Lab Report

Title: Lab 3
Notice: Dr. Bryan Runck
Author: Alexander Edstrom
Date: 12/1/21

Project Repository: *GIS5571/Lab3 at main · AlexanderEdstrom/GIS5571 (github.com)*
Time Spent: *8 hours*

Abstract

This lab will go over the process of using a cost surface to generate an optimal path. The lab will explore how different weighting schemes of the preferences will affect the optimal path chosen. The three paths will all start and end in the same places, making the differing decisions made by the changed weights most apparent. Between each of the paths being generated, the cost surface will be regenerated using different weights.

Problem Statement

The main focus of this lab will be recreating the cost surface each time to reflect a change in preference. This will be done by shifting the weights to or changing the data input to reflect the change in criteria of preference based on each situation. The final output will be three optimal paths, one from each set of weights and preferences. The construction of the original cost surface will be covered briefly. These requirements are shown in table 1 below.

Table 1. Requirements

#	Requirement	Defined As	(Spatial) Data	Attribute Data	Dataset	Preparation
1	Cost Surface	Cost surface generated in previous lab	Combination of layers to generate cost		Combination from USGS and MNGeoSpatial Commons (detailed in data section)	
2	Weights	The numerical reflection of real life parameters and preferences based on Dory's requirements and wants.	Map algebra done during cost surface creation			

Input Data

The data used to create the cost surface consists of a Sentinel 2 image of the area as well as a DEM from USGS. The image was obtained from USGS Earth Explorer using the online portal and the DEM was obtained from the MN GeoSpatial Commons hub using an API request. The weights are self generated based on the specifications laid out in the lab introduction. All of the external input data used is shown within table 2 below.

Table 2. Input Data

#	Title	Purpose in Analysis	Link to Source
1	MN DEM	USGS standard DEM clipped to the extent of MN, used in cost surface	Mn GeoSpatial Commons
2	Satellite Imagery	Sentinel 2 imagery of southeast MN, used in cost surface	Earth Explorer

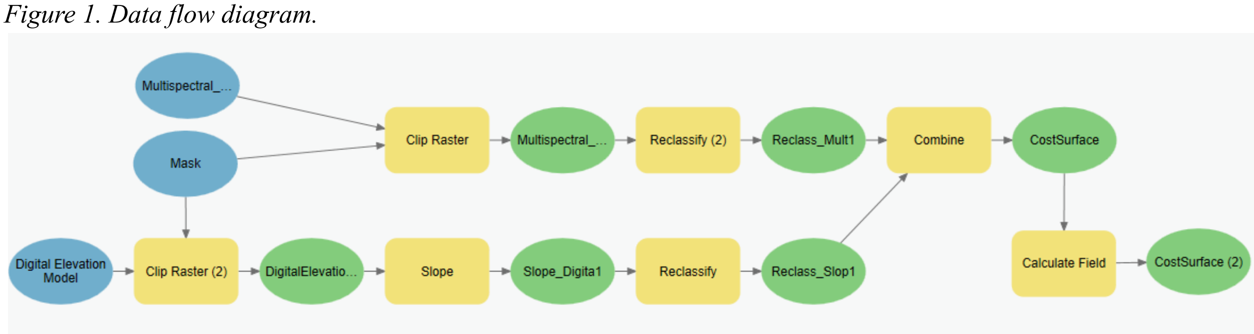
Methods

Creation of the cost surface:

The notebook for this lab found within the repo includes the process of creating the cost surface. The cost surface will combine the effects of slope and landcover to attach a travel cost to each cell. Starting with the Sentinel 2 image, and the DEM raster, a mask polygon was created to separate only the study area from the rest of the data in the layers. This is done to speed up computation times for the classification function as well as the slope calculation. First the slope is generated from the DEM clip by taking the first derivative of the elevation. The slope raster is then reclassified to a 1-10 cost scale, the greeter the slope, the higher the cost to represent the difficulty associated with traversing a steep surface.

