Lab Report

Title: Lab 3 – Solving Dory's Problem
Notice: Dr. Bryan Runck
Author: Mohsen Ahmadkhani
Date: 11/1/2021

Project Repository: https://github.com/mohsen-gis/GIS5571.git
Google Drive Link: -
Time Spent: 15 hours

Abstract
In this lab, I continue of developing an ETL process that proposes a few more optimal paths for Dory, who wants to go for Fly-Fishing in a state park near her house. I build a few more combinations of criteria for Dory to take this path.

Problem Statement

Table 1. The list of required data sets for the proposed study.

#	Requirement	Defined As	(Spatial) Data	Attribute Data	Dataset	Preparation
1	Landcover	The classification of land cover for the state of Minnesota	TIF Raster Zip file	Pixel values (landcover class)	MN GeoCommons	ETL
2	DEM	Elevation data for Winona, Wabasha, and Olmsted counties	TIF Raster geodatabase	Elevation	MN GeoCommons	ETL

Input Data

Table 2. Input data description.

#	Title	Purpose in Analysis	Link to Source
1	Landcover	For building the cost surface and find the optimum path	API
2	DEM	To convert to slope and build the cost surface and find the optimum path	API Source

Methods
In this part we were supposed to create a cost surface and propose a path from the origin to the destination. The origin point is at 44.127985, -92.148796 and the destination point is North Picnic area of Whitewater State Park. The subject is named Dory and here are her preferences:

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, though sometimes she doesn't mind if she's wearing her waders. Other than that, she just wants to take the path that is the most gradual in terms of slope.

Initially, based on the criteria described by Dory, I interpreted it as follows that the proposed path should:

- Avoid farm fields
- Avoid water bodies
- Avoid sharp slope

For this purpose, I took the data flow diagram depicted in figure 1. I first created a Point feature class for having two points as the origin and destination points. Next, I created a bounding box covering the area covering the origin and the destination. Next, I downloaded three sets of DEM files from MN GeoCommons for the three intersecting counties named Winona, Wabasha, and Olmsted. Then, clipped them each to the bounding box then all merged all the three clipped raster files to create a single merged raster. To clip the raster I used arcpy.management.Clip function. And the merging task was done using arcpy.ia.Merge function. Then, using the arcpy.sa.Slope function, I created the slope raster from the inputting DEM. The same clipping approach was done for the landcover raster file. Then I resampled all inputting raster files to 3-meter pixel sizes using arcpy.management.Resample function. Then, I reclassified the landcover data to field and non-field, and water and non-water classes using the arcpy.Reclassify_3d function. And finally I created the cost surface using the arcpy.sa.RasterCalculator function. Finally, I created the proposed optimum path using the OptimalRegionConnections function. The entire process is depicted in figures 12 to 16. Then, I started to build four different combinations of the criteria to experiment how different preferences impact the proposed path. It is noteworthy that to increase the computation speed, I down sampled the images from 30 meters resolution to 100 meters resolution. The produced images are displayed in figures 17 to 20.

