Lab Report 03 Part 01

Title: Cost Surface Model, 3 Options for Dory's Hike from Home to Whitewater State Park
Notice: Dr. Bryan Runck
Author: Jake Ford
Date: 11/29/2022

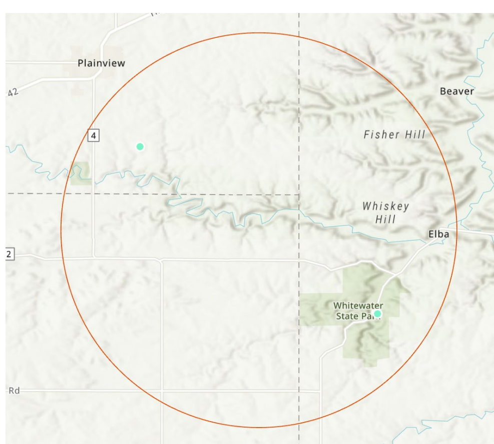
Project Repository: https://github.com/ThisFord/GIS5571-arc1.git
Time Spent:*30+*

Abstract

In this project we find the best route for Dory to take on her journey from her house to her favorite fishing spot in Whitewater State Park. We use raster analytics to create a Cost Path that follows the least cost path from one location to another based on user input. These decision maker criteria form the major constraints of our cost model. We make an ETL Pipeline to gather data from online sources, focus that data to our area of interest, standardize the data based on the constraints outlined by Dory, and combine the multiple raster sets into a weighted overlay which is then used in calculating the final cost path. Using weights that prioritize a path with minimal elevation change and account for landcover and hydrological features we calculate the best path for Dory to take. Because of the model's preference for shortest distance, the path ends up being a relatively straight line from start to finish, hewing along the borders of some steep slopes and farm fields. The resulting path represents the best path based on Decision Maker Criteria and is ready for scenario testing and adjustments and iteration based on user feedback. This allows for the path to be continually optimized, and for Dory to provide more precise metrics to shape the model toward a perfect fit.

Problem Statement

Dory wants to find the best route to her favorite fishing spot. She wants avoid farm fields and water that is too deep for wading; she prefers using bridges to cross streams and prefers the path with the most gradual slope. Using Map Algebra and raster analysis tools in ArcGIS Pro, we create a program to automatically determine the best path based on Dory's preferences. (*Creating a Cost Surface Raster—ArcGIS Pro | Documentation*, n.d.)



The area of interest for Dory's journey, the two dots represent the start and end points, while the red circle is the boundary of the study area.

Table 1. Requirements

#	Requirement	Defined As	Prep
1	ArcGIS Pro	Software for geospatial processing	
2	Jupyter Notebook in ArcPro	Python programming interface in Esri's ArcPro software	
3	Lucid Chart	Model sketching online software	Sketch out workflows
4	Model Builder in Arc Pro	For building Geoprocessing workflows in Arc Pro and visualizing the process	Create and execute workflows

Input Data

To calculate the best path we need to establish a area of interest around the defined start and end points. We begin by defining those points, then using a 2km buffer we establish a minimum bounding area between the points to account for possible back tracking. The county boundaries are used to filter the other input datasets prior to download to streamline the process. Elevation data is sourced from the Minnesota Geocommons and used to produce a slope model to help find the most even path. Hydrologic data comes from the same source and is used to put constraints on the path based on water feature size and class. Road data is used as an erase feature for the hydrological data, simulating bridge crossings in our weighted surface. Landcover data sourced from the MNDNR provides further constraints, allowing us to identify farm fields which Dory does not wish to traverse. The combination of these data sources into a single weighted cost surface comprises the bulk of the data processing necessary to construct the two inputs for a cost path analysis; a cost distance surface and a cost back link surface.

Table 2. Data for analysis

#	Title	Purpose in Analysis	Link to Source
1	Minnesota Roads	Defines constraint features	Mn Geospatial Commons
2	Table of Dory's coordinates: start and end points	Establishes start and end of route and defines AOI	Manual entry from brief
3	County Boundaries	Preprocessing ETL data	link
4	Elevation	Calculate slope of path	link
5	Hydrologic data	Defines constraint features	link
6	Landcover	Defines constraint features	link

Methods

This project will compile multiple feature layers and raster data sets into standardized weighted surfaces based on Dory's user preferences to find the optimal path of travel between her home and destination at Whitewater State Park. (Figure 1) describes the following workflow. We begin by wrangling data that pertains to the area of interest (aoi), make them into a comparable raster datatype, then we standardize the raster data through a reclassification process based on a scale of 1-10, using the user preferences as a guide. (*Multicriteria Decision Analysis in Geographic Information Science*, n.d.) After each data set has the raster cell values reclassified to the same scale based on our interpretation of Dory's preferences we combine all of the standardized surfaces into a single weighted surface where each input surface has a different percent impact of the final cell value. (Figure 2) This surface is used to create a cost distance raster and a back link raster, (Figure 3) which computes the cost of moving from one cell to another based on neighboring values. (*Creating a Cost Surface Raster—ArcGIS Pro | Documentation*, n.d.) These two surfaces form the final input to our cost path analysis. (*Using Weighted Overlay Analysis to Identify Areas That Are Natural and Accessible*, n.d.) In our initial cost path analysis the input surfaces are weighted as follows, generating the cost path in the figures below: landcover 50%, hydrological features 25%, slope 25%.

