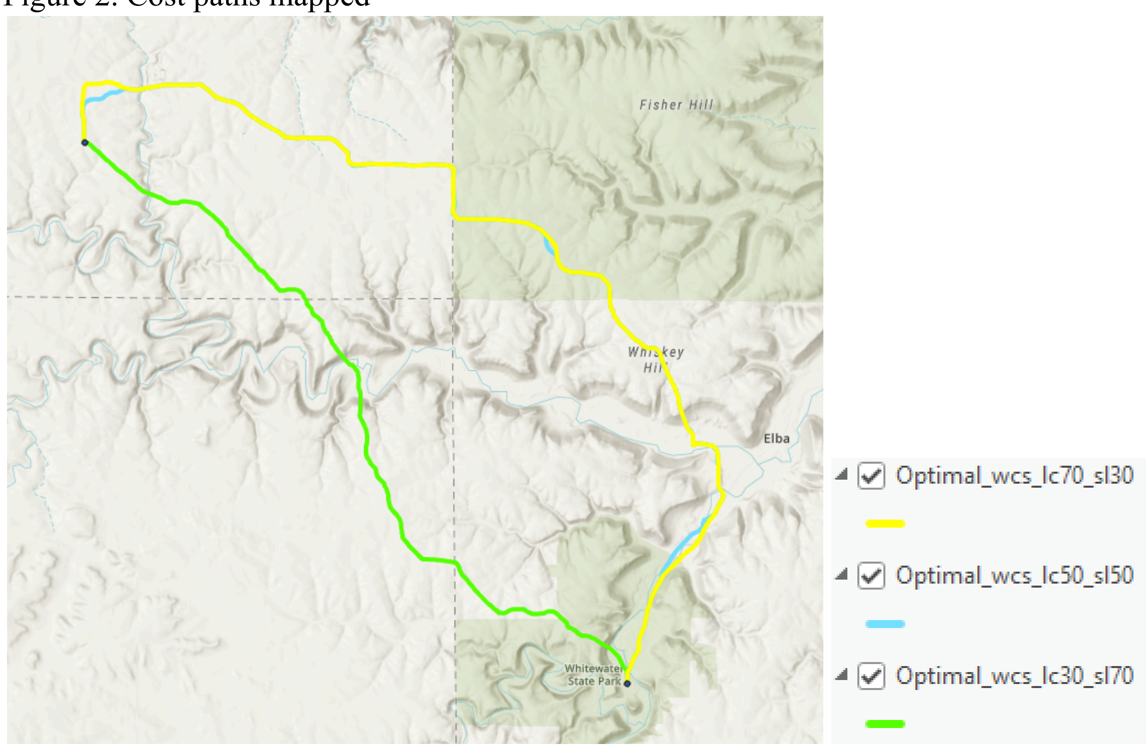
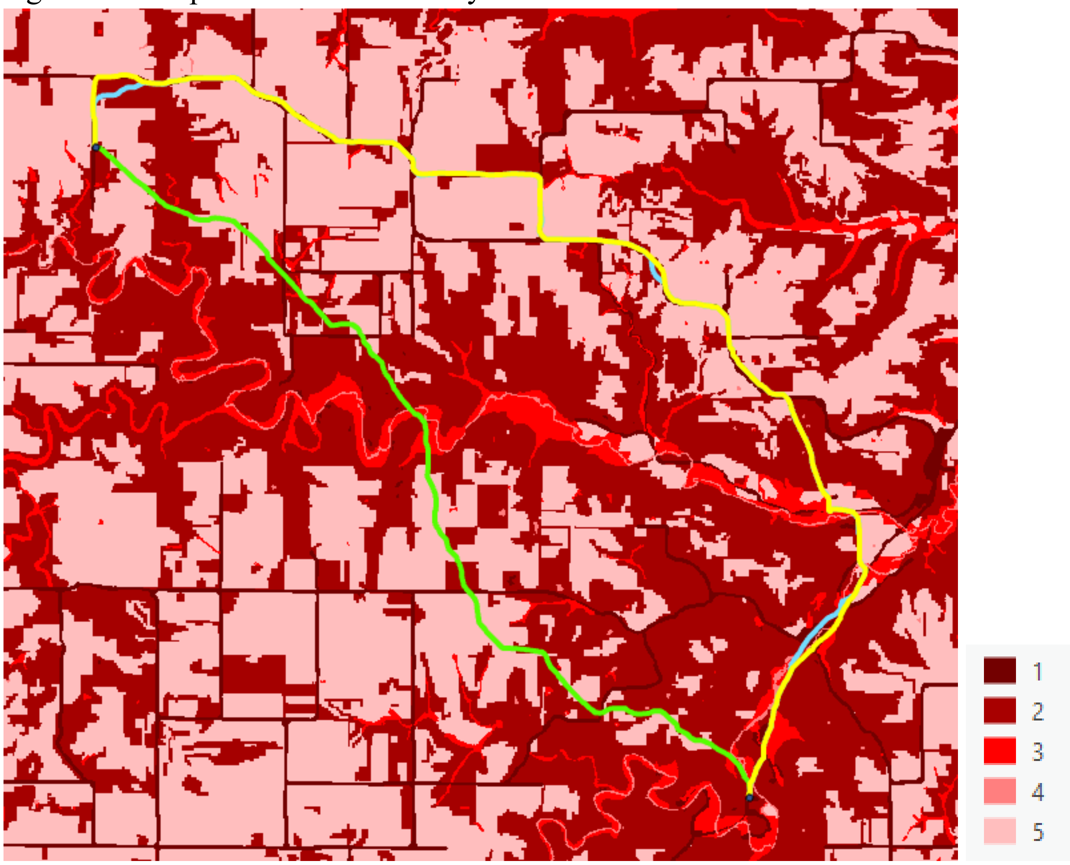
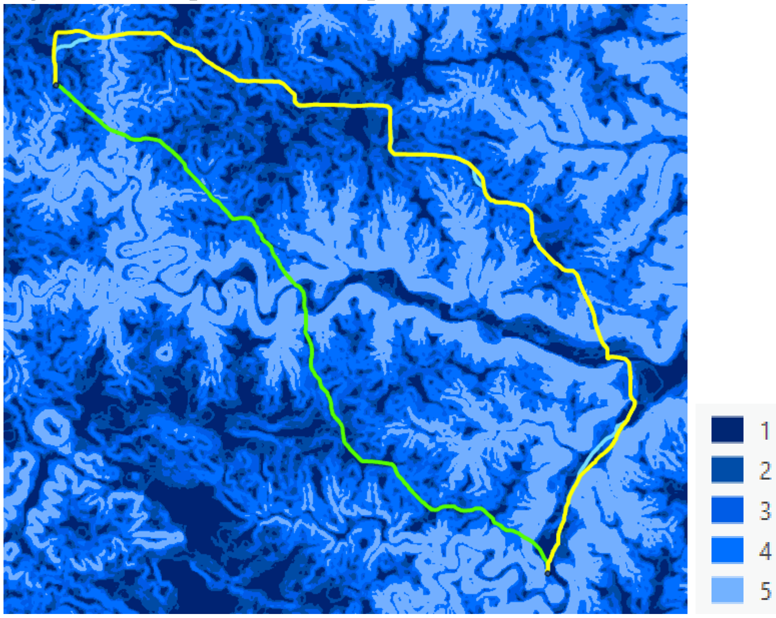


