Lab Report

Title: Lab 3 Part 1 (Adjustment to Lab 2 Part 2)
Notice: Dr. Bryan Runck
Author: Eric Gibson
Date: 11/28/2023

Project Repository: https://github.com/gibso632/GIS5571/tree/main/Lab3
Google Drive Link:
Time Spent: 4 hours

Abstract

This project introduces creating a cost surface within GIS software, specifically ArcGIS Pro, using an automated script to solve an example word problem. The problem specified a person named Dora who needed to find the best path to a picnic area within Whitewater State Park. The parameters included avoiding any farmland, as it can be swampy during the Spring, avoiding areas which are too steep to traverse, and potentially avoiding rivers if she did not have her waders. It was eventually determined to establish any farmland as areas to avoid (higher value), rivers as medium-risk areas (medium value), and slope as an increasing risk (greater the slope angle, the higher the value). Utilizing a land cover raster dataset and a digital elevation model (DEM), five separate cost surfaces were made with differing importance between land cover types and slope. This allowed for Dora to determine which element (slope or land cover) was preferred over the other when traversing. Originally for Lab 2, only two ways of weighting the slope and land cover were explored, but for Lab 3 Part 1 an extra way was determined to satisfy the requirements for Lab 3. This extra form of weighting was exponential by creating a cost surface formed by one raster dataset calculated to the power of another. This weighted areas which were undesirable for Dora much more heavily than using addition when calculating rasters.

Problem Statement

The issue being studied within this lab is regarding a person named Dora who enjoys fly fishing at Whitewater State Park, specifically at the North Picnic Area within the park. Her farm is located at 44.127985, -92.148796 and she has certain parameters for travelling by foot to the picnic area. These parameters are 1) avoiding farmland due to the muddiness of the soil in the Spring, 2) avoiding water features unless she has her waders, and 3) avoiding areas which are too steep to traverse. Outside of these three parameters, the path she could take is relatively open and she does not have a stated preference on which parameter is most important. Since she has waders, however, it is likely water features may not have as much weight as farmland or slope.

Table 1. Deliverables for this lab.

#	Requirement	Defined As	(Spatial) Data	Attribute Data	Dataset	Preparation

1	Points illustrating both Dora’s farm and the North Picnic Area	Point feature classes containing points where Dora’s house and the Whitewater State Park North Picnic Area are located.	Location points	Name of location	Google Places API with your API Key	Use arcpy.Point and arcpy.PointGeometry to create coordinate points.
2	Cost surface rasters	Five separate cost surface rasters depicting how different parameters could be weighted by Dora.	Rasters in GeoTIFF format	Cost between farmland, water, and slope	GeoTIFF rasters	Create slope and land use rasters weighted based on Dora’s parameters, then weight those rasters differently to provide multiple cost surfaces

Input Data

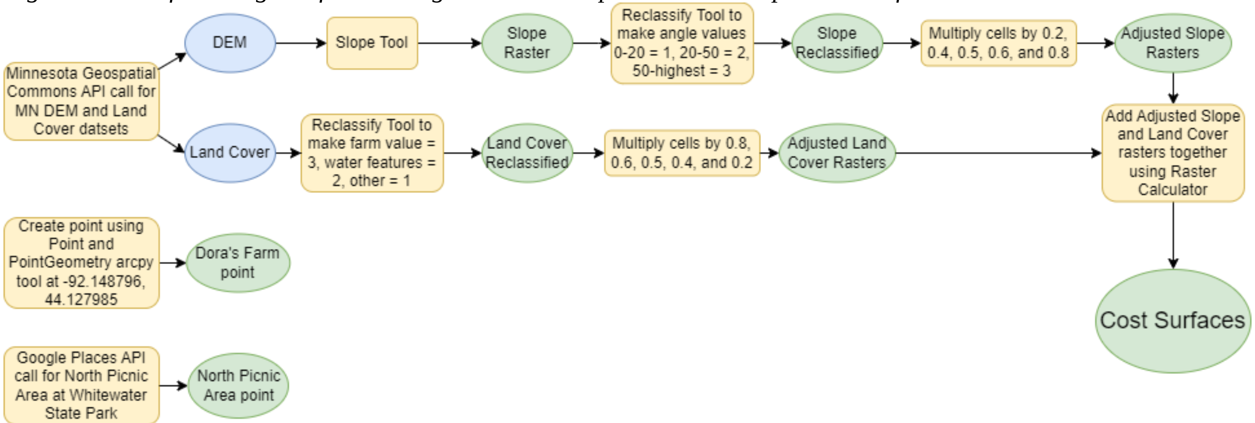
Three datasets are needed to create the final deliverables for this lab. This includes a GeoJSON gathered from the Google Places API containing the coordinates for the North Picnic Area in Whitewater State Park, a land use raster containing geographic data of farms within the area (in this case, the entirety of Minnesota), and a DEM of Minnesota. The land use raster will allow for the farmland and water features to be weighted based on the above parameters and the DEM can be converted to a slope raster depicting the angle of the slope with various weights established depending on the severity of the angle. The point data for Dora’s house will be created via the arcpy.Point and arcpy.PointGeometry tools.

