**Lab Report**

Title: Gone Fishin’, But Only On A Very Specific Route: Cost Surface Models

Notice: Dr. Bryan Runck

Author: Cecelia Isaac

Date: 03/04/2021

**Project Repository:** https://github.com/CeceliaAi/GIS5572/tree/master/Lab2

**Abstract**

The problem we will address in this lab is to create a cost surface model for a walking route through southern Minnesota. In our model, we will look for the driest, most level walking path for Dory. We will use several datasets to create the model: an agricultural raster of farm fields, a hydrology feature layer, and a DEM of Wabasha and Winona counties. To begin, we will build an ETL to download the DEM data from the DNR FTP server. Then, we use ArcPro to reclassify the rasters with weighted cells, then combine the rasters to create the cost surface, and finally compute the best path from the origin to the destination. The results will be the best path in the form of a polyline. We can first verify these results visually (e.g. does the route cross water only when there is a bridge?) and also through successful running of the tools. Since everyone’s process is different, the results across the class may not be comparable. In our final conclusions, we will discuss some of the choices made when weighting the data.

**Problem Statement**

We are creating a cost raster in order to compute the ideal path from Dory’s farm to her preferred fly fishing location. After obtaining the data, most of the lab is devoted to cleaning and processing data so that it is in the preferred format, or gives us the information we need. After we have the input and output rasters, computing the ideal route can be done through a final single tool. The water data comes in polygon format, and needs to be put into raster format. I also imported road data and simply erased the sections of water where a road line overlapped. The DEM is first created from LiDAR data, and then run through the Slope tool to get a slope raster. The agricultural data needed to be reclassified from many categories into two: one with farm fields, which are undesirable, and one with urban, non-farm fields, and trees, which are acceptable.

*Table 1. Data*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **#** | **Requirement** | **Defined As** | **Spatial Data** | **Attribute Data** | **Dataset** | **Preparation** |
| 1 | Buffer roads | Used to create breaks in the water feature layer | Road geometry |  | [Mn GeoSpatial Commons](https://gisdata.mn.gov/dataset/trans-roads-mndot-tis) | Crop to study area |
| 2 | Erase and reclassify rivers | Used as a barrier that cannot be crossed | Water polyline |  | [Mn Geospatial Commons](https://gisdata.mn.gov/dataset/us-mn-state-metc-water-lakes-rivers) | Crop to study area |
| 3 | Crop and reclassify cropland – Wabasha and Winona | Used to define ideal walking areas—farm fields are not preferred; developed areas, trees, and open fields are acceptable | Raster |  | [Mn Geospatial Commons](https://gisdata.mn.gov/dataset/agri-cropland-data-layer-2018) | Crop to study area |
| 4 | Convert to Slope and reclassify a DEM of Wabasha and Winona | Used to determine the most level slope on the route | Raster |  | [DNR FTP](ftp://ftp.lmic.state.mn.us/pub/gdrs/data/pub/us_mn_state_mda/agri_cropland_data_layer_2019.zip) | Combine |

**Input Data**

The data used will be a raster of agricultural land cover in Minnesota, a DEM of Winona and Wabasha counties, a polyline water layer, and a roads layer. The landcover data came from the DNR. The DEMs, roads, and water came from MnGeo.

*Table 2. Data sources*

|  |  |  |  |
| --- | --- | --- | --- |
| **#** | **Title** | **Purpose in Analysis** | **Link to Source** |
| 1 | Minnesota Roads | Raw input dataset for routing analysis from MNDOT | [Mn GeoSpatial Commons](https://gisdata.mn.gov/dataset/trans-roads-mndot-tis) |
| 2 | Lakes and Rivers | Raw input water polyline data | [Mn Geospatial Commons](https://gisdata.mn.gov/dataset/us-mn-state-metc-water-lakes-rivers) |
| 3 | Cropland | Landuse raster file | [Mn Geospatial Commons](https://gisdata.mn.gov/dataset/agri-cropland-data-layer-2018) |
| 4 | Elevation – Wabasha Winona | Digital elevation model of two counties | [DNR FTP](ftp://ftp.lmic.state.mn.us/pub/gdrs/data/pub/us_mn_state_mda/agri_cropland_data_layer_2019.zip) |