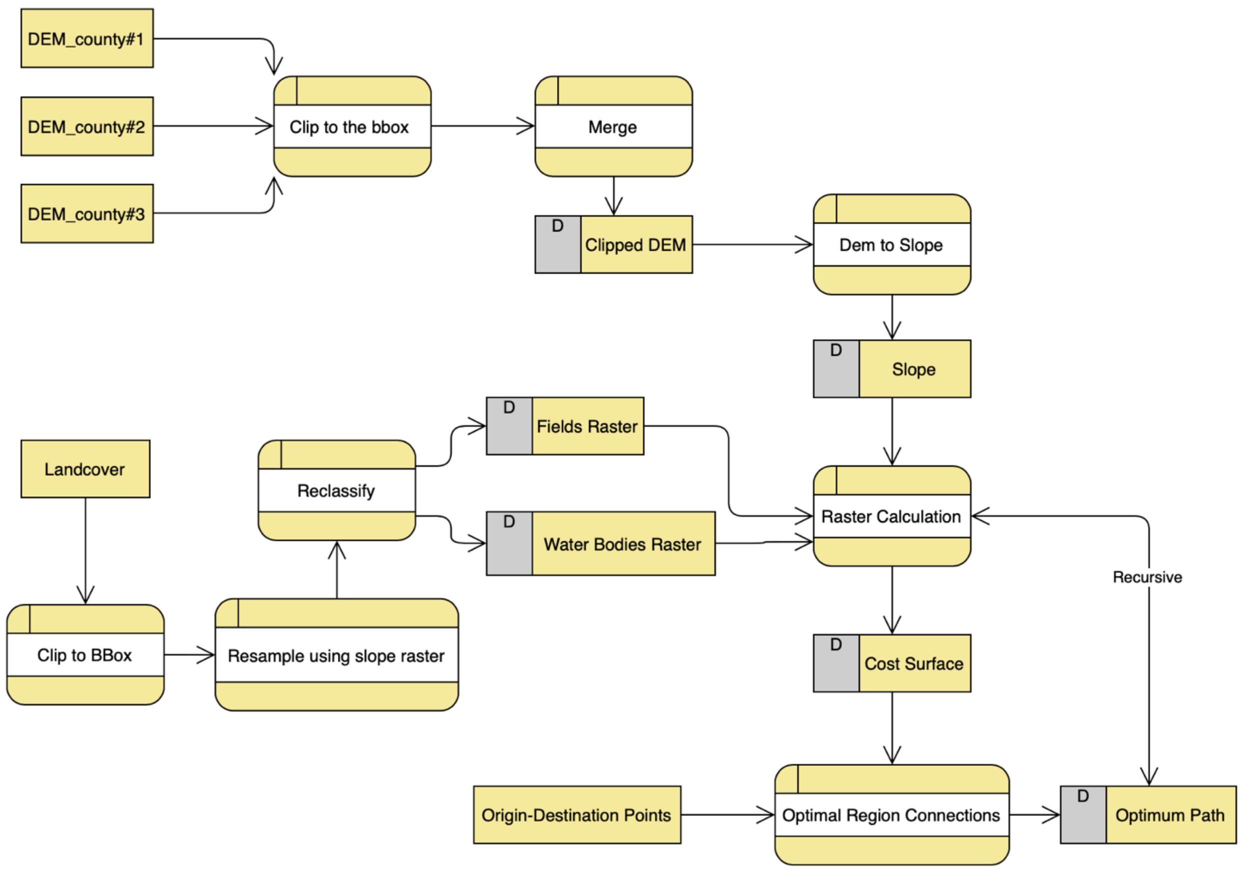

Figure 1. Data flow diagram.

Results

The results for this project are depicted in figures 2 to 10. These 9 figures are showing the outputs of all steps from data preparation to map algebra and cost surface creation. In this project I created six different combinations of criteria and cost surfaces. First, I considered the weights of -1 for Slope, -10 for farm, and -1 for water bodies. The resulting cost surface and corresponding optimal path are shown in figure 5. Next, I took another weighting approach considering the weights of -1 for Slope, -4 for farm, and -3 for water bodies. The resulting optimum path and cost surface are provided in figure 6. The extended part of this lab is experimenting the criteria combination of Slope (2), farm (1), water bodies (-2) shown in figure 7, Slope (-2), farm (-1), water bodies (+2) depicted in figure 8, Slope (-3), farm (-2), water bodies (-2) presented in figure 9, and finally Slope (+3), farm (+2), water bodies (+2) provided in figure 10.

