

## Lab Report

Title: Lab 2

Notice: Dr. Bryan Runck

Author: Cole Anderson

Date: 3/4/21

**Project Repository:** <https://github.com/and04671/GIS5572/tree/main/Lab2>

### Abstract

The purpose of this exercise was to find an optimal path, based on specified criteria, by creating a cost surface from an ETL data pipeline. The goal was to find a shortest path from Dory's home to a location in a nearby state park that did not cross water or farmfields and was as flat as possible. To accomplish this, a DEM from USGS and Crop Type layer from NASS/USDA were reclassified and combined. A distance accumulation was calculated from the source point, which was then used for an optimal path function. The results show a path that follows the criteria for the most part, with data resolution causing some water cross error. The outputs of this exercise show that suitability analysis is heavily dependent on operational definitions of criteria and data equality. The skill learned in the lab will be helpful in a final project regarding accessible movement.

### Problem Statement

Create a cost surface and find an optimal from Dory's farm (44.127985, -92.148796) to North Picnic Area in Whitewater State Park with following criteria: does not cross farm fields does not cross water bodies without bridges, and has minimal slope.

Table 1. Problem Statement Requirements

#	Requirement	Defined As	Spatial Data	Attribute Data	Dataset	Preparation
1	Path does not cross farmland	Farmland = 100/Not = 1	Local raster of land area	Land type	<a href="https://gisdata.mn.gov/dataset/agric-cropland-data-layer-2018">https://gisdata.mn.gov/dataset/agric-cropland-data-layer-2018</a>	
2	Path does not cross water bodies	Water = 100/Not = 1	Local raster of area	Land type	<a href="https://gisdata.mn.gov/dataset/agric-cropland-data-layer-2018">https://gisdata.mn.gov/dataset/agric-cropland-data-layer-2018</a>	
3	Path uses minimal slope surface	Min(Slope) = 1 Max(Slope) = 100	Slope	elevation	<a href="https://gisdata.mn.gov/dataset/elev-30m-digital-elevation-model">https://gisdata.mn.gov/dataset/elev-30m-digital-elevation-model</a>	Convert DEM to slope
4	Map distance accumulation/cost surface	Arcpy.sa.DistanceAccumulation				#1 raster + #2 raster
5	Map the optimal path from farm to picnic area	Source point [ ] to Dest point[ ] with min(Cost)				

## Input Data

The Minnesota DEM Model 30 M is a file from USGS, published by the MN DNR. It shows the elevation of Minnesota in NAD83 UTM-15 projection. The field 'VALUE' gives a number of feet above sealevel for each cell, while 'COUNT' gives the number of cells at said height. The 2018 Cropland Data Layer is a file from the National Agricultural Statistics Service and USDA, published by the Minnesota Department of Agriculture. It shows land type values in Minnesota in NAD83 UTM-15 projection. The field 'VALUE' gives the land cover type value (including water), field 'CLASS\_NAME' is the land cover type name, and field 'COUNT' the number of cells with specified land cover type.

Table 2. Input Data

#	Title	Purpose in Analysis	Link to Source
1	Minnesota Digital Elevation Model – 30 Meter Resolution	Finding lowest slope path	<a href="https://gisdata.mn.gov/dataset/elev-30m-digital-elevation-model">https://gisdata.mn.gov/dataset/elev-30m-digital-elevation-model</a>
2	Cropland Data Layer 2018, Minnesota	Path that does not cross water/farms	<a href="https://gisdata.mn.gov/dataset/agri-cropland-data-layer-2018">https://gisdata.mn.gov/dataset/agri-cropland-data-layer-2018</a>