**Methods**

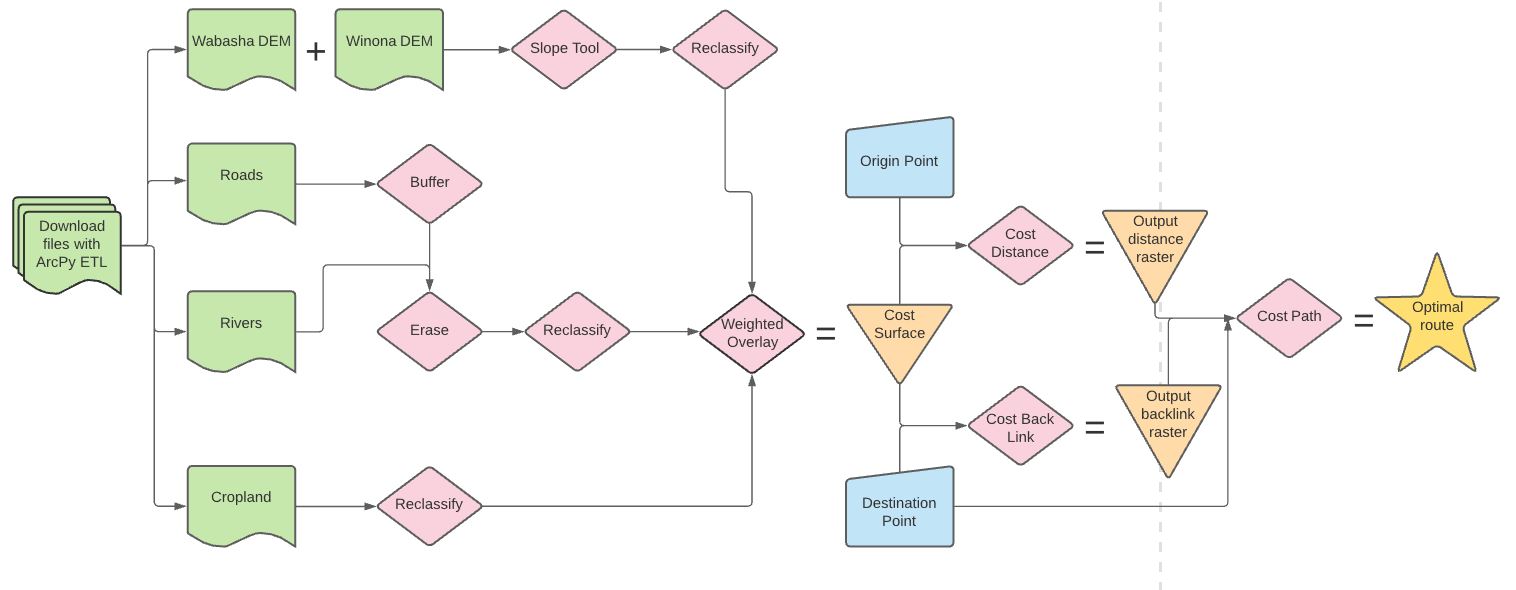
To create the ETL pipeline, I wrote Python code to download the data from MnGeo. Once the data had been loaded, I used ArcPro to crop to my study area (Figure 1). Then I loaded the DEM data. I combined the rasters of the two counties into one and then ran the Slope tool to get a slope output raster. I reclassified the slope data to weight the steeper slopes higher (and therefore less optimal).

Next, I cropped water and roads data to the study area. I buffered the roads layer and used that to erase areas of the river data, on the assumption that a road crossing the water implied a bridge that Dory could use. I had to fiddle with the buffer size in order to ensure the gap covered a whole pixel of space that would allow Dory to cross through. I then converted the water layer to a raster, and reclassified it.

