## Lab Report

Title: Lab 2 Part 2
Notice: Dr. Bryan Runck
Author: Eric Gibson
Date: 10/31/2023

**Project Repository:** https://github.com/gibso632/GIS5571/tree/main/Lab2
**Google Drive Link:**
**Time Spent:** 16 hours


### Abstract

This project introduces creating a cost surface within GIS software, specifically ArcGIS Pro, using an automated script to solve an example word problem. The problem specified a person named Dora who needed to find the best path to a picnic area within Whitewater State Park. The parameters included avoiding any farmland, as it can be swampy during the Spring, avoiding areas which are too steep to traverse, and potentially avoiding rivers if she did not have her waders. It was eventually determined to establish any farmland as areas to avoid (higher value), rivers as medium-risk areas (medium value), and slope as an increasing risk (greater the slope angle, the higher the value). Utilizing a land cover raster dataset and a digital elevation model (DEM), five separate cost surfaces were made with differing importance between land cover types and slope. This allowed for Dora to determine which element (slope or land cover) was preferred over the other when traversing.


### Problem Statement

The issue being studied within this lab is regarding a person named Dora who enjoys fly fishing at Whitewater State Park, specifically at the North Picnic Area within the park. Her farm is located at 44.127985, -92.148796 and she has certain parameters for travelling by foot to the picnic area. These parameters are 1) avoiding farmland due to the muddiness of the soil in the Spring, 2) avoiding water features unless she has her waders, and 3) avoiding areas which are too steep to traverse. Outside of these three parameters, the path she could take is relatively open and she does not have a stated preference on which parameter is most important. Since she has waders, however, it is likely water features may not have as much weight as farmland or slope.

*Table 1. Deliverables for this lab.*

| # | Requirement | Defined As | (Spatial) Data | Attribute Data | Dataset | Preparation |
|---|---|---|---|---|---|---|
| 1 | Points illustrating both Dora's farm and the North Picnic Area | Point feature classes containing points where Dora's house and the Whitewater State Park | Location points | Name of location | Google Places API with your API Key | Use arcpy.Point and arcpy.PointGeometry to create coordinate points. |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | North Picnic Area are located. | | | | |
| 2 | Cost surface rasters | Five separate cost surface rasters depicting how different parameters could be weighted by Dora. | Rasters in GeoTIFF format | Cost between farmland, water, and slope | GeoTIFF rasters | Create slope and land use rasters weighted based on Dora's parameters, then weight those rasters differently to provide multiple cost surfaces |

## Input Data

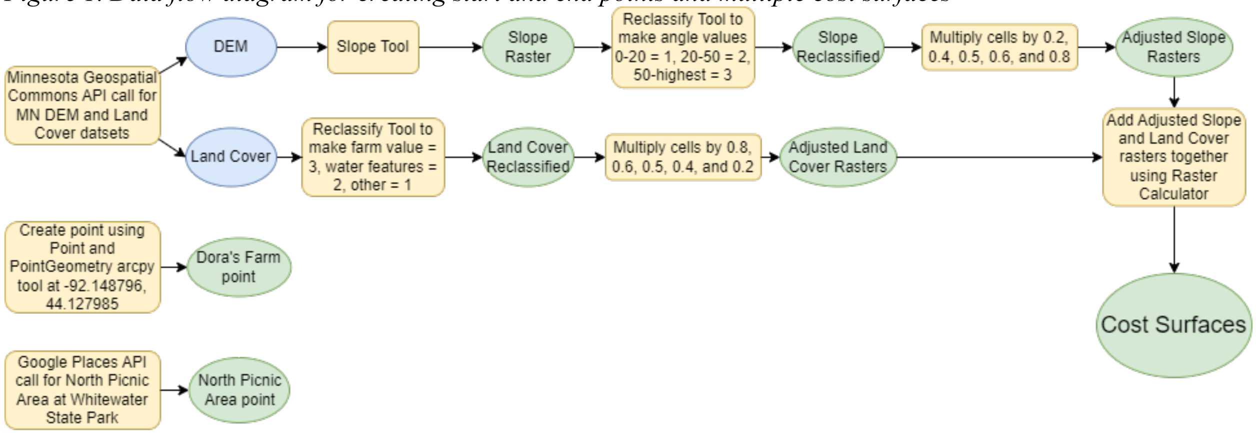
Three datasets are needed to create the final deliverables for this lab. This includes a GeoJSON gathered from the Google Places API containing the coordinates for the North Picnic Area in Whitewater State Park, a land use raster containing geographic data of farms within the area (in this case, the entirety of Minnesota), and a DEM of Minnesota. The land use raster will allow for the farmland and water features to be weighted based on the above parameters and the DEM can be converted to a slope raster depicting the angle of the slope with various weights established depending on the severity of the angle. The point data for Dora's house will be created via the arcpy.Point and arcpy.PointGeometry tools.

