Lab 3 Part One

Title: Create a Cost Surface with ArcGIS Jupyter Notebooks to Find Dory's Optimal Fishing Route (3 Routes)
Notice: Dr. Bryan Runck, Michael Felzan
Author: Alexander Danielson
Date: 11/30/2022

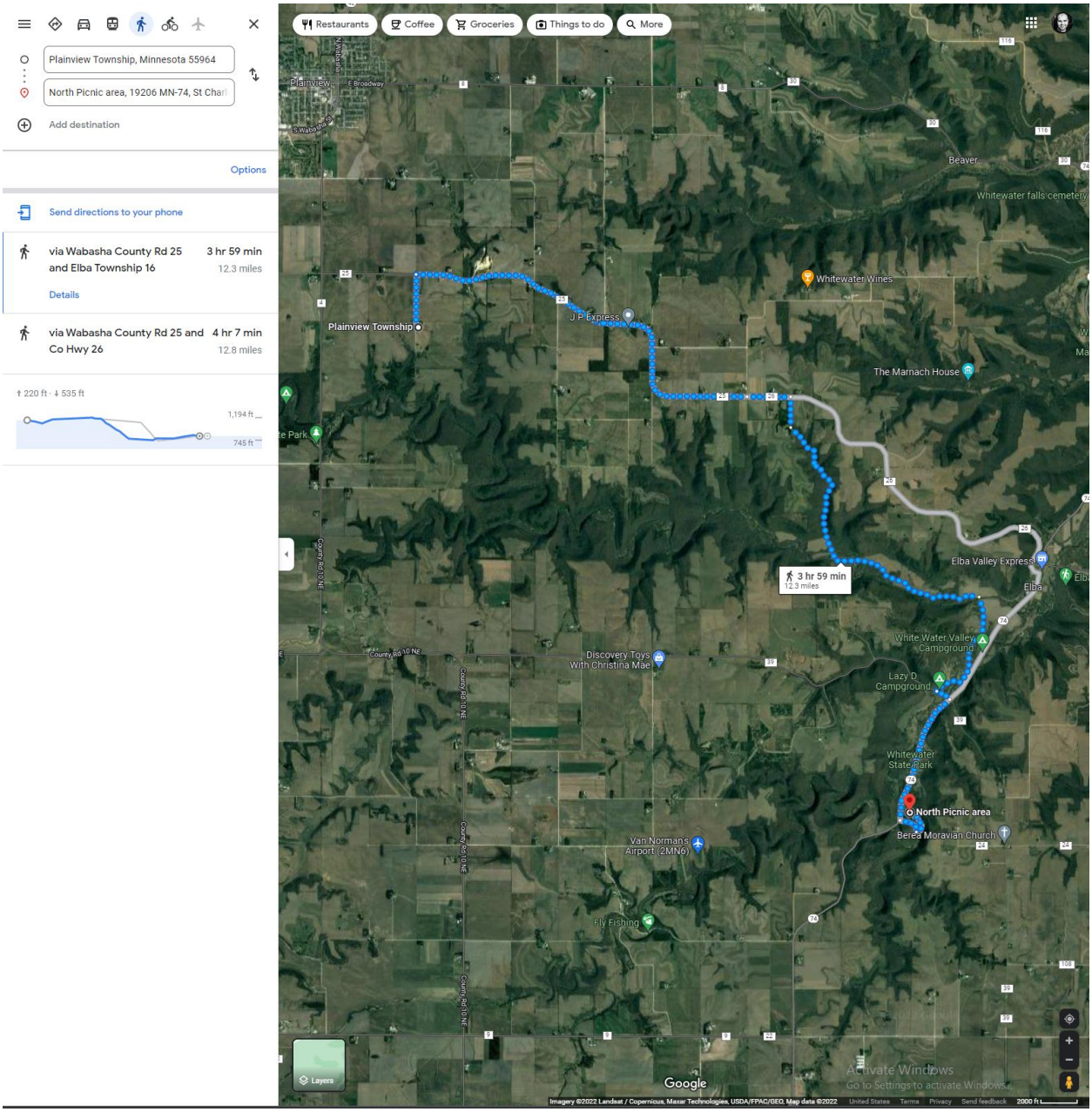
Project Repository: https://github.com/ardumn/GIS5571/tree/main/Lab3
Google Drive Link:
https://drive.google.com/drive/u/0/folders/11cDLxUWDv6QMdVnUu3WcG89pRrwmM9yB
Time Spent: 24 hours