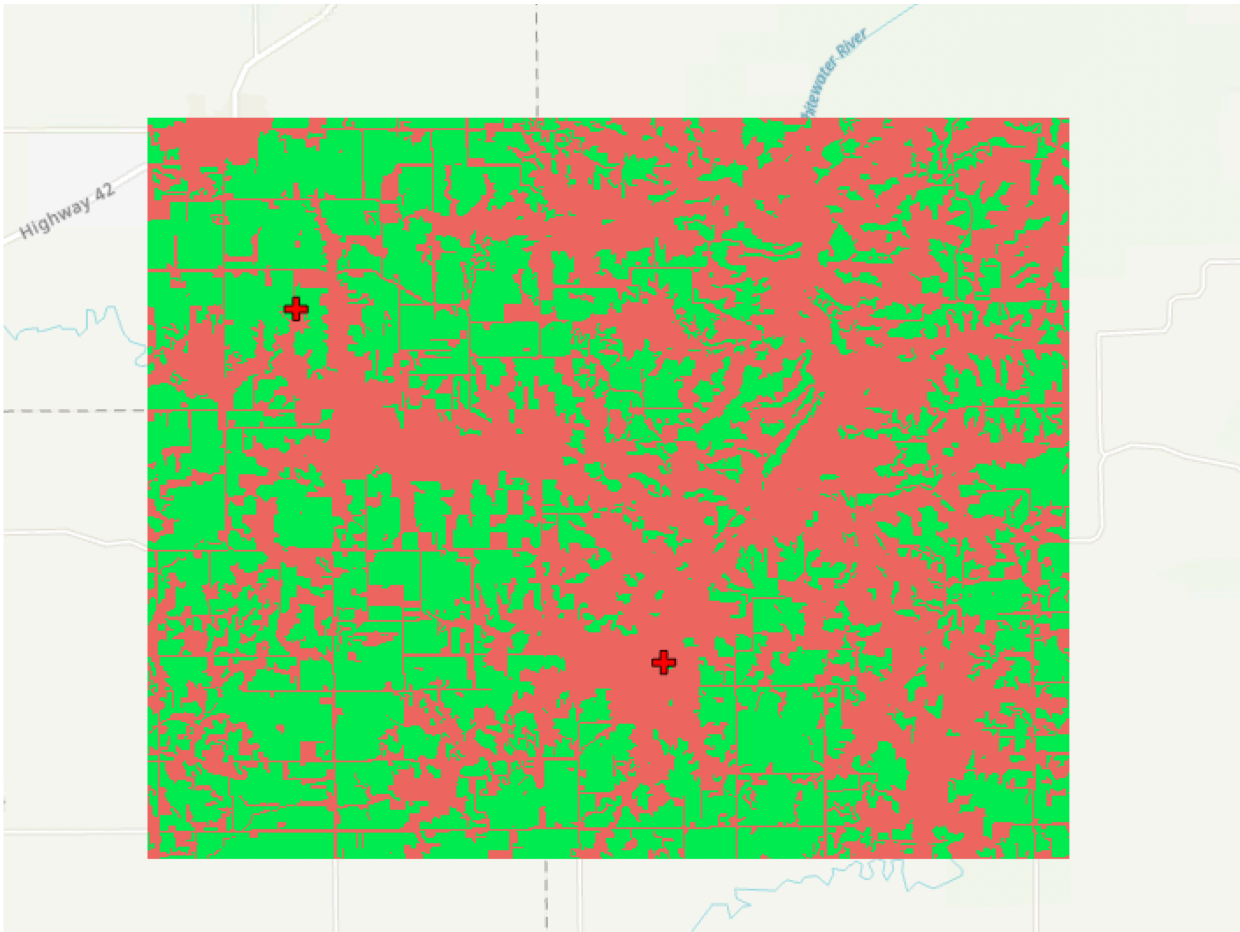

Figure 2. Farm (green) and non-farm (red) criterion raster (red cross marks indicate origin (top-left) and destination (bottom-right)).

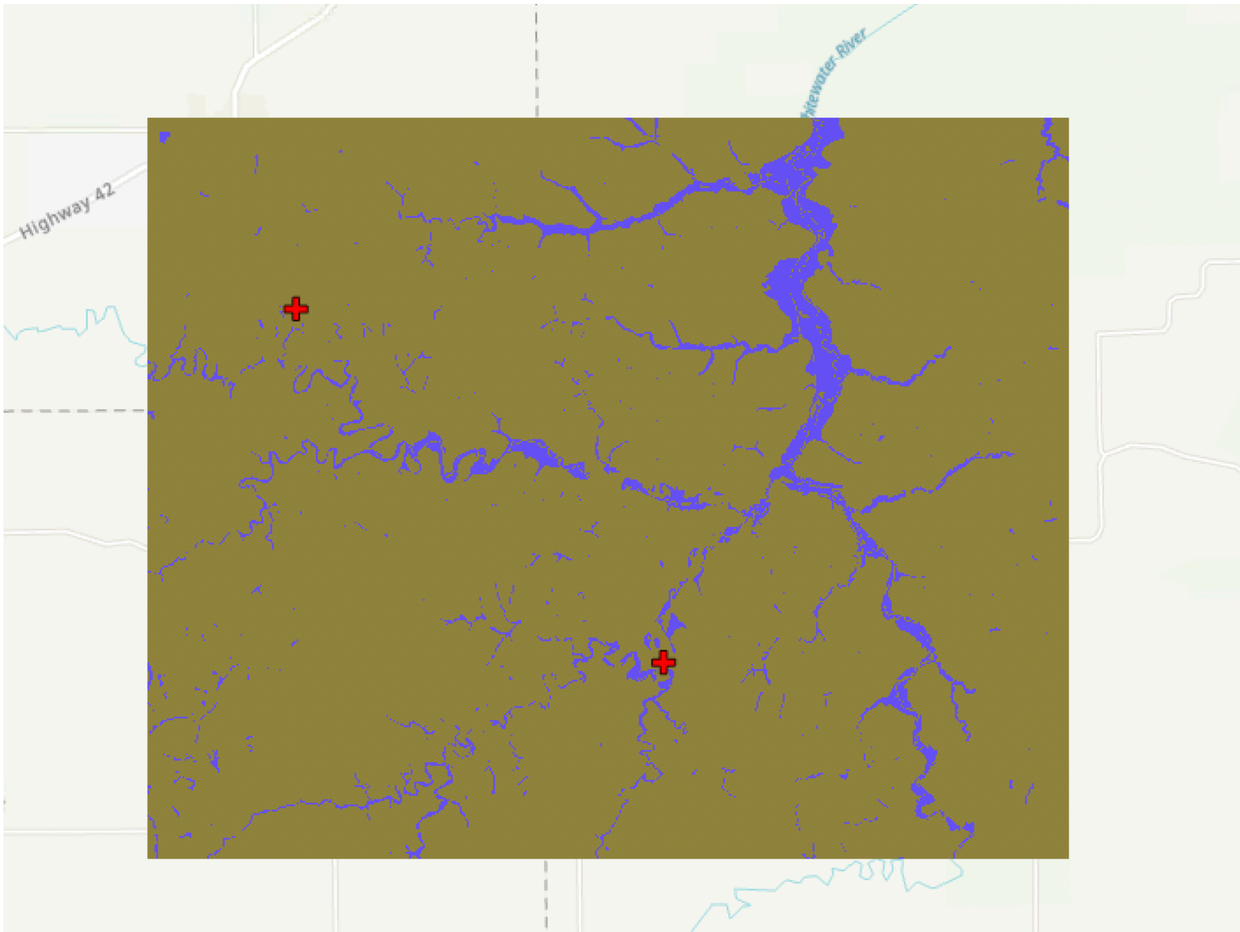

Figure 3. Water bodies criterion raster.