Table 2. Datasets used within this lab.

#	Title	Purpose in Analysis	Link to Source
1	North Picnic Area at Whitewater State Park from Google Places API	Providing a destination for Dora determined by the cost analysis raster.	Google Places API – need API Key
2	NLCD 2019 Land Cover, Minnesota	For determining areas of farms and water features to find the best route for Dora depending on land use	Minnesota Geospatial Commons API
3	Minnesota Digital Elevation Model – 30 Meter Resolution	For creating a raster depicting the slope to find areas which may be too steep to traverse.	Minnesota Geospatial Commons API

Methods

Figure 1. Data flow diagram for creating start and end points and multiple cost surfaces



The data flow diagram above depicts both creating the cost surfaces as well as the points separately. Creating the North Picnic Area point also involved the arcpy.Point and arcpy.PointGeometry functions, but the coordinates instead originated within the GeoJSON provided by the Google Places API (44.05311337010728, -92.04608057989272). The DEM, as mentioned above, was utilized to create a slope raster which was reclassified so that an angle between 0 and 20 degrees was given a value of 1, an angle between 20 and 50 degrees was given a value of 2, and an angle between 50 and the maximum angle of 80.2473 was given a value of 3. The land cover raster was reclassified so that any farmland (land use labelled as "Hay/Pasture" or "Cultivated Crops") was given a value of 3, water features (land use labelled as "Open Water", "Woody Wetlands", and "Emergent Herbaceous Wetlands") were given a value of 2, and any other land use was given a value of 1. It is also worth noting the land use raster contained values depicting each land cover type. This workflow ultimately describes a larger value as less optimal for Dora.

Once these rasters with the weighted land use and slope values were created, they were both multiplied by 0.2, 0.4, 0.5, 0.6, and 0.8 to create five separate rasters with differing overall weights using a for loop which looped through a range of 1-4. The rasters were multiplied by 0.5 separately since the loop only increased by 0.2 and each number within the range was used to number the rasters. Each raster was then added to each other using the **Raster Calculator** tool in Arcpy inversely, meaning the reclassified slope raster multiplied by 0.2 was added to the reclassified land use raster multiplied by 0.8, then the slope raster multiplied by 0.4 was added to the land use raster multiplied by 0.6, and so on. This output five different cost surfaces with different weights provided for the slope and land use. The cost surface of the sum between the slope raster multiplied by 0.2 and the land use raster multiplied by 0.8 determined areas to traverse if avoiding certain land uses were most important and vice versa for the cost surface created by the sum of the land use raster multiplied by 0.2 and the slope raster multiplied by 0.8. The cost surface created via the sum of both rasters multiplied by 0.5 determined areas to traverse if both land use type and slope were equally important to avoid. The rasters multiplied by 0.4 and 0.6 depicted cost surfaces where either slope or land use type was slightly more important to avoid.

Finally, using the **Optimal Path As Line** tool through arcpy, a vector polyline was created which moved through the various rasters depending on the value of each cell (the lower the cell value, the better chance the line will move to). This optimal path started at Dora's farm and ended at the North Picnic Area in Whitewater State Park, however all the paths were different, as each cost surface weighted land cover and slopes differently. Originally, the optimal paths were missing from Lab 2, so they will be added here for the Lab 3 implementation. Additionally, for Lab 2 only five total cost surfaces were created using two types of weighting, but with the Lab 3 Part 1 update, another cost surface and optimal path will be created to compare the different weighting techniques. This final form of weighting was performed by utilizing the **Raster Calculator** tool to set one raster, in this case the land cover raster, to the power of another, in this case the slope raster. This weighted the cost surface heavily on which areas were undesirable for Dora between both slope and farmland. Meanwhile, areas that may have had one undesirable factor (farmland or slope) were less heavily weighted than taking the sum of the two rasters.

Results

The following figures are ArcGIS Pro screenshots of the rasters from land use being the most important to avoid to an extreme slope being the most important to avoid. Red colors depict areas to avoid, blue colors depict possible routes.

Figure 2. Cost surface with land use weighted as (multiplied by) 0.8 and slope weighted as 0.2.