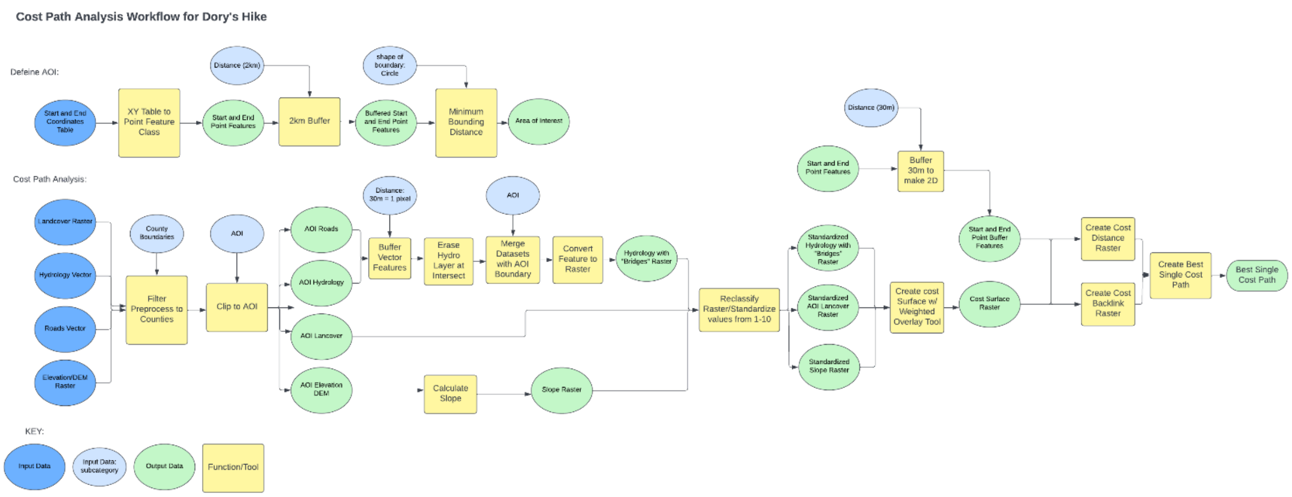


Figure 1, the project workflow.

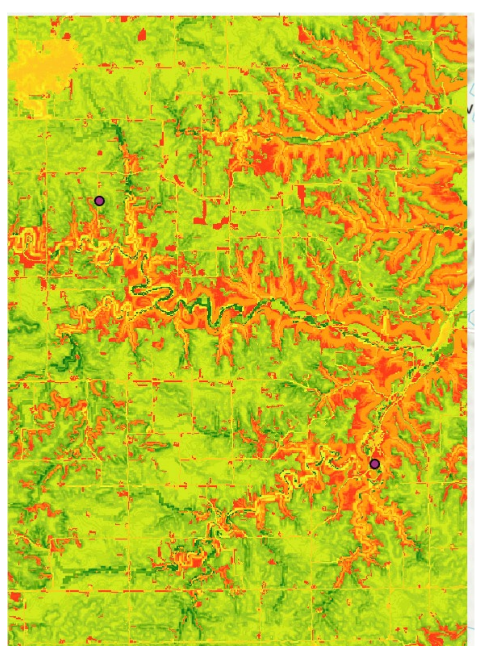


Figure 2, the final first cost surface, green represents lower cost, red higher.

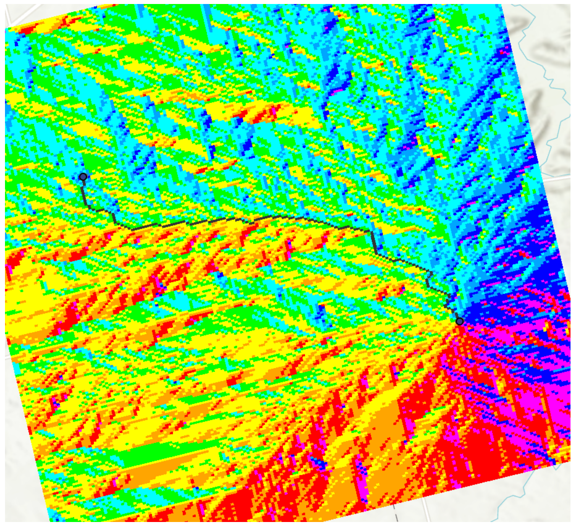


Figure 3, cost backlink surface with initial optimal path in black.