Figure 4. Slope criterion raster.

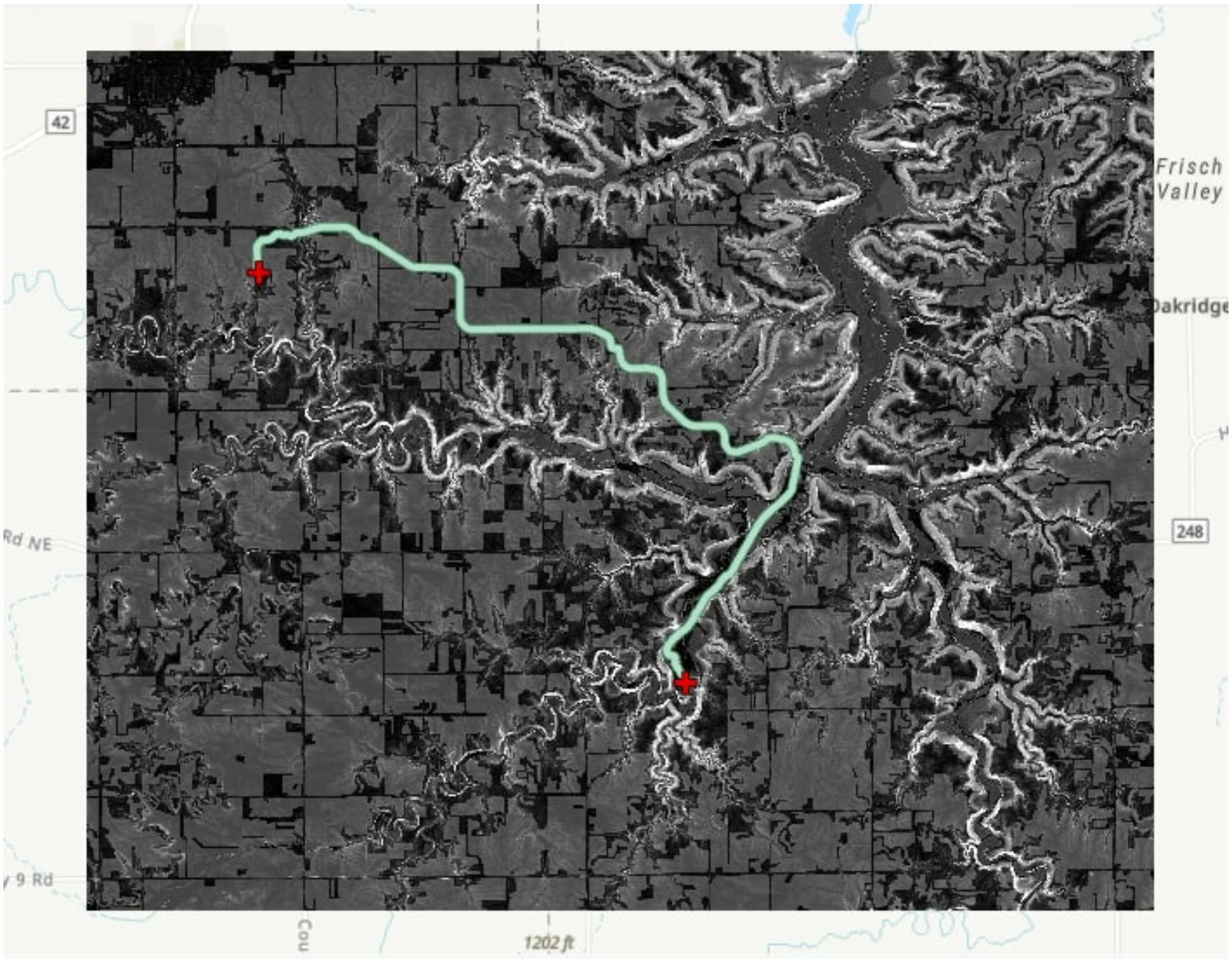

Figure 5. Scenario 1 resulting cost surface and optimum path: Slope (-1), farm (-10), water bodies (-1).