Abstract

How do I find the best fishing location in Mary Lake (Itasca State Park) from my home in Lake Alice Township? Though it may have been solved, this hypothetical question can be applied to a cost surface analysis in finding the best cost-effective path and minimizing factors based on user preference. In the case of Dory, who is enamored with fly fishing, and moved to Plainsview Township solely for this purpose, but has bias for her preferred route to the Whitewater State Park. These preferences being she prefers the spring season to fish but doesn't like to attempt walking muddy farm fields or maneuvering water bodies (unless she has her waders) to that of a bridge. Slope is preferable depending on graduality, with the confines of the area she's leaving in. Given all these factors, the study extent is established by using three counties and ancillary data extracted from MN GeoSpatial Commons to mask, clip, and reclassify then make weights. These weights assess the cost surface based on each of the parameters (data) that have been calculated and then produce a cost-effective distance (which Dory's Farm is inputted). This then is assessed on cost surface from all sources and a back link for each neighbor (North Picnic Area destination). This then should result in the final optimal route and can be reassessed many times to find different paths and applies to many other applications. ***(Re-edit: For this next phase, we change the weights, classifications, and data inputs to find variations in routes that Dory can take.)***

Problem Statement

Dory's predicament is to find an optimal route for fly fishing in Whitewater State Park from her farm (44.127985, -92.148796) in Plainsview Township, MN to the North Picnic Area (44.0544° N, 92.0448° W). The preferences outlined for her are that she enjoys spring hiking but disapproves of walking in farm fields of muddy terrain and crossing into water bodies if there is not a bridge (unless she has her waders). Additionally, she wants to have a gradual path in terms of slope (incline). Using a cost surface model and analysis, her preferences are inputted to find her most optimal path within the study area and given terrain. (See Map 1)



Map 1. Google Maps clip of network analysis estimate for a general overview.

#	Requirement	Defined As	(Spatial) Data	Attribute Data	Dataset	Preparation
1	County Boundaries, Minnesota	The standard Minnesota State County Boundary dataset is used by MNDNR and many other state agencies. It is maintained by the	County Geometry	Vector Boundary	MN GeoSpatial Commons	Numerous preparatory operations and extractions. First, create an ETL, then

		MNDNR Lands and Minerals Division.				establish the study extent for which Dory's Farm and North Picnic area reside in two different counties (using attributes and dissolve tool) This being the extent for cursory operations for Cost Surface creation.
2	NLCD 2019 Land Cover, Minnesota	Minnesota NLCD layer is derived from the full NLCD dataset. The national raster is clipped and projected to UTM Zone 15N.	Raster Classes	Pixel	MN GeoSpatial Commons	Used as a weight for Cost Surface analysis in optimal route. Once Counties have been Dissolved, NLCD is used by Land Use and Impervious surface to implicate suitable paths(some being wetlands and water).
3	Minnesota Digital Elevation Model – 30-Meter Resolution	30 Meter Digital Elevation Model (DEM) is a copy of the USGS 1:24,000 scale Level 2 DEMs for the State.	Raster Elevation	Pixel	MN GeoSpatial Commons	Used as a weight for Cost Surface analysis in optimal route. Once Counties have been Dissolved, DEM is used for Slope for rigidity/uphill areas.