## Methods

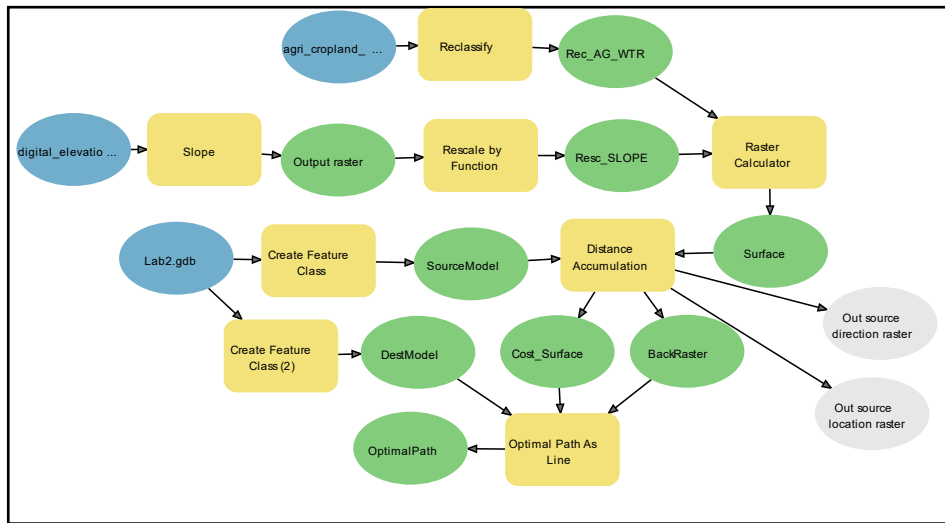
Function CKAN\_retrieval gathers required data from MN Geospatial Commons. The function takes parameters for a search query, result number, and resource number. A requests.GET request is sent to the CKAN API, containing the search query. The returned dictionary is converted to a JSON object, which can be parsed. The possible result and then possible resource numbers, as specified in the function parameters, find the correct dataset. This returns another URL with the particular dataset's address. Another GET request is sent to this new address to download the resource. The retrieved file is saved to disk as a .zip, where it is then extracted using the zipfile library module. The function is repeated for each desired dataset.

The cropland and DEM data must be standardized/reclassified to a common scale before creating a cost surface. In the cropland layer, raster cell values representing water or crop areas are converted to 100, and all others to 1, using the `arcpy.ddd.Reclassify` function. Desired land types are given the lowest possible cost and undesired the highest possible. In the DEM layer, the DEM is first converted to slope degrees using the `arcpy.ddd.Slope` function. The slope raster is then rescaled via `arcpy.sa.RescaleByFunction` with larger slopes corresponding to higher values between 1 and 100. The problem statement desires the most gradient slope, thus lower slopes are given the least travel cost. These two component rasters are then added together to create a weighted cost surface. To decrease processing time, the new surface raster is clipped to the rectangular extent (560000, 4850000), (600000, 4900000).

The source and destination points are created by creating two new feature classes and adding the respective source or destination point feature to each. The source point is created at (568097.73, 4886440.22) and the destination at (576512.44, 4878357.25). These are UTM coordinates, and were translated from the given degree coordinates. To add the points, an `InsertCursor` places a new row in each feature class.

The distance accumulation function `arcpy.sa.DistanceAccumulation` takes the source point and the clipped cost surface to find the cost to reach each raster cell from the source. The function outputs a new raster surface and a back-direction raster for use in the next step. The destination point, distance accumulation surface, and backraster are all used as inputs to find the final optimal path via function `arcpy.sa.OptimalPathAsLine`

Figure 2. Model Builder Diagram



## Results

In Figure 3, 4 and 5, Dory's farm is on the left-hand side of each image. The fly-fishing spot is in the bottom right corner. Figure 3. shows just the slope raster relative to the path. It is clear that the path runs down largely green, or low slope regions. The path enters a river valley about halfway along the trip; this area is very low slope, though it is close to water.

