Lab Report

Title: Lab 2 - Part 2

Notice: Dr. Bryan Runck

Author: Mattie Gisselbeck

Date: November 2, 2022

Repository: <https://github.com/mgisselbeck/GIS5571>

Time Spent: 20 hours

## Abstract

The objective of this analysis is to compare three different weighting approaches to generating a cost surface for determining optimal paths for Dory, within Dory’s preferences. The data was sourced from Minnesota Geospatial Commons and was scraped through an ETL in ArcGIS Pro via a Python notebook. Based on Dory’s preferences, the objective is to find a path that avoids farm fields, water bodies without a bridge, and has a gradual slope. The results are shown in the figures below (see Figure 2 through Figure 7). The data flow diagram above (Figure 1) shows all the variables and commands I applied in finding an optimal route for Dory. The results could be qualitatively verified by using the ‘ArcGIS Pro – Topographic’ by visually comparing the route with the hill shade of the map. In this lab, I was able to build off pre-existing knowledge with creating an ETL and building a cost path analysis model. The objectives of this lab helped me to gain practical applications of how I would create a cost path analysis through ArcPy or an open-source package.

## Problem Statement

The objective of this analysis is to compare three different weighting approaches to generating a cost surface for determining optimal paths for Dory, within Dory’s preferences. Dory’s preferences are: (1) Dory prefers to not walk through any farm fields, (2) she doesn’t like crossing water bodies if there isn’t a bridge (though sometimes she doesn’t mind if she’s wearing her waders), and (3) prefers a path with the most gradual slope (Runck, 2022).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **#** | **Requirement** | **Defined As** | **(Spatial) Data** | **Attribute Data** | **Dataset** | **Preparation** |
| 1 | County Boundaries in Minnesota | Wabasha, Winona, and Olmstead County | .shp (Vector) | N/A | [Minnesota Geospatial Commons](https://gisdata.mn.gov/dataset/bdry-counties) | ETL |
| 2 | 2012 Assessed Streams | Assessed Streams for Minnesota in 2012 | .shp (Vector) | N/A | [Minnesota Geospatial Commons](https://gisdata.mn.gov/dataset/env-assessed-streams-2012) | ETL |
| 3 | NCLD 2019 Land Cover, Minnesota | Land Cover Classification for Minnesota | TIFF (Raster) | Land Cover Classification | [Minnesota Geospatial Commons](https://gisdata.mn.gov/dataset/biota-landcover-nlcd-mn-2019) | ETL |
| 4 | Minnesota Digital Elevation Model - 30 Meter Resolution | Elevation (Wabasha, Winona, and Olmsted County) | TIFF (Raster) | Elevation | [Minnesota Geospatial Commons](https://gisdata.mn.gov/dataset/elev-30m-digital-elevation-model) | ETL |
| 5 | Roads, Minnesota, 2012 | Road Centerlines for all Public Roads in Minnesota | .shp (Vector) | N/A | [Minnesota Geospatial Commons](https://gisdata.mn.gov/dataset/trans-roads-mndot-tis) | ETL |

*Table 1. Required Data*

## Input Data

The table below is a collection of data from the Minnesota Geospatial Commons. Data was scraped through an ETL in ArcGIS Pro via a Python notebook. All the data described below will be used in a cost path analysis to find three optimal routes for Dory.