4	Stream Routes with Strahler Stream Order	Stream segments with Strahler stream order values assigned.	StreamLine Geometry	Vector	MN GeoSpatial Commons	Used as a weight for Cost Surface analysis in optimal route. Once Counties have been Dissolved, Strahler Streams are Clipped to study extent, Rasterized and used for Reclassificati on, and Raster Calculated to mitigate areas potentially wet.
5	XY Data.CSV		Aspatial	Tabular	N/A (Google Maps for coordinates)	XY Data CSV to convert XY Data to Point in ArcGIS Pro to display Dory's Farm and North Picnic Area

Table 1. Enumeration of the steps to establish the Cost Surface model about Dory's preferences.

Input Data

The data used is in line with the preferences and area that Dory's farm and Whitewater State Park reside in respectfully, all the county boundaries aren't necessary for this analysis, only three within the extent of the farmstead and state park. The counties act as the extent for all other layers to which the ancillary data will be masked and clipped, (after Olmstead, Wabasha, and Winona are exported). The NLCD 2019 data contains 12 classes ranging from Pasture, Bare Land, and Grassland to Developed land and is a subset of a larger NLCD dataset. Likewise, to that of the MN DEM originating from the USGS for elevation and rigidity of surfaces. Strahler Stream routes are unique, in that each stream is denoted with an ordinal ranking from 1 to 4 each number denoting the sallowness of the stream 1 being deepest and 4 being sallowest.

All data is inputted into a cost surface model and analysis and is suitable for the preferences outlined in the problem statement.

#	Title	Purpose in Analysis	Link to Source

1	County Boundaries, Minnesota	Study area by which the Optimal route for Dory will be investigated and the extent for inputting extracted masks for Landcover, DEMs, and Streams for Cost Surface operations.	MN GeoSpatial Commons
2	NLCD 2019 Land Cover, Minnesota	Classify Land Use features (Urban, Bare Land, Forest, Water, etc.) and Impervious Surfaces (Bridges and Roads) as a weight in the Cost Surface analysis for Dory's preferences for the North Picnic Area.	MN GeoSpatial Commons
3	Minnesota Digital Elevation Model – 30-Meter Resolution	Using the Slope tool to assess areas of uphill/downhill as a wight in the Cost Surface analysis for Dory's preferences to the North Picnic Area.	MN GeoSpatial Commons
4	Stream Routes with Strahler Stream Order	Rasterization of feature and Reclassification for further evidence and weighting for Cost Surface analysis for Dory's preferences to North Picnic Area.	MN GeoSpatial Commons

Table 2. Data requirements that apply to the study extent of Dory's Farm and the North Picnic Area (Whitewater State Park).

Methods

As described in the input data, the methodology used for the lab was to construct the study area by which the locales of Dory's farm and the Whitewater State Park reside in. By using the coordinates of Dory's farm and locating the coordinates of North Picnic Area via Google Maps, the data were tabulated into an Excel file and exported as CSV, for better functionality in ArcGIS Pro. Utilizing the XY to Point tool the points were displayed in the Map on ArcGIS Pro as a vector point in which the study area (as stated) was surmised. Once the three counties were selected, and dissolved by the points, with further extraction of the NLCD, DEM, and clipping of the streams, the raster analytics phase proceeds.

To this extent, the reclassification of each raster dataset is confirmed to Dory's preference stated in the problem statement, bridges and roads from impervious surfaces, land use of fields, barren land, water bodies, etc. slope for the DEM, and finally streams. Each is classified uniquely, especially that of Streams for NODATA values, since they have a surplus of these values, and need Raster Calculator to populate these values with 0s to have weights registered correctly. Then the Cost Surface can analyze multiple times from Dory's Farm and the Picnic Area based on the Distance to the nearest source (NPA) and each neighborhood and the least-accumulative cost for Back-Link. Thus, producing the optimal final route for Dory for her to find fly fish.