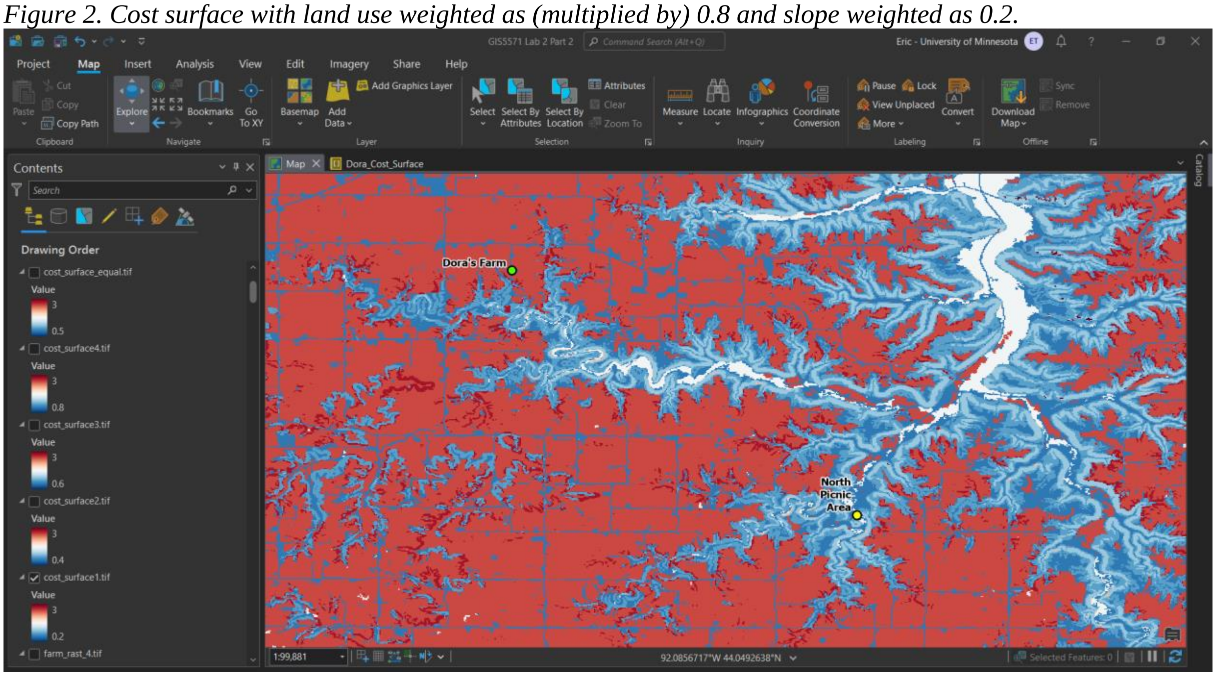


Figure 3. Cost surface with land use weighted as (multiplied by) 0.6 and slope weighted as 0.4.

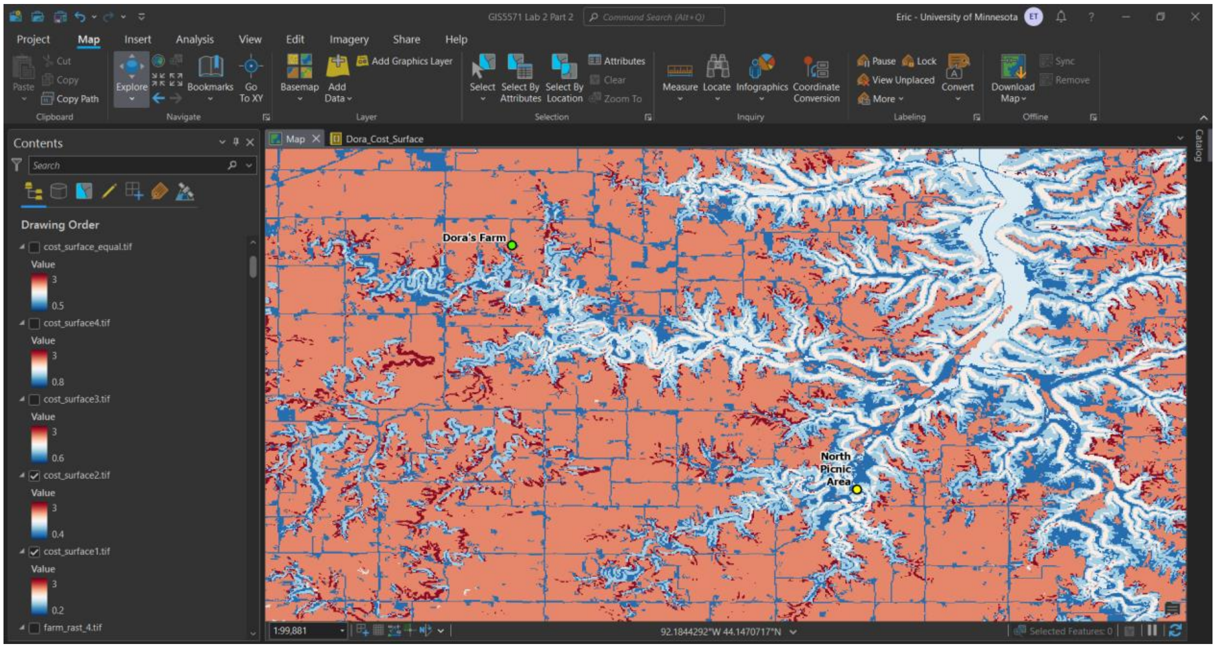


Figure 4. Cost surface with land use and slope weighted the same (multiplied by 0.5 each).