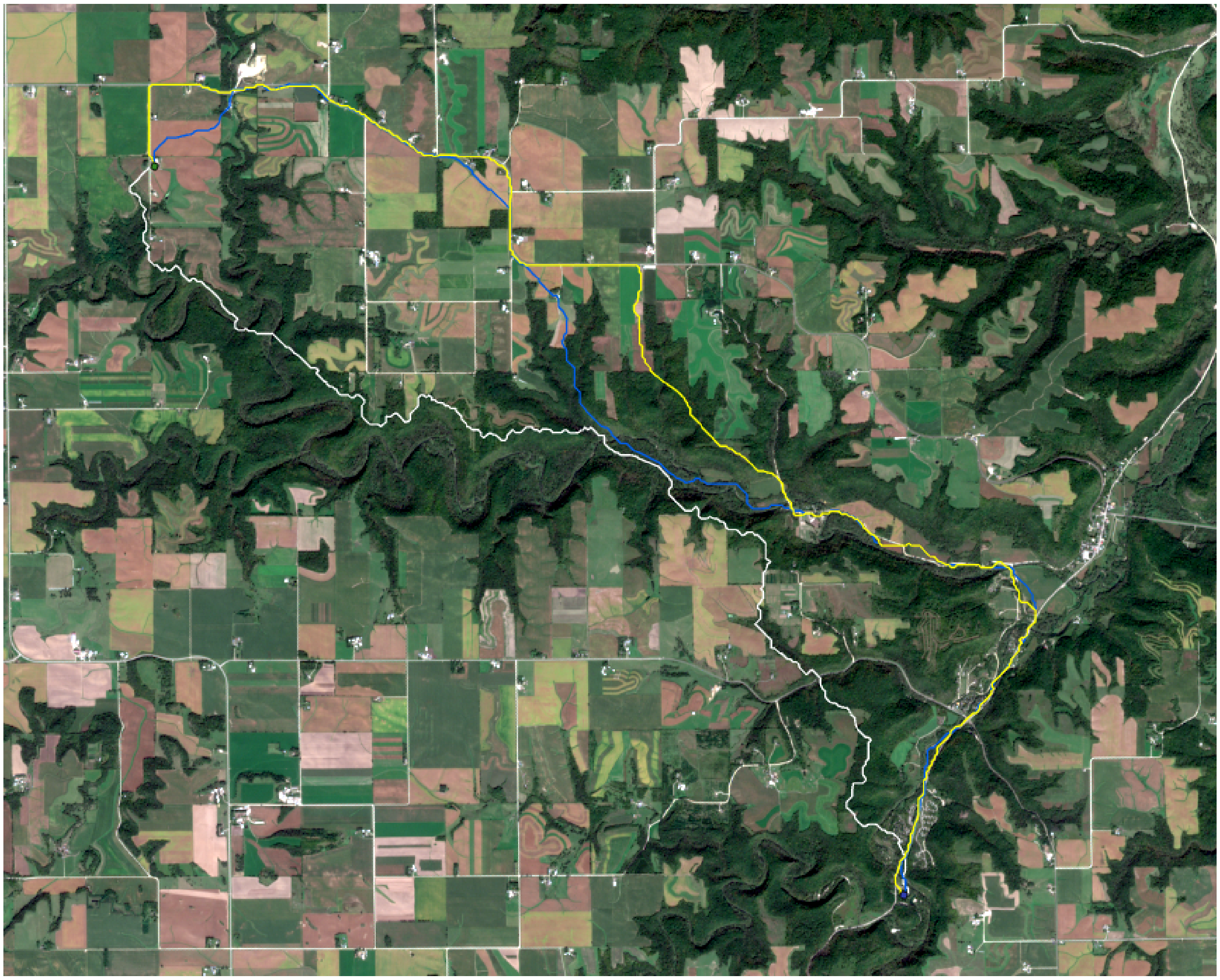
The Sentinel 2 image clip is then classified using a supervised pixel based random trees method. The image is classified into land use to which a travel cost can be associated with. These cost values are based on the preferences outlined in the prompt. The slope and the classified rasters are then combined and a new field is created to simply add the two cost variables. This means that both of the two variables are weighted the same amount. The workflow for this is shown in figure 1.

Figure 1. Data flow diagram.



The creation of the optimal path using this cost surface consists of generating a distance accumulation raster and a back direction raster based on the source of the route (in this case Dory's house) and the cost surface. After that, an optimal path can be found by inputting the destination, distance accumulation raster, and back direction raster. This will output a single most optimal path. To generate multiple paths according to different preferences, the calculator for cost in each of the slope and land cover rasters can be multiplied by a percent before they are added to build the cost surface. I will explore three different combinations of weights to show how the path can differ based on preference weights.

Figure 2. The 3 optimal paths.



Results

The main results are summarized within figure 2 above.

Path 1:
The first path is unweighted, both the slope and the land cover have no percentage associated with them to alter the impact they have on the path. This path is shown in figure 2 as the white line. This one favored the forested areas, only leaving the forest where a road intersected the treeline. It looks to have crossed the river once at a very narrow spot.

Path 2:
For this path the weight of the slope was increased heavily, the ratio between the two variables was set to .7x for the slope and .3x for the land cover. This path is shown in blue in figure 2. This path mostly follows the road, which makes sense as it would almost certainly have the least elevation change from cell to cell by default.

Path 3:
For this path the weight of the land cover was increased heavily; the ratio between the two variables was set to .7x for the land cover and .3x for the slope. This path is shown in yellow in figure 2. This path follows the road even closer than path 2 does, this also makes sense as roads/developed areas are labeled as the easiest land cover type to traverse in the classification cost formula.

Results Verification

Qualitative verification is somewhat straightforward. Looking at the original image while keeping each set of weights/parameters in mind will let you come to a general idea of what an ideal path would look like. In the case of path 1 it makes sense that it would follow the forest in the most direct path to the end point while only crossing the river if absolutely necessary. Quantitative verification is harder in this situation as the travel preferences were converted into number values by my discretion, so there is an implicit bias to the weights of the variables due to my interpretation of the prompt.

Discussion and Conclusion

The main purpose of this project was to explore how changing the weights of the cost surface can influence the optimal paths produced even though the same process is being done to the same starting data. There is a certain amount of uncertainty baked into the process of bridging the gap between quantitative and qualitative data. There are certainly methods, like competing hypotheses, that can reduce the effect of this uncertainty but they require extra set up and consideration.

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	**22**
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	**25**
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	**18**
		100	**93**