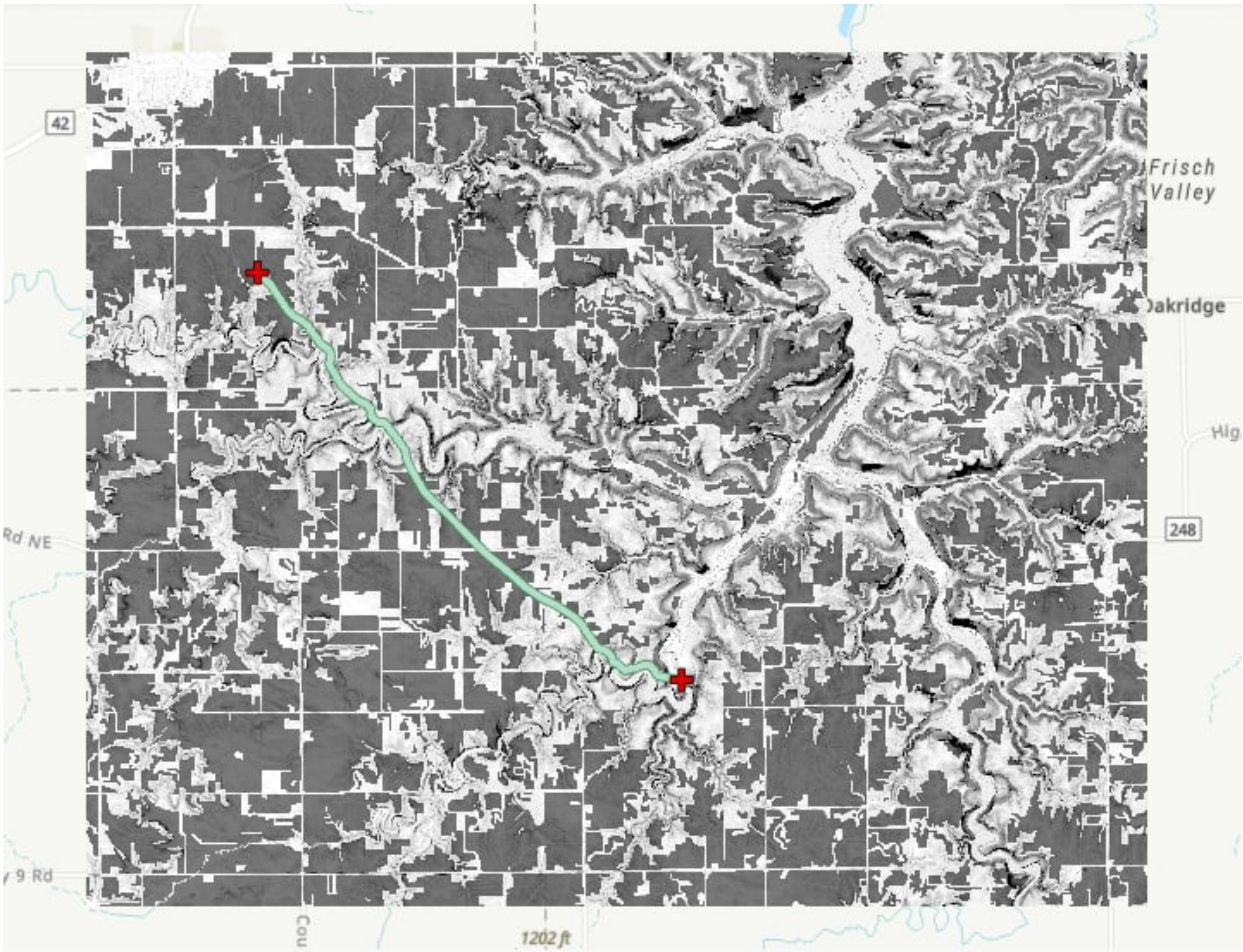

Figure 6. Scenario 2 resulting cost surface and optimum path: Slope (-1), farm (-4), water bodies (-3).