*Table 2. Datasets used within this lab.*

| # | Title | Purpose in Analysis | Link to Source |
|---|---|---|---|
| 1 | North Picnic Area at Whitewater State Park from Google Places API | Providing a destination for Dora determined by the cost analysis raster. | Google Places API – need API Key |
| 2 | NLCD 2019 Land Cover, Minnesota | For determining areas of farms and water features to find the best route for Dora depending on land use | Minnesota Geospatial Commons API |
| 3 | Minnesota Digital Elevation Model – 30 Meter Resolution | For creating a raster depicting the slope to find areas which may be too steep to traverse. | Minnesota Geospatial Commons API |

## Methods

*Figure 1. Data flow diagram for creating start and end points and multiple cost surfaces*




The data flow diagram above depicts both creating the cost surfaces as well as the points separately. Creating the North Picnic Area point also involved the arcpy.Point and arcpy.PointGeometry functions, but the coordinates instead originated within the GeoJSON provided by the Google Places API (44.05311337010728, -92.04608057989272). The DEM, as mentioned above, was utilized to create a slope raster which was reclassified so that an angle between 0 and 20 degrees was given a value of 1, an angle between 20 and 50 degrees was given a value of 2, and an angle between 50 and the maximum angle of 80.2473 was given a value of 3. The land cover raster was reclassified so that any farmland (land use labelled as "Hay/Pasture" or "Cultivated Crops") was given a value of 3, water features (land use labelled as "Open Water", "Woody Wetlands", and "Emergent Herbaceous Wetlands") were given a value of 2, and any other land use was given a value of 1. It is also worth noting the land use raster contained values depicting each land cover type. This workflow ultimately describes a larger value as less optimal for Dora.

Once these rasters with the weighted land use and slope values were created, they were both multiplied by 0.2, 0.4, 0.5, 0.6, and 0.8 to create five separate rasters with differing overall weights using a for loop which looped through a range of 1-4. The rasters were multiplied by 0.5 separately since the loop only increased by 0.2 and each number within the range was used to number the rasters. Each raster was then added to each other using the **Raster Calculator** tool in Arcpy inversely, meaning the reclassified slope raster multiplied by 0.2 was added to the reclassified land use raster multiplied by 0.8, then the slope raster multiplied by 0.4 was added to the land use raster multiplied by 0.6, and so on. This output five different cost surfaces with different weights provided for the slope and land use. The cost surface of the sum between the slope raster multiplied by 0.2 and the land use raster multiplied by 0.8 determined areas to traverse if avoiding certain land uses were most important and vice versa for the cost surface created by the sum of the land use raster multiplied by 0.2 and the slope raster multiplied by 0.8. The cost surface created via the sum of both rasters multiplied by 0.5 determined areas to traverse if both land use type and slope were equally important to avoid. The rasters multiplied by 0.4 and 0.6 depicted cost surfaces where either slope or land use type was slightly more important to avoid.

## Results

The following figures are ArcGIS Pro screenshots of the rasters from land use being the most important to avoid to an extreme slope being the most important to avoid. Red colors depict areas to avoid, blue colors depict possible routes.