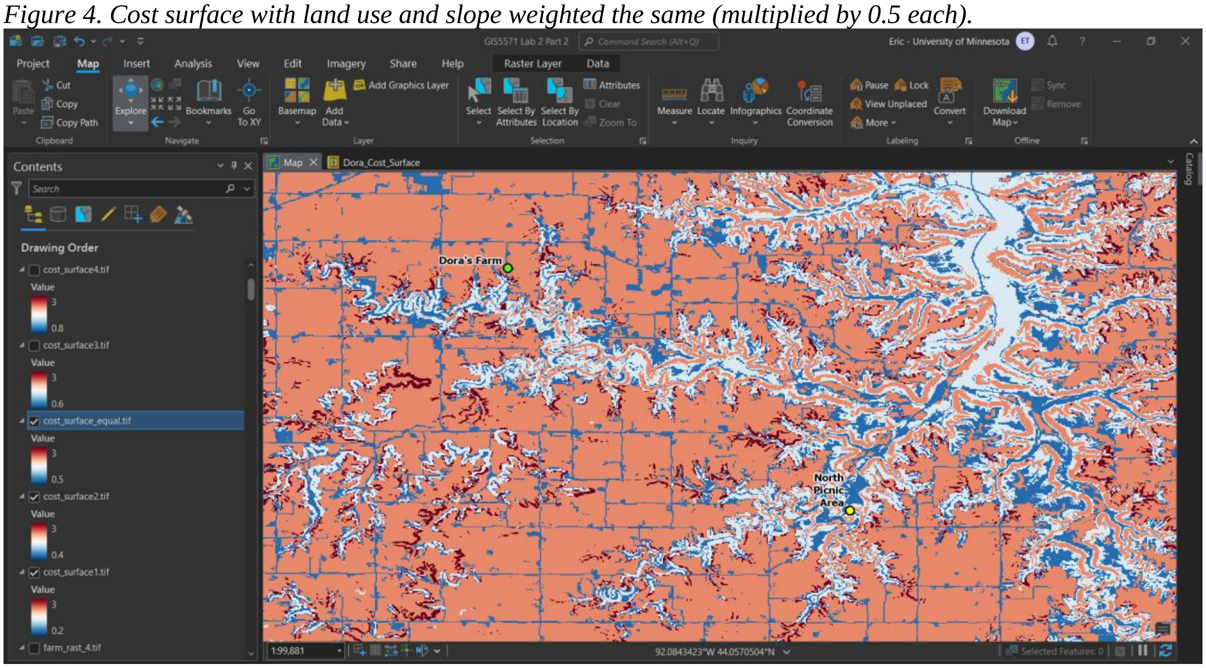


Figure 5. Cost surface with land use weighted as (multiplied by) 0.4 and slope weighted as 0.6.