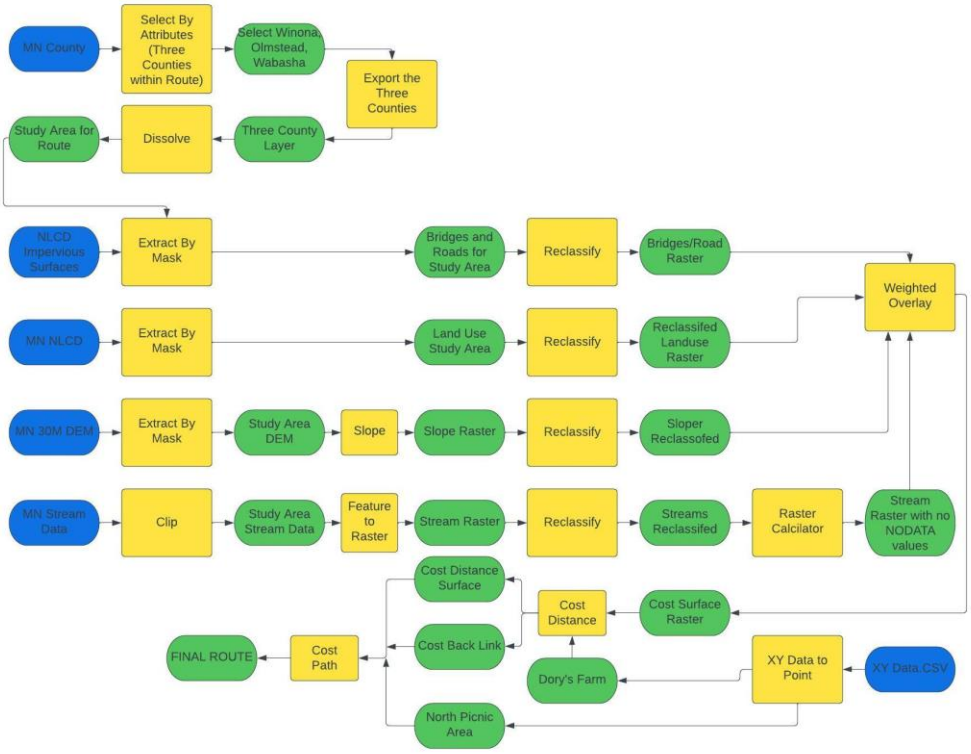
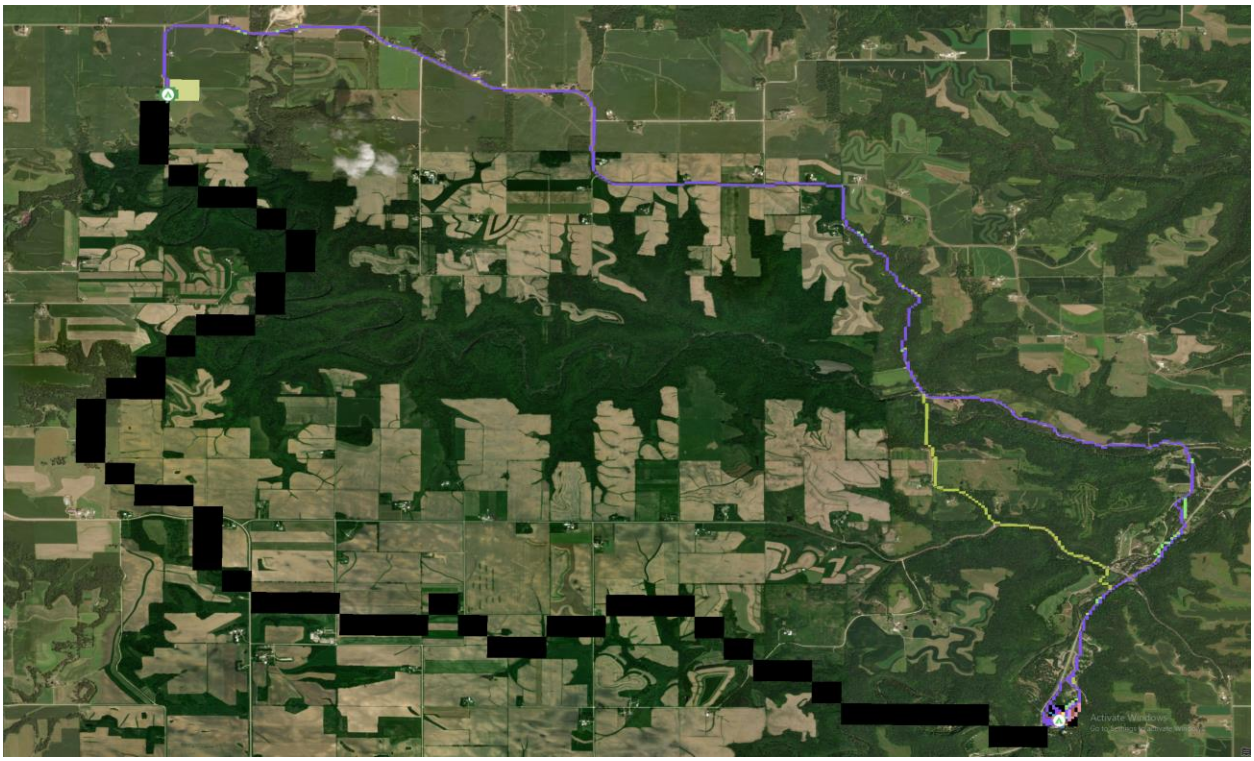