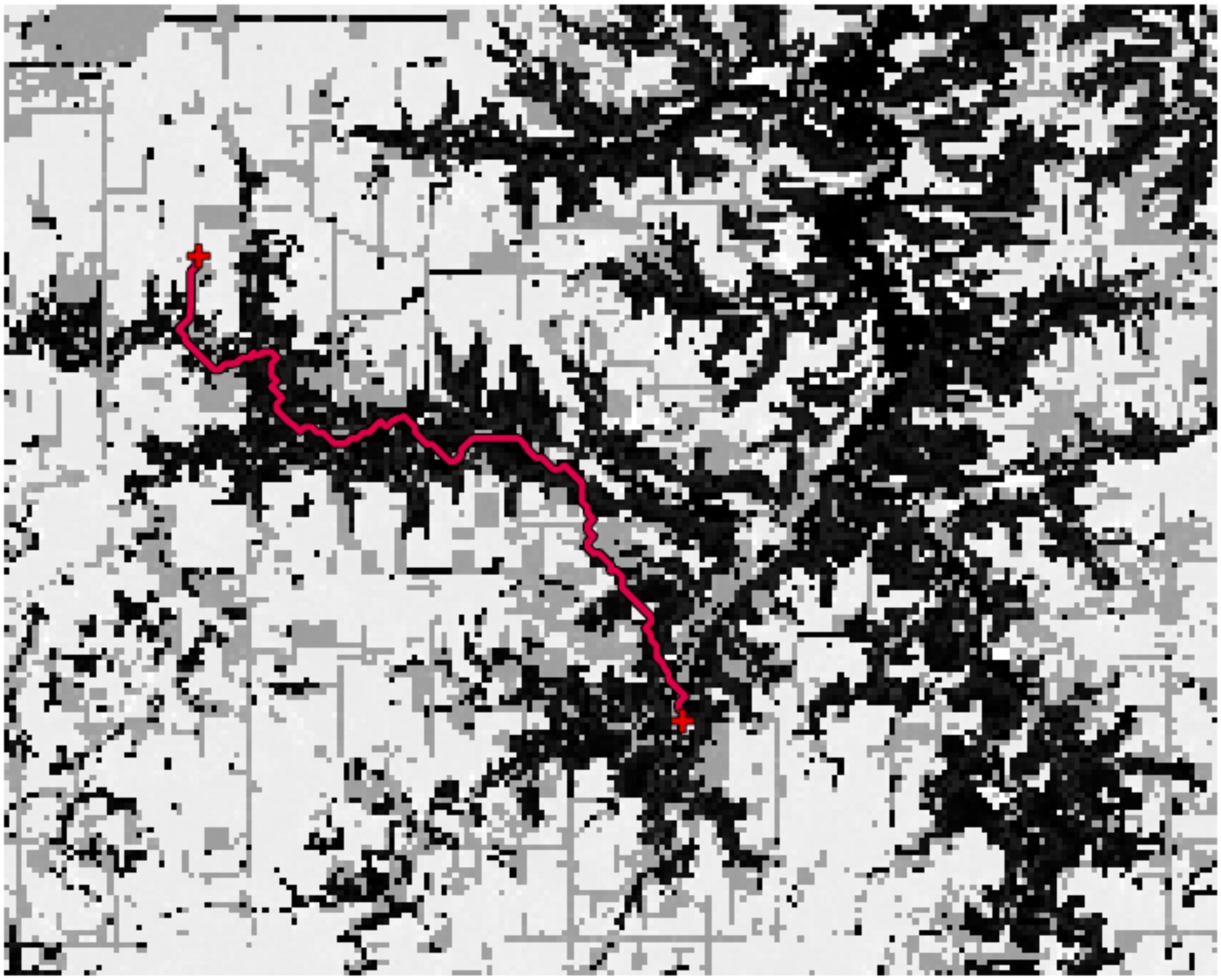

Figure 7. Scenario 3 resulting cost surface and optimum path: Slope (2), farm (1), water bodies (-2).