Figure 3. Optimal Path Output with Slope Raster

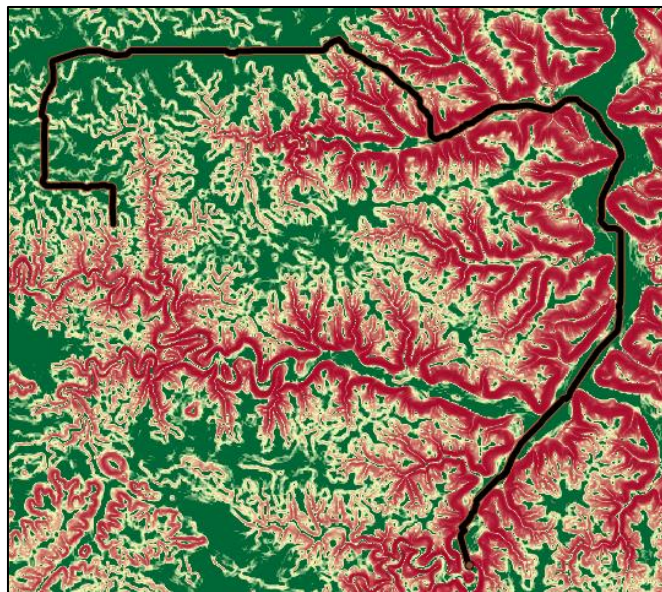


Figure 4. shows the final reweighted cropland raster, where purple areas are acceptable, and green ones are not. The path enters a solid purple block in the second half.

Figure 4. Optimal Path Output with Cropland Raster

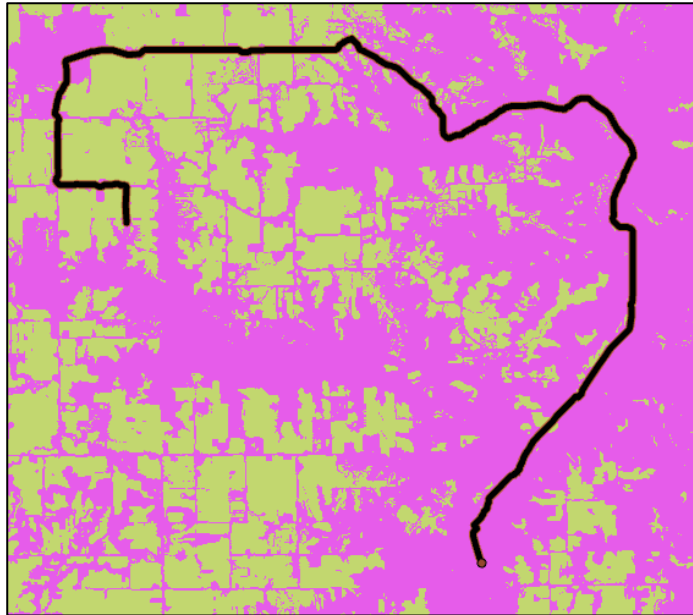
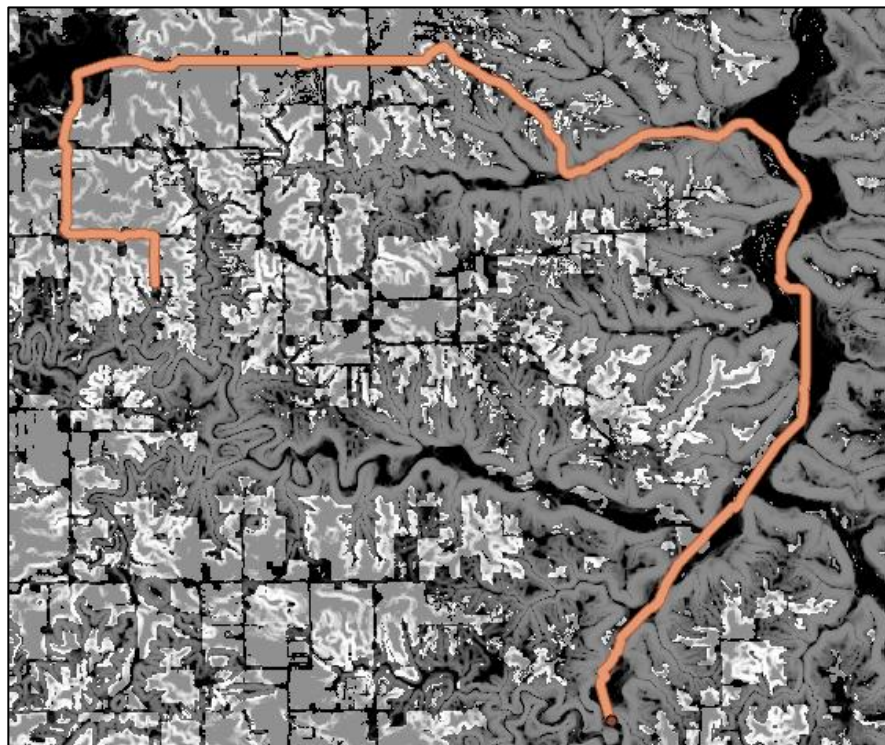


Figure 5. shows the final cost raster. Areas like roads and plateaus end up being desirable areas, while the bluffs on either side of the river basin are the least so. Roads are often narrow corridors between farm fields in the map.