I reclassified the croplands data into two classes. Farm fields of all types were one class, and developed land, fallow fields, forests, and sunflower patches were another. Finally, I created two separate point layers to act as my origin and destination points.

First I used Distance Accumulation to create surface rasters, but when I put them into the Optimal Path tool, they did not work. I switched to using the Legacy tools. I used Weighted Overlay to create the cost surface (Figure 2). I used water as a barrier and made less steep cells, as well as cells in the non-farm-field category, more desirable. After this, I used Cost Distance and Cost Back Link to create two rasters from the origin and destination points. Then, I put everything into the Cost Path tool. The output of this tool is the optimal path for Dory to walk (Figure 3).

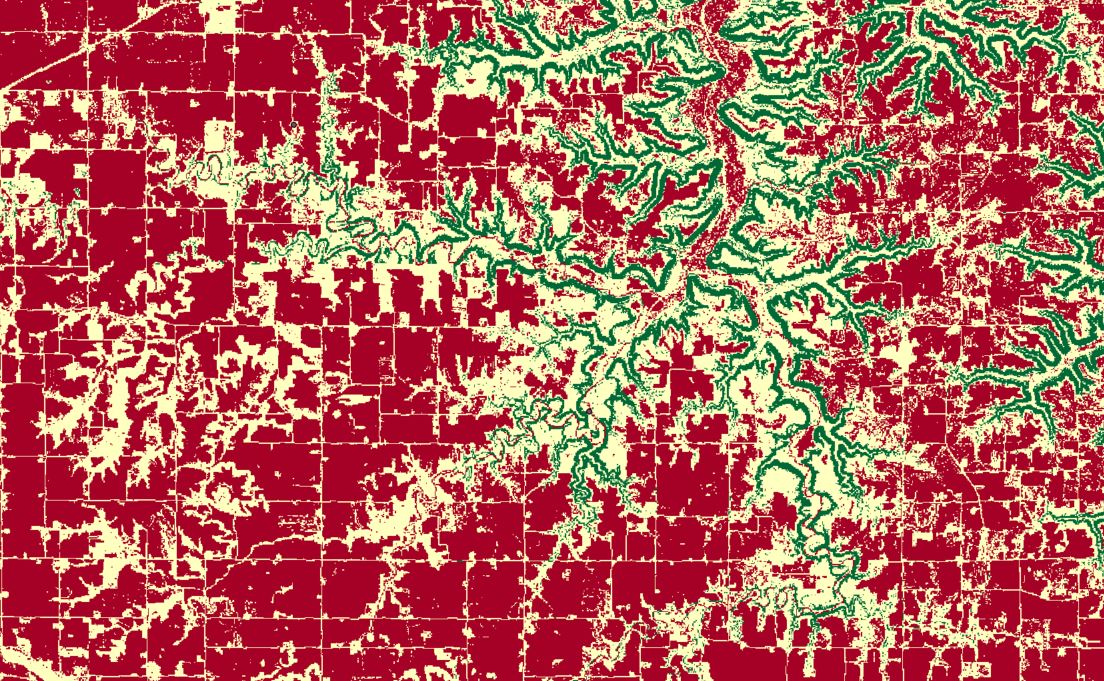
*Figure 1. One big data flow diagram.*

**

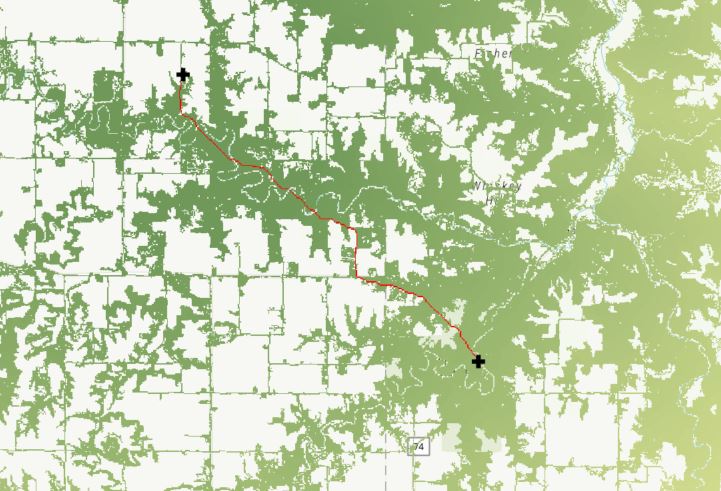
**Results**

The results are a geodatabase containing a collection of cost rasters weighted to the specifics of Dory’s preferences. There is also a polyline along the optimal path for Dory to walk. These results address the problem statement by producing the desired output (Figure 3). The results also include an ETL and the constructed cost surface (Figure 2), which is based on our interpretation of how the input data should be classified and weighted.

*Figure 2. Cost Surface*

**

*Figure 3. Output polyline route*

**

**Results Verification**

The results can be verified by error-free code and tool outputs. They can also be verified visually by the polyline output, and checking that this output follows the rules set out by the cost surface. In our case, that means water is a barrier and that farm fields are an undesirable route.

**Discussion and Conclusion**

I learned how to properly format raster data in order to combine it to create a cost surface. In order to answer the main problem, most of the work was done on the front end to prepare everything. The main problem itself was then very simple, as all the component parts were loaded into one tool (Cost Path). During the process of preparing the data, I also made a lot of choices