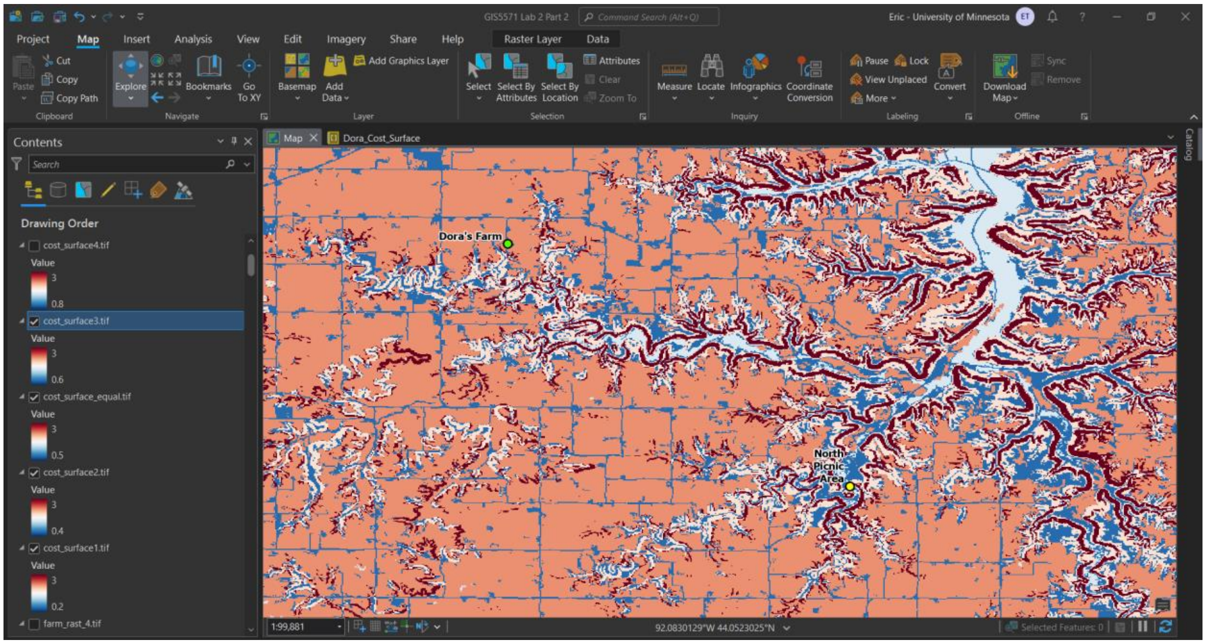
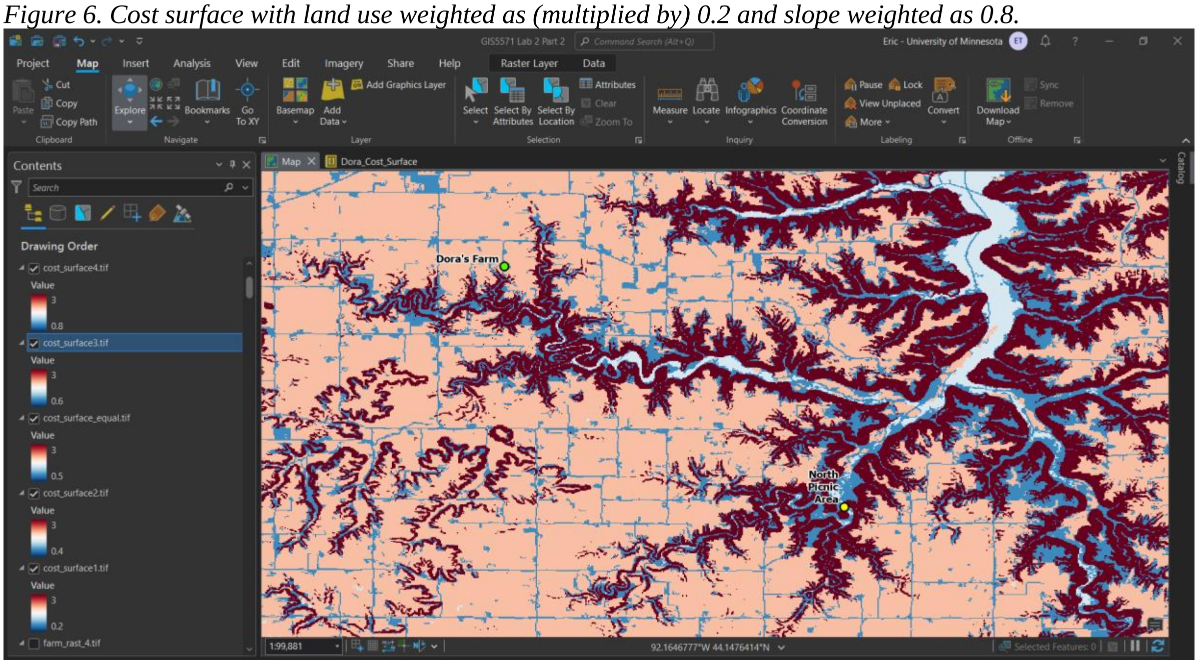


Figure 6. Cost surface with land use weighted as (multiplied by) 0.2 and slope weighted as 0.8.



These rasters feature a look at cost surfaces from where land uses are the most weighted for avoidance to where slope is the most weighted for avoidance comparatively. The point labelled "Dora's Farm" is Dora's starting point and the point labelled "North Picnic Area" is

Dora’s endpoint created via the arcpy.PointGeometry(arcpy.Point(*coordinates*)) function. The gradual change from farmland having the greatest avoidance value to slope having the greatest can be seen via the change in the location of the red coloration.