Figure 3. Cost paths with land use layer enabled



The green cost path is a bit harder to understand, though by enabling the slope layer, we can see that the path generally is sticking to areas with lower slope scores - the darker areas (Figure 4). It crosses some steep areas, but they are relatively short, so Dory is not picking up too much cost relative to the total length of the path by making these choices.

Figure 4. Cost paths with slope classes visible



It is clear that even in a situation where the two factors are weighted the same (blue line), that the land cover values are having a larger impact on the overall path, as it follows the weighting scheme that favors land cover (yellow line), more closely. It is only when we more heavily weight the slope classification (green line) that a more distinct path emerges. Additional weighting testing could reveal other potential paths and allow us to determine exactly at what point the routing changes significantly.

Table 4 shows the total path cost for each of the three optimal routes along with their weighting scheme.

Table 4. Path cost for each weighting scheme

Name	Slope weight	Land cover weight	Path cost
Optimal_wcs_lc30_sl70	70	30	32096.1152348
Optimal_wcs_lc50_sl50	50	50	31315.5918
Optimal_wcs_lc70_sl30	30	70	26982.17188

We can see that the weighting scheme which favors land classification is the cost path which produces the least amount of cost to Dory. However, unless we know more about what Dory's preferences truly are, we cannot claim that this is actually the "best" path. These scenarios are only potential scores given many assumptions about Dory's preferences.

Results Verification
At each step of the pipeline, I was checking the outputs to ensure that they were making sense and were ending up in the correct location. Additionally, the fact that two of the weighting schemes produced similar results suggests that the weighting was having the intended impact.

Discussion and Conclusion
Ultimately the most interesting part of this analysis remains the understanding that it provided about the way that our input data and assumptions impact the outputs we get. I am not able to fully understand Dory's preferences, and had to make some assumptions about how she prefers to travel to her favorite fishing spots. That includes assumptions about how to reclassify the two input rasters to match her preferences.

However, for this lab, I focused more on how multiple weighting schemes have an impact on the overall path that is optimal for her to take. It's clear from the results that there is a weighting scheme which produces the least overall "cost" to Dory, but this does not mean it is the best path for her to take. If indeed Dory values land cover more than the difficulty of slope crossing, then we could conclude that she should take that route. But because I had to make guesses about how to balance her preferences, I can only consider this to be an estimate.

The lesson to take away from this lab and the previous lab is that model builders need to be aware of what is driving the "best" results from their models. Both the classifications and the weighting schemes are important factors, and should be carefully considered and checked against model goals.

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**