|  |  |  |  |
| --- | --- | --- | --- |
| **#** | **Title** | **Purpose in Analysis** | **Link to Source** |
| 1 | County Boundaries in Minnesota | To be used in the cost surface equation to find optimal routes for Dory | [Minnesota Geospatial Commons](https://gisdata.mn.gov/dataset/bdry-counties) |
| 2 | 2012 Assessed Streams | To be used in the cost surface equation to find optimal routes for Dory | [Minnesota Geospatial Commons](https://gisdata.mn.gov/dataset/env-assessed-streams-2012) |
| 3 | NCLD 2019 Land Cover, Minnesota | To be used in the cost surface equation to find optimal routes for Dory | [Minnesota Geospatial Commons](https://gisdata.mn.gov/dataset/biota-landcover-nlcd-mn-2019) |
| 4 | Minnesota Digital Elevation Model - 30 Meter Resolution | To calculate the slope and add the output into the cost surface analysis to find optimal routes for Dory | [Minnesota Geospatial Commons](https://gisdata.mn.gov/dataset/elev-30m-digital-elevation-model) |
| 5 | Roads, Minnesota, 2012 | To be used in the cost surface equation to find optimal routes for Dory | [Minnesota Geospatial Commons](https://gisdata.mn.gov/dataset/trans-roads-mndot-tis) |

*Table 2. Input Data*

## Methods

Based on Dory’s preferences, the objective is to find three paths that avoid farm fields, water bodies without a bridge, and have a gradual slope. The starting point, Dory’s Farm, is 44.127985, -92.148796 and end point, picnic area in Whitewater State Park, is 44.054852, -92.045780.

**Diagram

Description automatically generated**

*Figure 1.* [*Data Flow Diagram*](https://github.com/mgisselbeck/GIS5571/blob/main/Lab2/Part%202/Graphics/Lab2_Part2_DataFlowDiagram.png) *for Cost Path Analysis.*

### Part 1.1: Import Packages and Request Data from Minnesota Geospatial Commons

(See Lab 2 – Part 1 Python Notebook)

### Part 1.2: Create a Study Extent

To create the study extent, I used ‘Feature Class to Feature Class’ to apply a SQL expression that selects only Winona, Olmsted, and Wabasha and create a new feature class.

|  |
| --- |
| # Create Study Extent (Feature Class to Feature Class)  arcpy.conversion.FeatureClassToFeatureClass("mn\_county\_boundaries", r"C:\Users\gisse015\Documents\ArcGIS\Projects\Lab2\_2\Lab2\_2.gdb", "StudyExtent", "CTY\_NAME = 'Wabasha' Or CTY\_NAME = 'Winona' Or CTY\_NAME = 'Olmsted'", 'AREA "AREA" true true false 19 Double 0 0,First,#,mn\_county\_boundaries,AREA,-1,-1;PERIMETER "PERIMETER" true true false 19 Double 0 0,First,#,mn\_county\_boundaries,PERIMETER,-1,-1;CTYONLY\_ "CTYONLY\_" true true false 19 Double 0  # Dissolve County Boundaries  arcpy.management.Dissolve("StudyExtent", r"C:\Users\gisse015\Documents\ArcGIS\Projects\Lab2\_2\Lab2\_2.gdb\StudyExtent\_Dissolve", None, None, "MULTI\_PART", "UNSPLIT\_LINES") |

### Part 1.3: Impervious Routes

The impervious roads were extracted by mask to match the study extent. Reclassification of impervious roads was executed by using the reclassify command (See Table 3).

|  |
| --- |
| # Extract by Mask  Extract\_Roads = arcpy.sa.ExtractByMask("NLCD\_2019\_Land\_Cover\_Impervious\_Descriptor.tif", "StudyExtent\_Dissolve"); Extract\_Roads.save(r"C:\Users\gisse015\Documents\ArcGIS\Projects\Lab2\_2\Lab2\_2.gdb\Extract\_Roads")  # Reclassify Roads (Scale: 1-10) (See Table 3)  arcpy.ddd.Reclassify("Extract\_Roads", "Class\_Name", "Unclassified 10;'Primary road' 1;'Secondary road' 1;'Tertiary road' 1;'Non-road non-energy impervious' 2;'LCMAP impervious' 3;'Wind turbines' 7", r"C:\Users\gisse015\Documents\ArcGIS\Projects\Lab2\_2\Lab2\_2.gdb\Reclass\_Roads", "DATA") |

|  |  |
| --- | --- |
| Value | New |
| Unclassified | 10 |
| Primary | 1 |
| Secondary | 1 |
| Tertiary | 1 |
| Non-road Impervious | 2 |
| LCMAP Impervious | 3 |
| Wind Turbines | 7 |