Lab 3 Part 1

For Lab 3 Part 1, a third form of weighting was added to compare three different ways of weighting for suitability analysis of paths for Dora. The third form of weighting was exponential where the cells within the land cover raster were calculated via using the exponent of the value of the relating cell in the slope raster and vice versa. The results of the optimal paths from all three forms of weighted can be viewed below:

Figure 7. Optimal path using Reclassify weighting

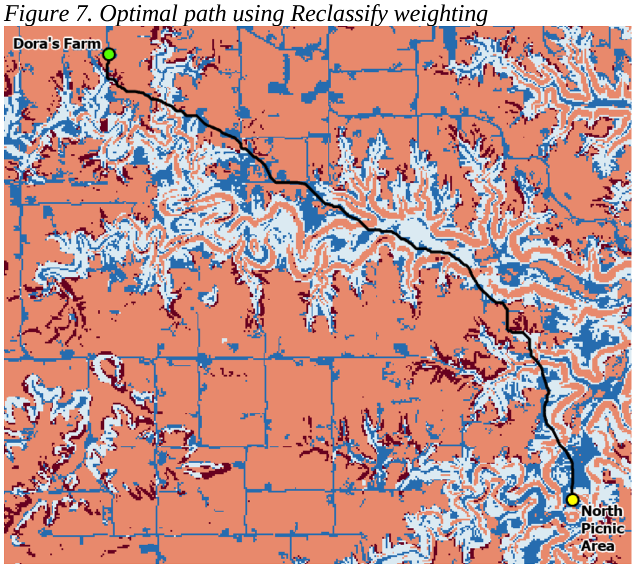


Figure 8. Optimal path when farmland is least desirable

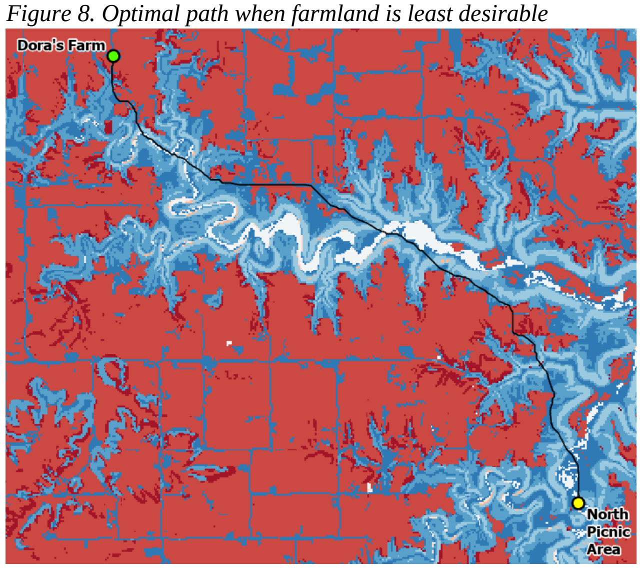


Figure 9. Optimal path when farmland is slightly less desirable

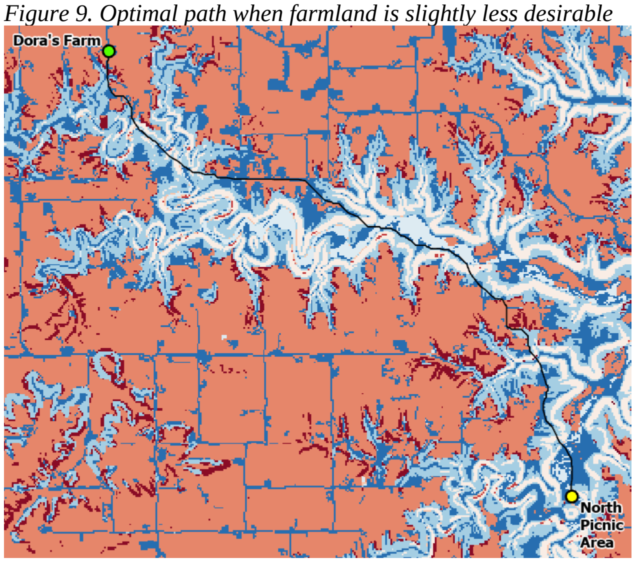


Figure 10. Optimal path when slope is slightly less desirable

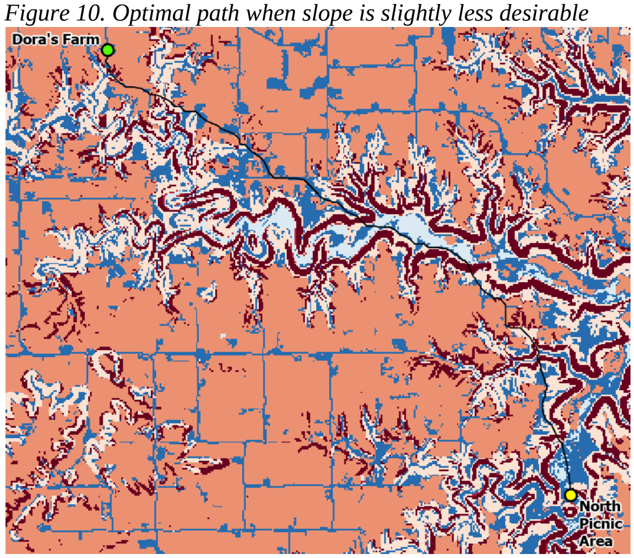


Figure 11. Optimal path when slope is least desirable

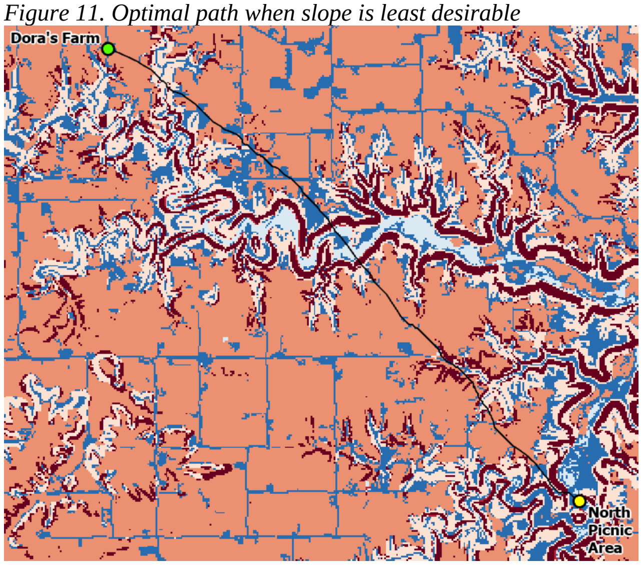
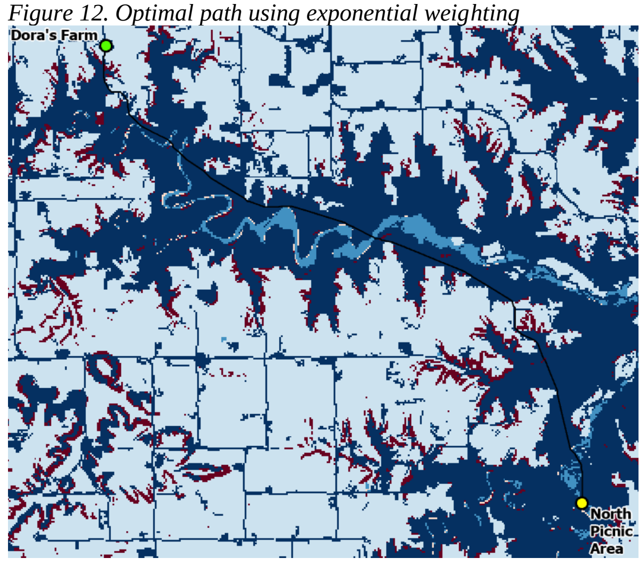


Figure 12. Optimal path using exponential weighting



The optimal path depicted in **Figure 1** was created using weighting based on reclassifying the land cover and DEM rasters of which the rest of the rasters were built off of. **Figures 8 – 11** contain optimal paths created via weighting by multiplying the rasters by 0.2, 0.4, 0.6, and 0.8 to find the paths if slope or land cover was most important to avoid (0.8) to least important to avoid (0.2) depending on Dora's preference. **Figure 12** depicts the optimal path by the final weighting added on through Lab 3 Part 1, which was exponential weighting where the land cover weighted raster was calculated to the power of the weighted slope raster.

Results Verification