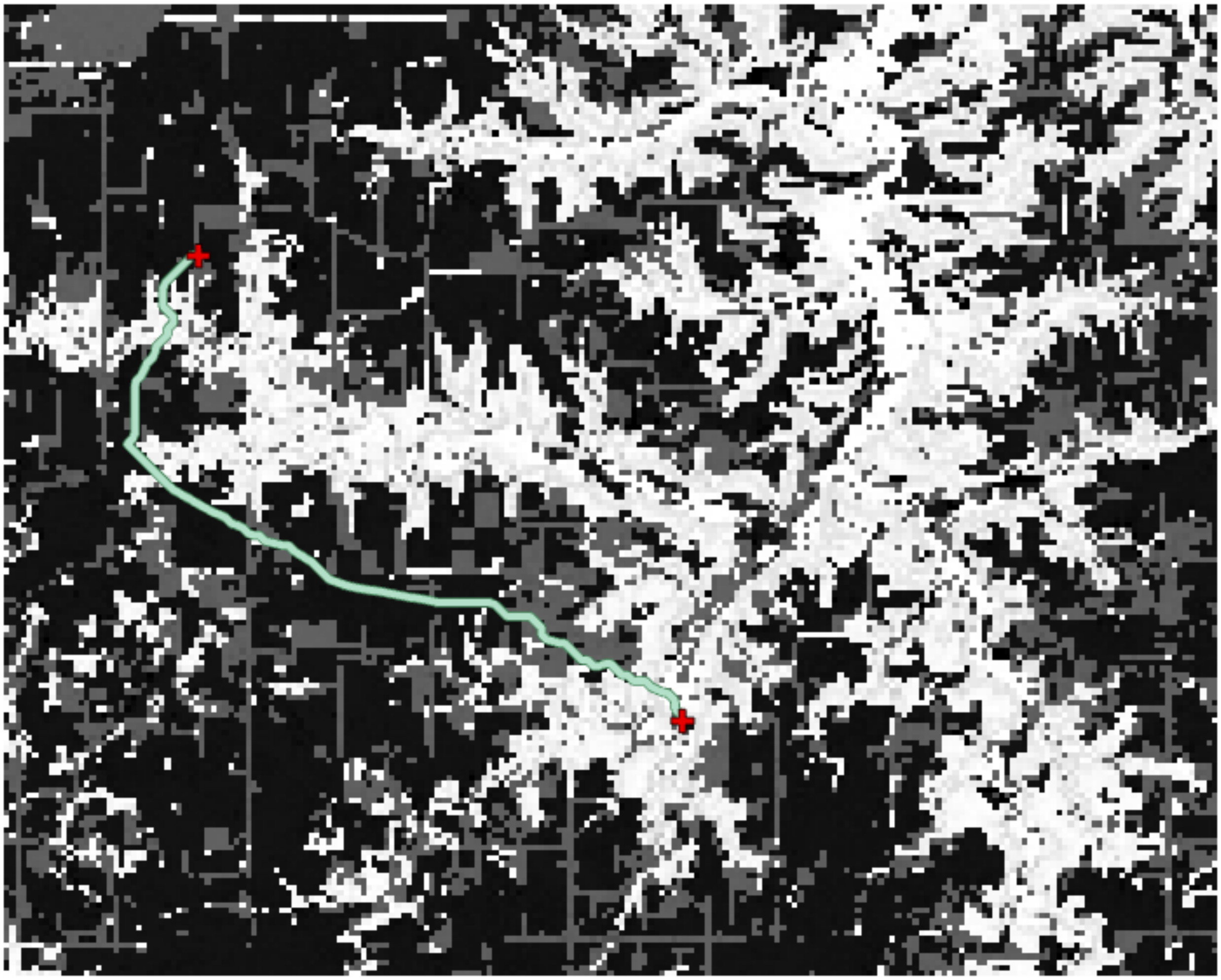

Figure 8. Scenario 4 resulting cost surface and optimum path: Slope (-2), farm (-1), water bodies (+2).

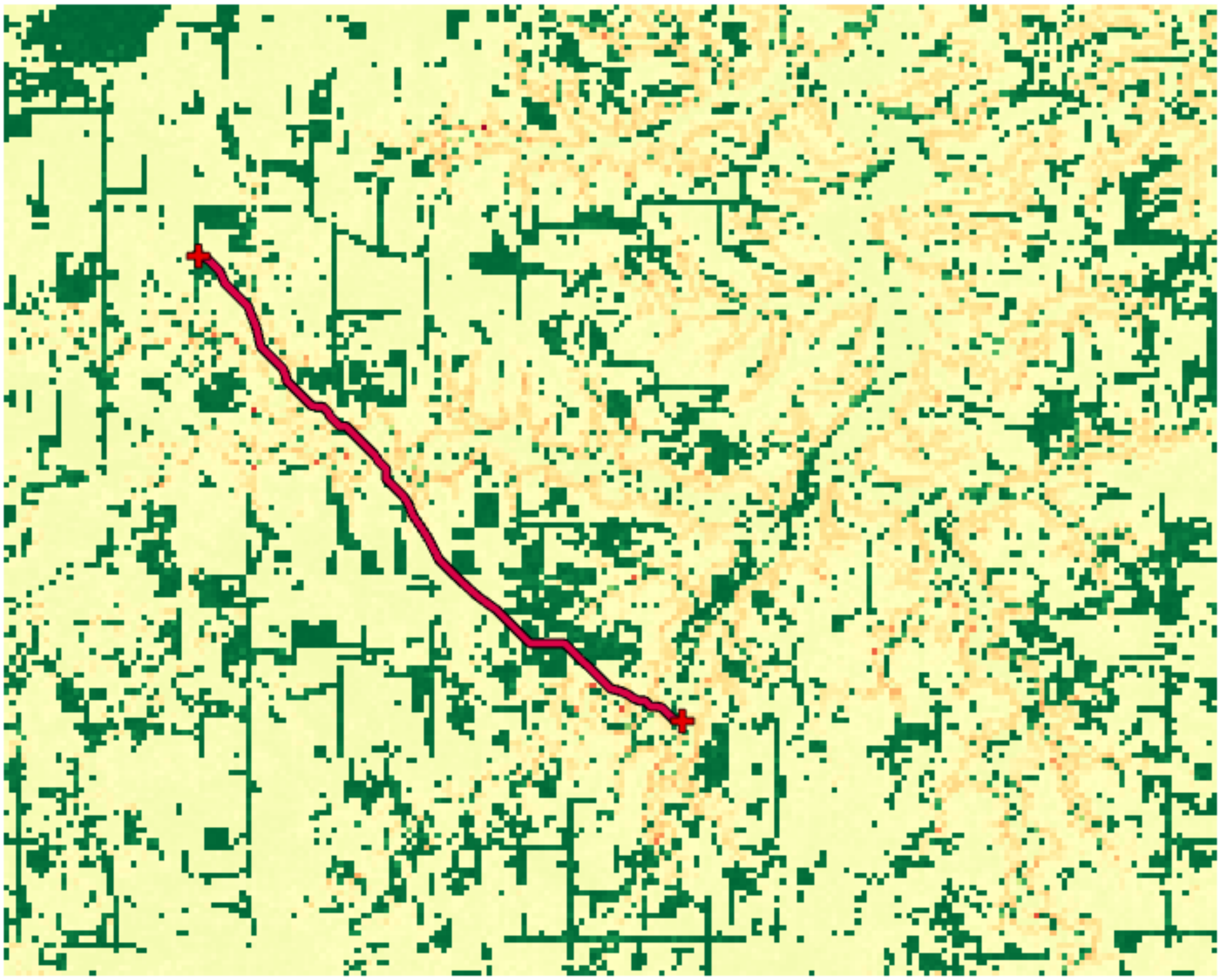

Figure 9. Scenario 5 resulting cost surface and optimum path: Slope (-3), farm (-2), water bodies (-2).