*Table 3. Reclassification for Impervious Roads.*

### Part 2.4: Digital Elevation Model (DEM)

The ‘Extract by Mask’ command was used to clip the DEM into the study extent. To calculate the slope, the ‘Slope’ tool was ran. The slope was reclassified by using the ‘Reclassify’ command (See Table 4).

|  |
| --- |
| # Extract by Mask  Extract\_DEM = arcpy.sa.ExtractByMask(r"C:\Users\gisse015\Documents\ArcGIS\Projects\Lab2\_2\elev\_30m\_digital\_elevation\_model.gdb\digital\_elevation\_model\_30m", "StudyExtent\_Dissolve"); Extract\_DEM.save(r"C:\Users\gisse015\Documents\ArcGIS\Projects\Lab2\_2\Lab2\_2.gdb\Extract\_DEM")  # Calculate Slope  arcpy.ddd.Slope("Extract\_DEM", r"C:\Users\gisse015\Documents\ArcGIS\Projects\Lab2\_2\Lab2\_2.gdb\DEM\_Slope", "PERCENT\_RISE", 1, "PLANAR", "METER")  # Reclassify Slope (Scale: 1-10) (See Table 4)  arcpy.ddd.Reclassify("DEM\_Slope", "VALUE", "0 3 1;3 6 2;6 10 3;10 15 4;15 20 8;20 25 9;25 30 10;30 40 10;40 60 10;60 100 10;100 1000 10", r"C:\Users\gisse015\Documents\ArcGIS\Projects\Lab2\_2\Lab2\_2.gdb\Reclass\_DEM\_2", "DATA") |

|  |  |  |
| --- | --- | --- |
| Start | End | New |
| 0 | 3 | 1 |
| 3 | 6 | 2 |
| 6 | 12 | 3 |
| 12 | 15 | 4 |
| 15 | 18 | 8 |
| 18 | 21 | 9 |
| 21 | 24 | 10 |
| 24 | 27 | 10 |
| 27 | 30 | 10 |
| 30 | 533.47 | 10 |

*Table 4. Reclassification for DEM Slope.*

### Part 2.5: NLCD

The NLCD was extracted by ‘Extract by Mask’ command to fit the study extent and then reclassified (See Table 5).

|  |
| --- |
| # Extract by Mask  Extract\_NLCD = arcpy.sa.ExtractByMask(r"C:\Users\gisse015\Documents\ArcGIS\Projects\Lab2\_2\tif\_biota\_landcover\_nlcd\_mn\_2019\NLCD\_2019\_Land\_Cover.tif", "StudyExtent\_Dissolve"); Extract\_NLCD.save(r"C:\Users\gisse015\Documents\ArcGIS\Projects\Lab2\_2\Lab2\_2.gdb\Extract\_NLCD")  # Reclassify NLCD  arcpy.ddd.Reclassify("Extract\_NLCD", "NLCD\_Land", "'Open Water' 10;'Developed, Open Space' 2;'Developed, Low Intensity' 2;'Developed, Medium Intensity' 2;'Developed, High Intensity' 2;'Barren Land' 5;'Deciduous Forest' 7;'Evergreen Forest' 7;'Mixed Forest' 7;Shrub/Scrub 7;Herbaceous 7;Hay/Pasture 9;'Cultivated Crops' 9;'Woody Wetlands' 10;'Emergent Herbaceous Wetlands' 10", r"C:\Users\gisse015\Documents\ArcGIS\Projects\Lab2\_2\Lab2\_2.gdb\Reclass\_NLCD", "DATA") |

|  |  |
| --- | --- |
| Value | New |
| Open Water | 10 |
| Developed, Open Space | 1 |
| Developed, Low Intensity | 2 |
| Developed, Medium Intensity | 2 |
| Developed, High | 2 |
| Barren Land | 5 |
| Deciduous Forest | 7 |
| Evergreen Forest | 7 |
| Mixed Forest | 9 |
| Shrub/Scrub | 9 |
| Herbaceous | 9 |
| Hay/Pasture | 9 |
| Cultivated Crops | 10 |