The output needed to provide Dora with information regarding a path to take to the North Picnic Area at Whitewater State Park was relatively open. Outside of parameters mentioning avoiding farmland, potentially water features, and a steep slope, there was no determination of what should be weighted more: land use or slope. This led to the creation of several different cost surfaces with varying weights for both and the outputs in the **Results** section seem to replicate that. **Figure 2** depicts the cost surface where land use is the most weighted and **Figure 6** depicts the cost surface where slope is the most weighted with **Figures 3** to **5** gradually moving from land use being the most weighted to slope being the most weighted. When viewing these figures, there is a clear pattern showing this gradual change, confirming the validity of the cost surfaces.

Additionally, it is important to note there is also varying weights within the land use itself. Since Dora has waders, while she prefers to avoid water features, she can still traverse through them. This is why the water features were reclassified with a value of 2 instead of 3 and the cost surfaces do seem to validate this as well. When viewing **Figure 2**, while the land use is the most weighted, the water features only have a moderate avoidance value. This clearly describes how there is a difference between the weight of water features and farmland within the land use dataset.

The optimal paths, for the most part seem to align with the correct projections for what was hypothesized. The polylines tend to run along areas of less value from Dora's Farm to the North Picnic Area, but there are a few exceptions. **Figure 11** depicts a cost surface for if Dora believed slope to be least desirable, however the path seems to move directly through the highest value slopes. Presumably, this is because the reclassified rasters were used instead of the original slope rasters, meaning the values were simply not large enough for the optimal path to avoid these areas. Using the **Reclassify** tool in arcpy is useful for more discreet data, like land cover, but probably not continuous data, like slopes.