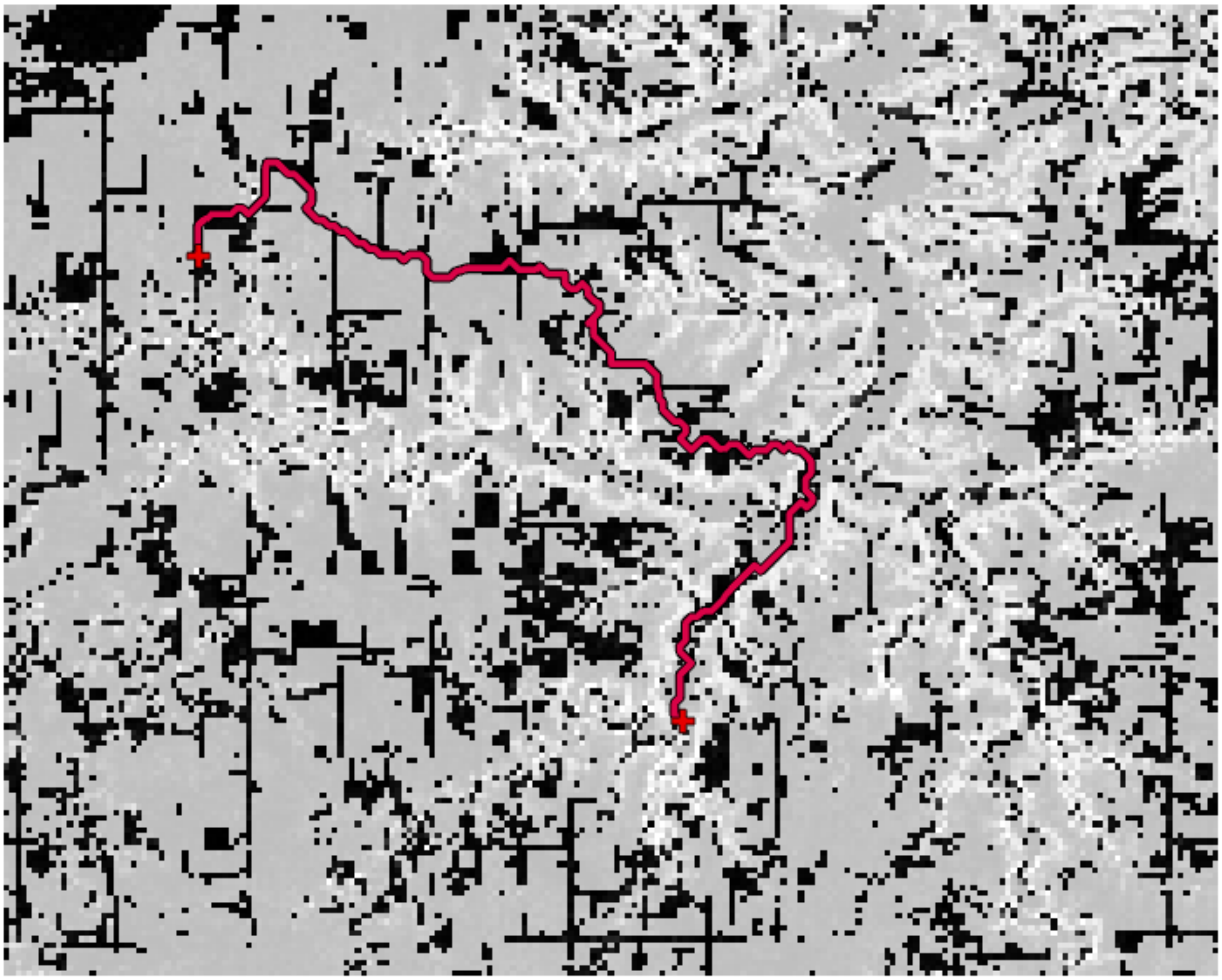

Figure 10. Scenario 6 resulting cost surface and optimum path: Slope (+3), farm (+2), water bodies (+2).

Results Verification

Verification for each step was done by comparing the results with the hill shade of the base map in the ArcGIS Pro context. In other words, I used visual verification of the results in this step. Also, since the data datasets were downloaded from verified sources like MN GeoCommons and PRISM, they are already verified and are trustworthy. I also double checked to have consistent and unified geographical projection systems for all files.

Discussion and Conclusion

It was a practical experience working with raster files and run map algebra and raster calculation to generate a cost surface model. The most significant challenge/limitation that I had was the load of the project. All the processes required a strong pc and high memory so that the execution be successful and fast. Unfortunately, I had to use my desktop computer at my office remotely which is not a strong system. Most of the time the computer was crashing and running out of memory so that all the process had to stop until I problem shoot the computer on campus. Also, I was hoping to be able to create many more maps for the cost surface part with many different combinations of the weights, but I couldn’t run more than six times.

API for data collection

All the APIs are provided as the links in table 2.

References

None.

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	**26**
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	**21**
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	**95**