about how the data would be weighted based on my interpretation of the instructions. There were also some practical decisions to be made. For example, downloading the data took a very long time, so I only used two counties, when it is possible the true best route would cut through a third county. I also used Erase on some of my data, which may not be the best choice or even an option in future situations. The data also leaves out important real-world scenarios such as impassable areas or closed-off roads.

**References**

Boomiverse. (n.d.). *FTP Error—550 Failed to change directory* [Blog]. Retrieved February 14, 2021, from <https://community.boomi.com/s/article/ftperror550failedtochangedirectory>

ESRI. (n.d.-a). *LAS Dataset To TIN (3D Analyst)* [Documentation]. Retrieved February 14, 2021, from <https://pro.arcgis.com/en/pro-app/latest/tool-reference/3d-analyst/las-dataset-to-tin.htm>

ESRI. (n.d.-b). *LAS To Raster function* [Documentation]. Retrieved February 14, 2021, from <https://pro.arcgis.com/en/pro-app/latest/help/analysis/raster-functions/las-to-raster-function.htm>

*ftplib—FTP protocol client*. (n.d.). [Documentation]. Retrieved February 14, 2021, from <https://docs.python.org/3/library/ftplib.html>

Kite. (n.d.). *How to download a file from an FTP server in Python*. Retrieved February 14, 2021, from <https://www.kite.com/python/answers/how-to-download-a-file-from-an-ftp-server-in-python>

*Lidar Elevation Data for Minnesota: 2008-2012*. (n.d.). Retrieved February 14, 2021, from [https://www.mngeo.state.mn.us/chouse/elevation/lidar.html#](https://www.mngeo.state.mn.us/chouse/elevation/lidar.html)

*Python: Download a file from an FTP server*. (n.d.). [Forum]. Retrieved February 14, 2021, from <https://stackoverflow.com/questions/11768214/python-download-a-file-from-an-ftp-server>

*Requests: HTTP for Humans*. (n.d.). [Documentation]. Retrieved February 14, 2021, from <https://requests.readthedocs.io/en/master/>

Rockikz, A. (n.d.). *How to Download and Upload Files in FTP Server using Python*. Retrieved February 14, 2021, from <https://www.thepythoncode.com/article/download-and-upload-files-in-ftp-server-using-python>

**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** |