Results

Initial results of the weighted surface with land cover weighted at 50%, hydrological features 25%, slope 25% (Figure 4A.) This represents the optimal path for Dory with minimal distance and minimal elevation change (i.e. slope) as the highest priority

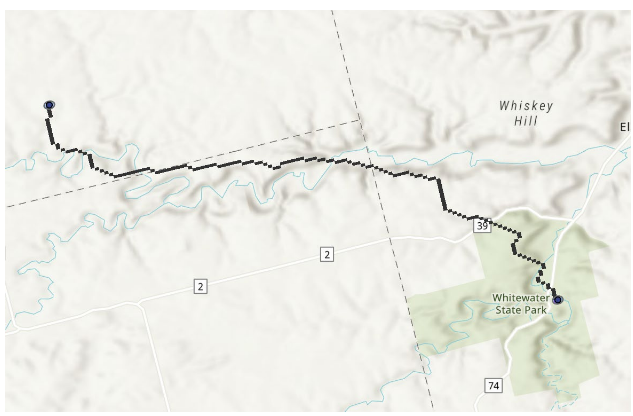


Figure 4 showing the first route over the cost surface compiled with the above weights.

Looking at the topographic and hydrological feature layers, (Figures 5 and 6) she crosses some streams and gullies that are weighted heavily. This appears somewhat counter intuitive until we consider that the final weighted layer has set her preference to minimize distance as highest priority. The next two scenarios adjust the weights to minimize water crossings, which would likely change the optimal path. Further feedback from the user is required to determine the fit of the optimal path model to the desired outcome. The automation of the script makes adjustments and iteration a low-cost feature of the decision optimization process. (*Creating a Cost Surface Raster—ArcGIS Pro | Documentation*, n.d.)

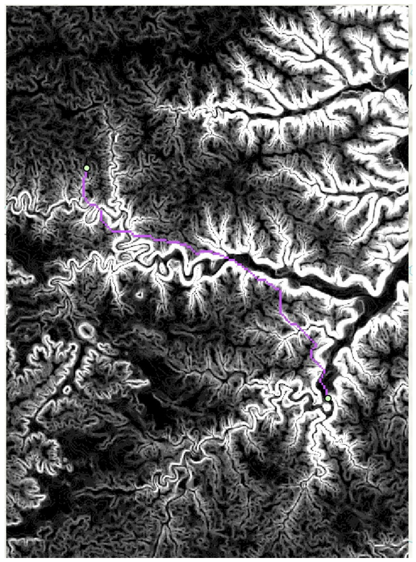


Figure 5 shows the cost path over the slope derived weighted surface, lightest areas have the greatest slope.

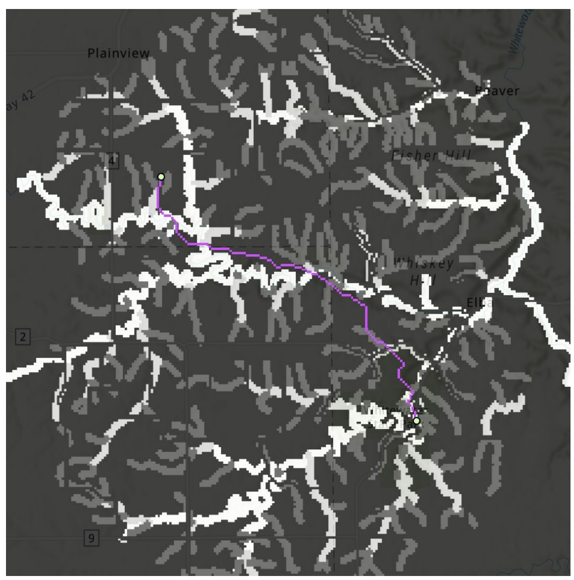


Figure 6 shows cost path over the hydrology weighted surface.

The first cost path does a satisfactory job of considering landcover classes, hewing closely to the wooded regions near the stream and avoiding farmland. The second and third paths created with alternate weights both weighted landcover quite low, and slope very high, as a result the paths cut across farm fields to keep a flat path. (Figure 7) The final figure in this section shows the three paths together with varying weights on an Esri basemap for clarity. (Figure 8)