*Table 5. Reclassification for NLCD.*

### Part 2.6: Streams with Strahler Stream Order

Streams with Strahler Stream Order were clipped to the study extent by using the “Clip” tool and then converted into a raster. The reason behind converting the feature to a raster is so we can use it in the raster calculator and weighted overlay. The streams were reclassified using the ‘Reclassify’ command. The ‘Raster Calculator’ command was implemented to convert the null values within the dataset to 1 using the SQL expression: Con(IsNull("Reclass\_Streams"),1, "Reclass\_Streams".

|  |
| --- |
| # Clip Streams  arcpy.analysis.Clip("streams\_with\_strahler\_stream\_order", "StudyExtent\_Dissolve", r"C:\Users\gisse015\Documents\ArcGIS\Projects\Lab2\_2\Lab2\_2.gdb\Streams\_Clip", None)  # Feature to Raster  arcpy.conversion.FeatureToRaster("Streams\_Clip", "SO\_VALUE", r"C:\Users\gisse015\Documents\ArcGIS\Projects\Lab2\_2\Lab2\_2.gdb\Feature\_Stre1", r"C:\Users\gisse015\Documents\ArcGIS\Projects\Lab2\_2\Lab2\_2.gdb\Reclass\_DEM\_Slope")  # Reclassify Streams  arcpy.ddd.Reclassify("Feature\_Stre1", "Value", "1 1;2 2;3 3;4 7;5 8;6 9;8 10", r"C:\Users\gisse015\Documents\ArcGIS\Projects\Lab2\_2\Lab2\_2.gdb\Reclass\_Streams", "DATA")  # Raster Calculator (Note: Command only worked in Geoprocessing pane in ArcGIS Pro)  output\_raster = arcpy.ia.RasterCalculator(' Con(IsNull("Reclass\_Streams"),1, "Reclass\_Streams")'); output\_raster.save(r"c:\Users\gisse015\documents\ArcGIS\Projects\Lab2\_2\Lab2\_2.gdb\RC\_Streams") |

### Part 2.7: Dory’s Farm (Start Point) and Picnic Area (End Point)

|  |
| --- |
| # Dory's Farm (Start Point)  arcpy.management.XYTableToPoint("Start.csv", r"C:\Users\gisse015\Documents\ArcGIS\Projects\Lab2\_2\Lab2\_2.gdb\Start\_Point", "LONG", "LAT", None, 'GEOGCS["GCS\_WGS\_1984",DATUM["D\_WGS\_1984",SPHEROID["WGS\_1984",6378137.0,298.257223563]],PRIMEM["Greenwich",0.0],UNIT["Degree",0.0174532925199433]];-400 -400 1000000000;-100000 10000;-100000 10000;8.98315284119521E-09;0.001;0.001;IsHighPrecision')  # Picnic Area (End Point)  arcpy.management.XYTableToPoint("End.csv", r"C:\Users\gisse015\Documents\ArcGIS\Projects\Lab2\_2\Lab2\_2.gdb\End\_Point", "LONG", "LAT", None, 'GEOGCS["GCS\_WGS\_1984",DATUM["D\_WGS\_1984",SPHEROID["WGS\_1984",6378137.0,298.257223563]],PRIMEM["Greenwich",0.0],UNIT["Degree",0.0174532925199433]];-400 -400 1000000000;-100000 10000;-100000 10000;8.98315284119521E-09;0.001;0.001;IsHighPrecision')  # Point to Raster - Dory's Farm (Start Point)  arcpy.conversion.PointToRaster("Start\_Point", "OBJECTID", r"C:\Users\gisse015\Documents\ArcGIS\Projects\Lab2\_2\Lab2\_2.gdb\Start\_RasterPoint", "MOST\_FREQUENT", "NONE", None, "BUILD")  # Point to Raster - Picnic Area (End Point)  arcpy.conversion.PointToRaster("End\_Point", "OBJECTID", r"C:\Users\gisse015\Documents\ArcGIS\Projects\Lab2\_2\Lab2\_2.gdb\End\_RasterPoint", "MOST\_FREQUENT", "NONE", None, "BUILD") |