*Figure 2. Cost surface with land use weighted as (multiplied by) 0.8 and slope weighted as 0.2.*

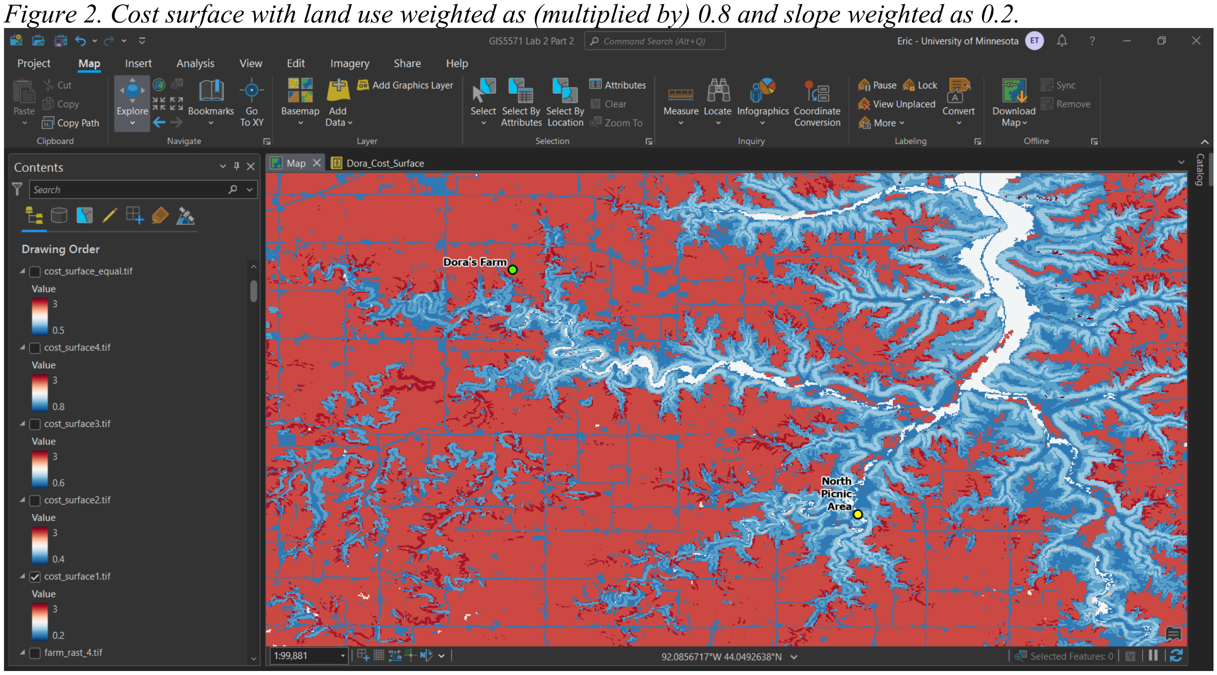



*Figure 3. Cost surface with land use weighted as (multiplied by) 0.6 and slope weighted as 0.4.*