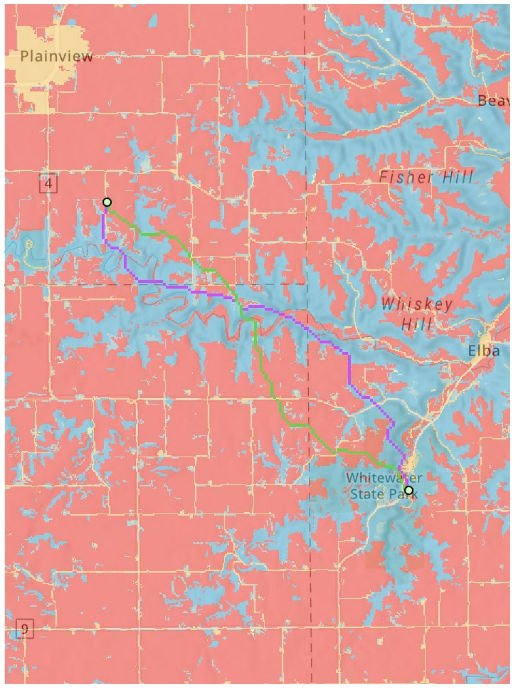


Figure 7, first cost path in purple, and second in green with landcover map.

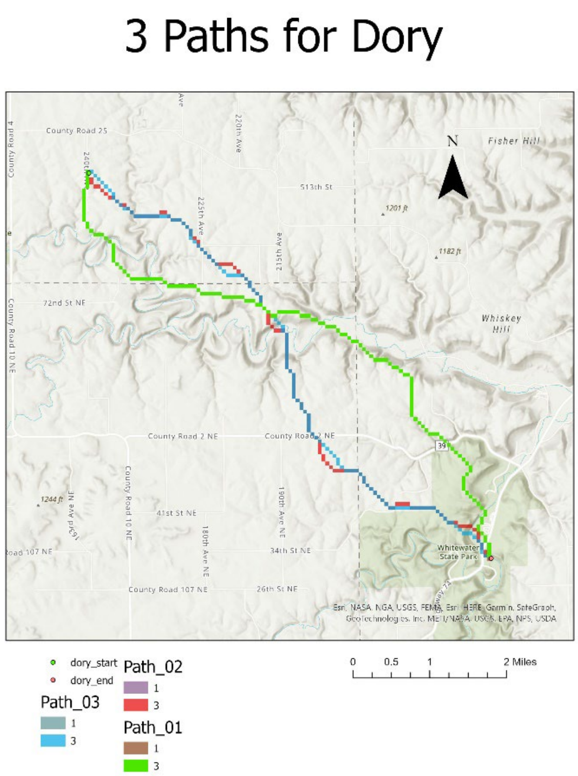


Figure 8, the three cost paths .

Cost path 3 and 2 share much of their geometry, even with very different weights, the effect of the slope on the path is apparent. Path one, the initial path weighted landcover ass the biggest constraint, so it hews to the wooded, non-agricultural lands near the stream, and as a result has a much higher slope total over the path.

Results Verification

Because the best path is ultimately determined by the decision maker’s preference, the validation process will be necessarily qualitative. In order to assess the accuracy of our model we present Dory with the determined best fit path and the geographic features underneath, along with an analysis of the weights assigned to each criterion as justification for the route. We can adjust weights based on the user’s appraisal of the fitness of the path based on the map features, and do test runs on the path in the real world. Ideally, we would present 3 alternative paths based on different weight configurations to test against one another, then use the decision-maker feedback to adjust weights in the model. Using a scenario type analysis model fits this process well and can be achieved at little cost or inconvenience, as most days Dory makes this trip. This will allow us to home in, iteratively, on a best fit path for the user. (*Modeling Biological Systems*, n.d.; *Multicriteria Decision Analysis in Geographic Information Science*, n.d.)

Discussion and Conclusion

Creating a cost path model is a complex task. There are many ways to construct the model, with multiple features and weights to be considered. A major result from this project is that it demonstrates that the problem and decision criteria need to be explicitly and comprehensively defined in order to have an accurate, useful and direct model for path optimization. In this case, success is hard to determine because the constraints are hard to define concretely. As it is, the path described above fits the brief, but may not be the absolute best fit, as the fit has not been defined in absolute terms.

Currently with much more work, the paths have been fixed to display correctly and account for all the proper weighted layers. The three paths show great options, 2 very different paths and 2 quite similar but with extra care taken to find the flattest path.

References

Creating a cost surface raster—ArcGIS Pro | Documentation. (n.d.). Retrieved November 6, 2022, from https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/creating-a-cost-surface-raster.htm

Modeling Biological Systems: (n.d.). Retrieved November 25, 2022, from https://link-springer-com.ezp1.lib.umn.edu/book/10.1007/b106568

Multicriteria Decision Analysis in Geographic Information Science. (n.d.). Retrieved November 25, 2022, from https://link-springer-com.ezp1.lib.umn.edu/book/10.1007/978-3-540-74757-4

Using weighted overlay analysis to identify areas that are natural and accessible. (n.d.). ArcGIS API for Python. Retrieved November 6, 2022, from https://developers.arcgis.com/python/samples/calculating-cost-surfaces-using-weighted-overlay-analysis/

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**