### Part 1.8: Cost Path A

|  |
| --- |
| # Weighted Overlay (Cost Surface)  Cost\_Surface = arcpy.sa.WeightedOverlay("('Reclass\_Streams\_A' 25 'Value' (1 1; 2 2; 3 3; 7 8; 8 9; 9 10; 10 10; NODATA NODATA); 'Reclass\_DEM\_Slope\_A' 25 'Value' (2 2; 3 3; 4 4; 7 8; 8 9; 10 10; NODATA NODATA); 'Reclass\_NLCD\_A' 25 'Value' (2 2; 5 6; 7 8; 9 9; 10 10; NODATA NODATA); 'Reclass\_Roads\_A' 25 'Value' (1 1; 2 2; 3 3; 7 8; 10 10; NODATA NODATA));1 10 1"); Cost\_Surface.save(r"C:\Users\gisse015\Documents\ArcGIS\Projects\Lab2\_2\Lab2\_2.gdb\Cost\_Surface\_A")  # Cost Distance  Cost\_Distance = arcpy.sa.CostDistance("DorysFarm", "out\_raster", None, r"C:\Users\gisse015\Documents\ArcGIS\Projects\Lab2\_2\Lab2\_2.gdb\Direction\_Raster", None, None, None, None, ''); Cost\_Distance.save(r"C:\Users\gisse015\Documents\ArcGIS\Projects\Lab2\_2\Lab2\_2.gdb\CostDis")  # Cost Path (Best Single)  Cost\_Path = arcpy.sa.CostPath("PicnicArea", "CostDis", "Direction\_Raster", "BEST\_SINGLE", "ID", "INPUT\_RANGE"); Cost\_Path.save(r"C:\Users\gisse015\Documents\ArcGIS\Projects\Lab2\_2\Lab2\_2.gdb\CostPath") |

### Part 1.9: Cost Path B

|  |
| --- |
| # Weighted Overlay (Cost Surface)  # Cost Distance  # Cost Path (Best Single) |

### Part 1.10: Cost Path C

|  |
| --- |
| # Weighted Overlay (Cost Surface)  # Cost Distance  # Cost Path (Best Single) |

## Results

The results are shown in the figures below (see Figure 2 through Figure 7). The main themes of the lab were preparing data in an ETL pipeline, creating a cost surface to find an optimal route. The data flow diagram above (Figure 1) shows all the variables and commands I applied in finding an optimal route for Dory.

*Map

Description automatically generated*

*Figure 7. Results for Part 2.10: Cost Path (Optimal Route for Dory)*

## Results Verification

The results could be qualitatively verified by using the ‘ArcGIS Pro – Topographic’ by visually comparing the route with the hill shade of the map. There are many route outputs an analyst can get based off how they chose to reclassify. Using reliable data sources like Minnesota Geospatial Commons helps to ensure quality results if used correctly.

# Discussion and Conclusion

In this lab, I was able to build off pre-existing knowledge with creating an ETL and building a cost path analysis model. The objectives of this lab helped me to gain practical applications of how I would create a cost path analysis through ArcPy or an open-source package. The biggest roadblocks for this lab were creating the cost distance output for the cost path analysis and how computationally intensive it was for my computer. In the future, I hope to apply what I learned in this lab to learn how to learn how to code using packages like numpy and rasterio.

## References

Runck, B. 2022. GIS 5571: Lab 2.

https://github.com/mgisselbeck/GIS5571/blob/main/Lab2/Lab2\_Instructions.pdf

## 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 | **27** |
| **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 | **27** |
| **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 | **19** |
|  |  | 100 | **97** |