Lab Report 3.1

Title: Cost surfaces, optimal paths, and weighted multi-criteria decision analysis
Notice: Dr. Bryan Runck
Author: Evening Hade
Date: 29 November 2024

Project Repository: https://github.com/eveningsaria/gis5571/tree/main/Lab3
Google Drive Link: N/A
Time Spent: 15 hours (excluding Lab 2, Part 2 freak-out time)

Abstract

Using API queries, elevation and land cover rasters were downloaded from MNDNR. These were used to make a cost surface, which was then used to create an optimal path from one point (a farmhouse) to another (the south bank of the North Branch of the Whitewater River). Different weighting schema were used to generate four "optimal" paths.

Problem Statement

The purpose of this exercise is to examine how changing the weight of a field in a multi-criteria decision analysis can change the outcome. The focus of this exercise was on using map algebra to make cost surfacse with different weights for their categories. Skills making API requests and ETL pipelines were combined with map algebra and optimal path analysis.

Table 1. Types of data used.

#	Requirement	Defined As	(Spatial) Data	Attribute Data	Dataset	Preparation
1	Elevation	.las data determining elevation from lidar	Point feature	Terrain elevation	MNDNR	Uploaded to MNDNR website as a .laz file, needs laszip to unzip
2	Land cover	.tif file with different categories of land use based on a colormap	Raster image	Categorized land cover	MNDNR via USGS	None

Input Data

Four vector .las files were downloaded from the MNDNR as .laz files, while raster .tif files were downloaded from MN Geospatial in a .zip file folder. A sample of the entire Whitewater River area was chosen due to the large size of .las data files. The .las data contained point feature data, but it needed to be unzipped and turned into a .las dataset before it could be used. The land cover data were .tif files with different attributes for the State of Minnesota. These were saved in a .zip file and needed to be extracted before they could be used. There were other supplemental files in the folder, but the only file used was the statewide land cover raster.

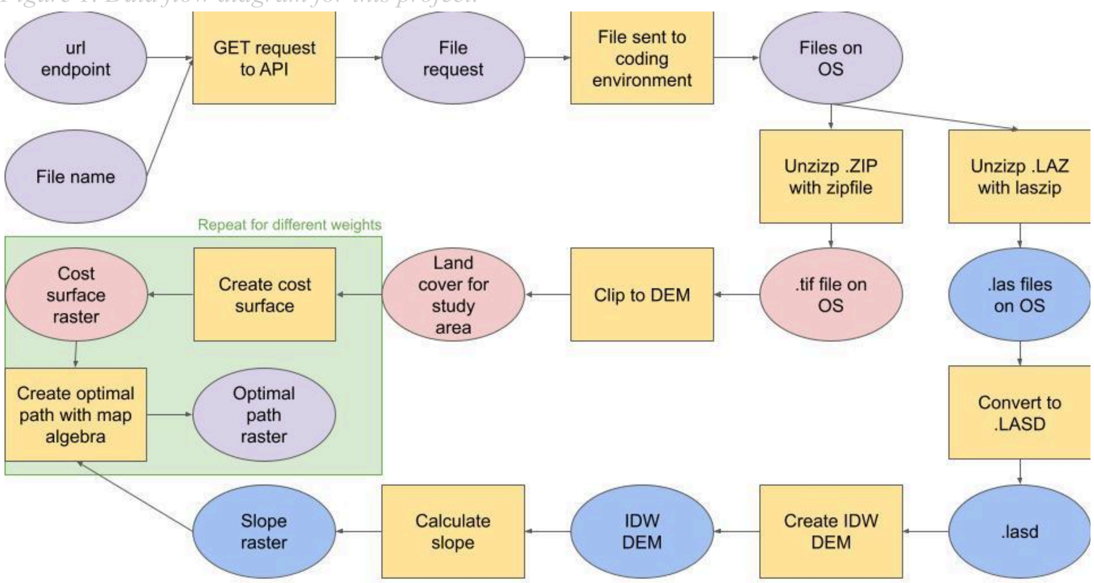
Table 2. Datasets used in this exercise.