Finally, the final weighting for Lab 3 Part 1 was performed by calculating the exponent of one raster with another, leading areas containing farmland and a sharp slope to be much higher valued than one or the other. This is clearly demonstrated when visualizing the cost surface in **Figure 12** where the surface is mostly blue even in areas of farmland and sharp slopes, but a deep red in the few areas where farmland and a sharp slope coexist. As a comparison, all other cost surfaces were created by adding the slope and land cover datasets together. This exponential weighting formed an optimal path which is much more direct from the starting to end points than all others.

In terms of all three weighting types, using the **Reclassify** tool to form weights for different land covers and slope angle ranges was useful for moving forward with other weighting techniques, especially regarding land cover. As mentioned above, however, continuous data is not as useful within this weighting technique because the optimal paths may still move through areas of higher values. While it was necessary for the land use calculations, it was not necessary and even problematic for the slope calculations. The second form of weighting via multiplying both the land cover and slope by 0.2, 0.4, 0.6, and 0.8 each against each other worked well for determining which optimal paths to take if Dora preferred sharper slopes or farmland to traverse through. This helped through choosing what her preferences were but did not necessarily choose paths based on if both steep slope and farmland were present. For example, for the cost surface created via multiplying the land cover dataset by 0.2 and the slope dataset by 0.8 (**Figure 6**), since land cover has such a low value, it will not weight areas where steep slopes and land cover overlap by as much. While Dora prefers farmland in this case, she would still like to avoid farmland as much as possible, so this could potentially become an issue for this form of weighting. For this reason, the third form of weighting was performed, exponential weighting, which created a raster that weighted areas where slope and farmland was present heaviest, while other areas were lightly weighted, including areas of just slope or farmland. Still, this does output an optimal path that moves through areas of farmland and steep slopes, so this may not be the best form of weighting for Dora. Instead, it is likely a combination of all these types would be best.

Discussion and Conclusion

This lab was a great introduction to creating a cost surface and will help immensely moving forward, especially considering my Final Project will likely utilize a cost surface to create trails for high pointing. I learned there can be multiple weights within the same cost surface depending on circumstances and needs of individuals, companies, organizations, etc. This is important to understand and create cost surfaces with various parameters which may not be able to be enforced otherwise. This relates to the problem with this lab because Dora not only has parameters and preferences between land use and slope, she also has certain preferences regarding specific land use types. Land uses are all within the same dataset, at least for this study, so understanding how to provide weights for certain attributes is important to provide the correct deliverables.

Originally for Lab 2 I did not provide the optimal paths, so I have updated this lab with the optimal paths included as well introduced a third form of weighting to satisfy the requirements of Lab 3 Part 1. This final type of weighting allows Dora to see another path which weights mostly for areas where slope and farmland combine. Any area outside of that, including areas of farmland and a steep slope are not highly weighted as long as those features are not in the same area. This path is preferrable if Dora is fine with traveling through farmland or steep slopes, but not both at the same time. All three different types of weighting provide Dora with as many options as possible based on her preferences.

References

"Where Developers Learn, Share, & Build Careers." *Stack Overflow*, 2008, stackoverflow.com/. Accessed 31 Oct. 2023.

Multiple prompts. *ChatGPT*, October 2023 version, OpenAI, 31 October 2023, chat.openai.com.

Chu, Justin. "Overview." *Google*, developers.google.com/maps/documentation/places/web-service/overview. Accessed 09 Oct. 2023.

National Land Cover Database. "NLCD 2019 Land Cover, Minnesota." *Minnesota Geospatial Commons*, 2019, gisdata.mn.gov/dataset/biota-landcover-nlcd-mn-2019.

USGS. "Minnesota Digital Elevation Model - 30 Meter Resolution." *Minnesota Geospatial Commons*, 2004, gisdata.mn.gov/dataset/elev-30m-digital-elevation-model.

ESRI. "ArcPy: Documentation: Arcgis Developers." *Documentation | ArcGIS Developers*, 2023, developers.arcgis.com/documentation/arcgis-add-ins-and-automation/arcpy/.

ESRI. "ArcGIS pro Geoprocessing Tool Reference." *ArcGIS Pro Geoprocessing Tool Reference-ArcGIS Pro | Documentation*, 2023, pro.arcgis.com/en/pro-app/latest/tool-reference/main/arcgis-pro-tool-reference.htm.

Self-score

Fill out this rubric for yourself and include it in your lab report. The same rubric will be used to generate a grade in proportion to the points assigned in the syllabus to the assignment.

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	**20**

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	**24**
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	**16**
		100	**88**