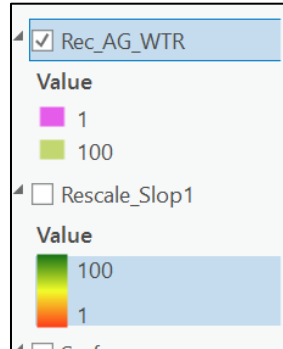
Figure 5. Optimal Path Output with Cost Surface



## Results Verification

The optimal path is not perfect. The raster granularity for land type must not have been small enough for narrow streams like the Whitewater River to appear. The cropland layer does not show any 'prohibited' areas along this particular riverbed, and thus the path is allowed to freely cross the stream. The path would be ideal otherwise. Looking at Figure 3. and a hill shade map that the ground-truth steep areas are red in the figure. Figure 6. shows that red values are closer to 100 than the flat areas. Looking at Figure 4. and a satellite image, the areas marked as green farm fields are indeed farm fields. Figure 6. shows again that undesirable areas are marked as high travel costs for both component rasters. Possible methods to improve this optimal path are discussed in the next section.

Figure 6. Values by Raster Display Color



## Discussion and Conclusion

Optimal path is an important element of my final project, which will rely on finding low slope and short distance accurately. It is now obvious that a high-resolution DEM is going to be required to achieve this goal, as well as an accurate vector layer of ground features. I now have the fundamental skills to accomplish the project after this lab.

Suitability analysis seems extremely subject to change based on how parameters and datasets are defined. Even if you use weighted and Boolean logic correctly on raster datasets, the data you choose initially heavily affect the result. Streams not appearing in the land use raster due to cell size means that the path technically broke the rules given to it in the real world, but did not in the data layer. The cell size in the slope DEM is 30 meters; there can be a very large amount of height fluxuation inside a 30 meters cell. By reducing the DEM cell size and cutting out all water using a water polygon layer, the accuracy of the final path could certainly be improved with enough time.

Again, Model Builder proved an invaluable tool in this exercise. After writing the code by hand, the path would not appear with any amount of adjustment. The Model Builder eliminated any possible typing errors and ensures inputs and outputs are correct

## References

- ESRI. (2020, July 14). Distance Analysis: Identifying Optimal Paths Using Rasters. Retrieved March 03, 2021, from <https://www.esri.com/training/catalog/60109c7e8106ed0454f90b25/distance-analysis%3A-identifying-optimal-paths-using-rasters/#!/arcgis-online-training/>
- ESRI. (2021). Arcgis pro geoprocessing tool reference. Retrieved March 02, 2021, from <https://pro.arcgis.com/en/pro-app/latest/tool-reference/main/arcgis-pro-tool-reference.htm>
- N.A.S.S. (2018). Cropland data Layer 2018, Minnesota. Retrieved March 03, 2021, from <https://gisdata.mn.gov/dataset/agri-cropland-data-layer-2018>



U.S.G.S. (2004). Minnesota digital elevation model - 30 meter resolution. Retrieved March 03, 2021, from <https://gisdata.mn.gov/dataset/elev-30m-digital-elevation-model>

### 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
<b>Structural Elements</b>	All elements of a lab report are included ( <b>2 points each</b> ): 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	<b>26</b>
<b>Clarity of Content</b>	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 ( <b>12 points</b> ). There is a clear connection from data to results to discussion and conclusion ( <b>12 points</b> ).	24	<b>23</b>
<b>Reproducibility</b>	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	<b>27</b>
<b>Verification</b>	Results are correct in that they have been verified in comparison to some standard. The standard is clearly stated ( <b>10 points</b> ), the method of comparison is clearly stated ( <b>5 points</b> ), and the result of verification is clearly stated ( <b>5 points</b> ).	20	<b>18</b>
		100	<b>94</b>