#	Title	Purpose in Analysis	Link to Source
1	4342-28-60.las	LiDAR .LAS surface elevation data for part of Wabasha County, MN	MNDNR
2	4342-28-61.las	LiDAR .LAS surface elevation data for part of Wabasha County, MN	MNDNR
3	4342-29-60.las	LiDAR .LAS surface elevation data for part of Wabasha County, MN	MNDNR
4	4342-29-61.las	LiDAR .LAS surface elevation data for part of Wabasha County, MN	MNDNR
5	NCLD_2019_Land_Cover.tif	Colormap raster of land cover data for MN with an attribute table explaining the map categories	MN Geospatial Commons

Methods

The .las data from MNDNR was manipulated first using Python and the modules `arcpy`, `requests`, and `os`. To start, the data was accessed by using `requests.get()` on a url endpoint for the data and saved to a file folder in `os`. Next, the data was extracted using the DNR-provided laszip.exe. The unzipped .las files were added to ArcGIS Pro and turned into an LAS dataset such that it could be added to the map. This LASD was then converted into an IDW-interpolated DEM. This layer was used to make a slope raster. Next, the land cover data from MN Geospatial Commons was accessed by using `requests.get()` on a url endpoint for the data and saved as a zip file folder in `os`. The data was extracted using `ZipFile` and the relevant data was added to the map as a raster. This was then clipped using the DEM extent. Three different cost surfaces were created using this land cover data. Combined with the slope raster, four different optimal path analyses were conducted and saved as rasters.

Figure 1. Data flow diagram for this project.



Results
Even while keeping the same input rasters, the cost surface and optimal path analyses ended up being noticeably different. First, an optimal was created where all of the decision maker's criteria were equally weighted. Second, an optimal where avoiding water was devalued compared to avoiding fields, yet all other factors remained equal was created. Third, a cost surface where the cost surfaces values from the previous optimal path stayed the same, but land cover had double the weight was created, followed by one where slope had a greater weight than land cover. Finally, another criteria, avoiding roads, was given weight in the cost surface, creating a fifth optimal path. Some of the paths covered what visually appeared to be longer distances than others. There did not seem to be an optimal path where all of the criteria could be met; thus, to create the "true optimal path", the decision maker would need to weight the criteria herself. As the results stand, the decision maker will need to choose from the multiple different paths created programmatically with different criteria weights assigned by the mapmaker.

Results Verification
These results can be replicated using ArcGIS Pro's Jupyter Notebook integration. The large file size of the .las files required me to take a subsample, meaning that other analyses focusing on the entire Whitewater River area will be significantly different. To clarify, I am not creating "the" optimal path, but four paths optimized based on the criteria and weights I manipulated programmatically for my decision making analysis.

Discussion and Conclusion
This exercise proves that it would be impossible to create the optimal path without further input from the decision maker(s). Instead, I was able to create an optimal path, four of them to be exact, based on my own speculations. This class provided my first experience making cost surfaces and optimal paths. At first, I found this impossible because I didn't understand how I was supposed to identify water and farm fields from an .las elevation dataset. Then, a classmate told me I was meant to import a land cover raster as well. From there, the process was embarrassingly simple, given all the difficulty I had with lab 2, part 2. I conclude that this step would be a valuable thing to add to the lab directions, and that this exercise broadened my perspectives in a way I am grateful for.

Figure 2. Optimal path where all categories were equals placed over the land cover raster for easier navigability. The path is a grey line.

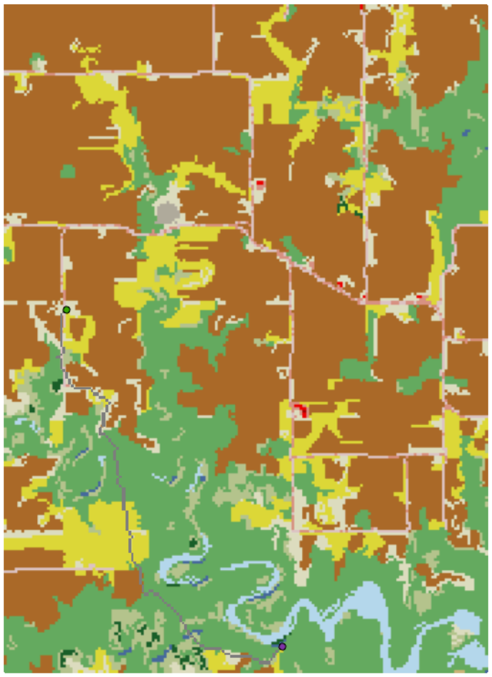


Figure 3. Second optimal path where avoiding water was devalued placed over the land cover raster for easier navigability. Again, the path is a grey line.