Figure 1. Enumeration of Cost Surface model steps and end result for a final route for Dory's Path to North Picnic Area.

Results

For this next phase, Map 2 now shows three different optimal paths from reclassifying and using different weights, the process for running the Cost Surface model proved resourceful and reliable, it is a matter of classification and enumeration units that are used. This time buffer operations were utilized around the river and origin/destination points so that it could be distinguishable and more defined compared to the original black path. Likewise, the number of weights that are used and run multiple times to find many potential routes, not just one fixed path is the point of a cost surface analysis and the predicates that are inputted into the model. It can assume that if you were to apply Cost Surface techniques to a similar procedure in finding the least cost path that using a plethora of other raster products and tools would implicate errors in data analysis and in the classification of said data to match the output of routes.



Map 2. Four Optimal Paths for Dory:1. Black Original Rough Path with One Weight for Each DEM and Slope 2. Purple Line More Defined Path with Higher Weighted DEM (4) and NLCD (2) 3. Darker Green with 20 Classes DEM and 20 Classes NLCD Both Equally Weighted 4. Lighter Green Both 10 Classes and Higher Weighted NLCD.

Results Verification

As depicted in the first map, there are four (roughly) paths that can be taken, with a slight bias in the classification and error in how each class was produced. The model itself is reliable and can be used to be reproduced for other purposes and analysis and can avail Dory in multiple ways which shed like travel for the North Picnic Area. But also applicable in slope and watershed analysis in this part of the state and others.

Discussion and Conclusion

Using Cost Surface analysis is almost parallel to that of a Network Analysis in finding the least cost path to a route, instead of vector analysis we are performing raster analysis and minimizing/mitigating preferences on a cell basis with a time travel basis. In juxtaposition, the cost surface model seems to take more factors into account than network analysis, or about the same amount of factors, for real-world implications if you have more weights and products to be inserted into your model. To minimize the amount of bias and error is to use the most reliable data sources, code correctly, have the same cell sizes as other classifications and weights, and run multiple analyses to find the best route based on the study you're training to perform

References

Create a network analysis layer. (n.d.). Arcgis.com. Retrieved October 30, 2022, from https://pro.arcgis.com/en/pro-app/latest/help/analysis/networks/new-network-analysis-layer.htm

Conditional evaluation with Con—ArcMap. (n.d.). Arcgis.com. Retrieved October 29, 2022, from https://desktop.arcgis.com/en/arcmap/latest/tools/spatial-analyst-toolbox/conditional-evaluation-with-con.htm

Understanding cost distance analysis. (n.d.). Arcgis.com. Retrieved October 29, 2022, from https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/understanding-cost-distance-analysis.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	**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	**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	**98**