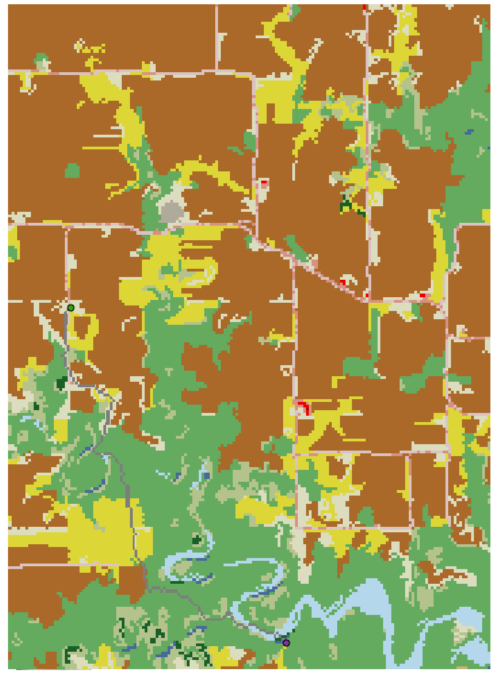


Figure 4. Optimal path where avoiding fields is given given high weight placed over the land cover raster.

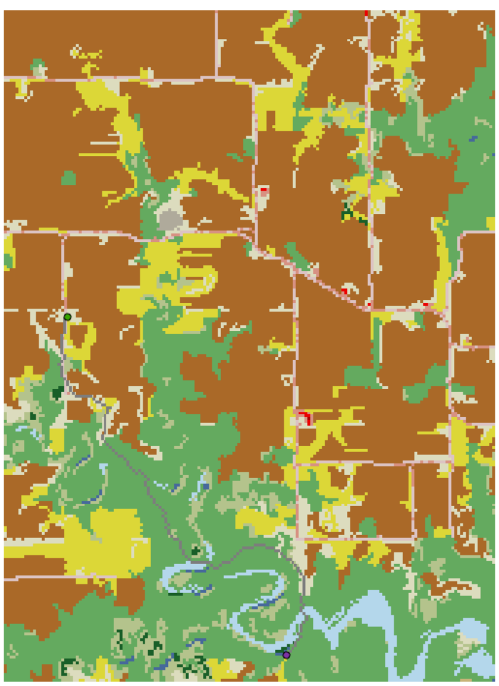


Figure 5. Fourth optimal raster where low slopes are high weight placed over the land cover raster.

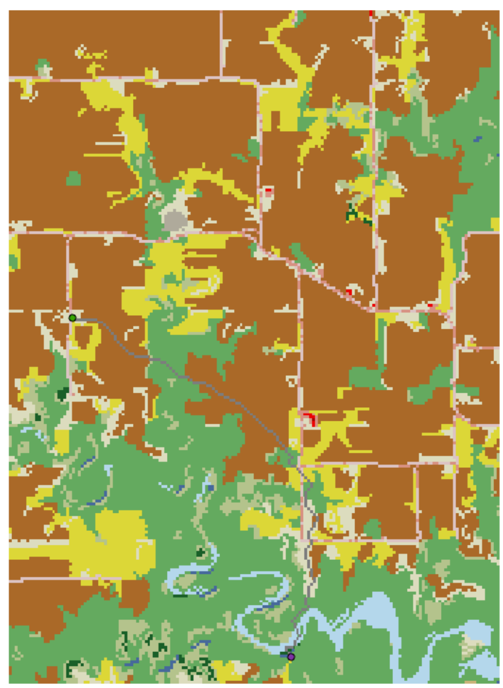
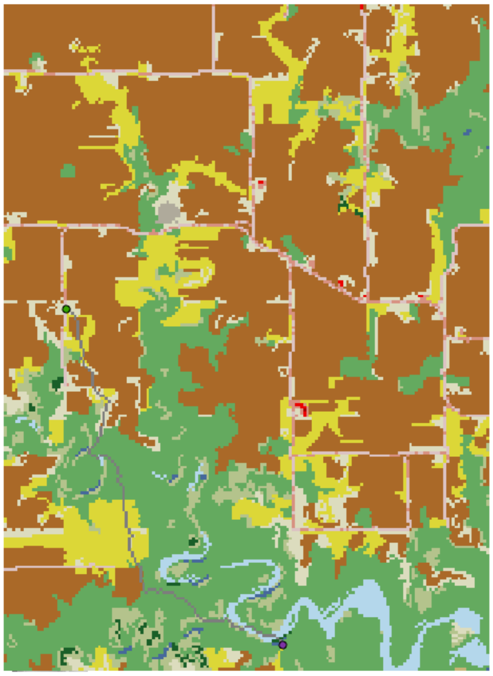


Figure 6. Optimal path where roads are an added category placed over the land cover raster.



References

Minnesota Department of Natural Resources. (2012). *LiDAR Elevation Data for Wabasha County*. Minnesota GIS Data Resources. https://resources.gisdata.mn.gov/pub/data/elevation/lidar/county/wabasha/laz/

Minnesota Department of Natural Resources. (2019). *NLCD 2019 Land Cover, Minnesota*. MN Geospatial Commons. https://prism.oregonstate.edu/normals/

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**