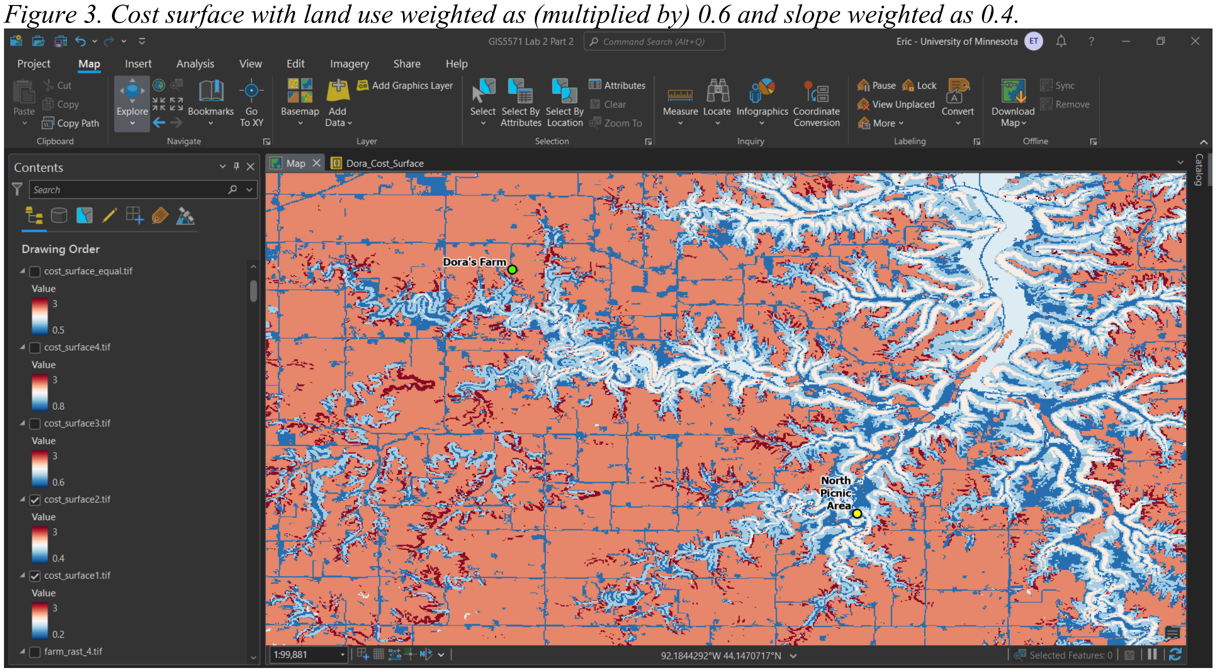


*Figure 4. Cost surface with land use and slope weighted the same (multiplied by 0.5 each).*

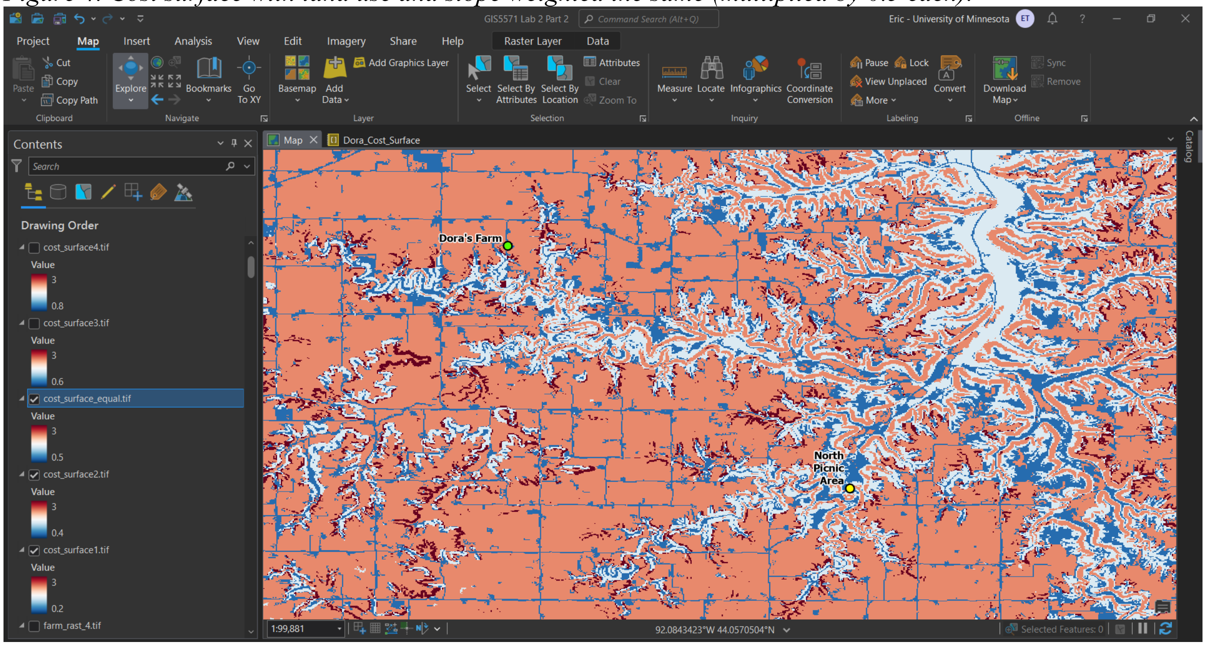



*Figure 5. Cost surface with land use weighted as (multiplied by) 0.4 and slope weighted as 0.6.*

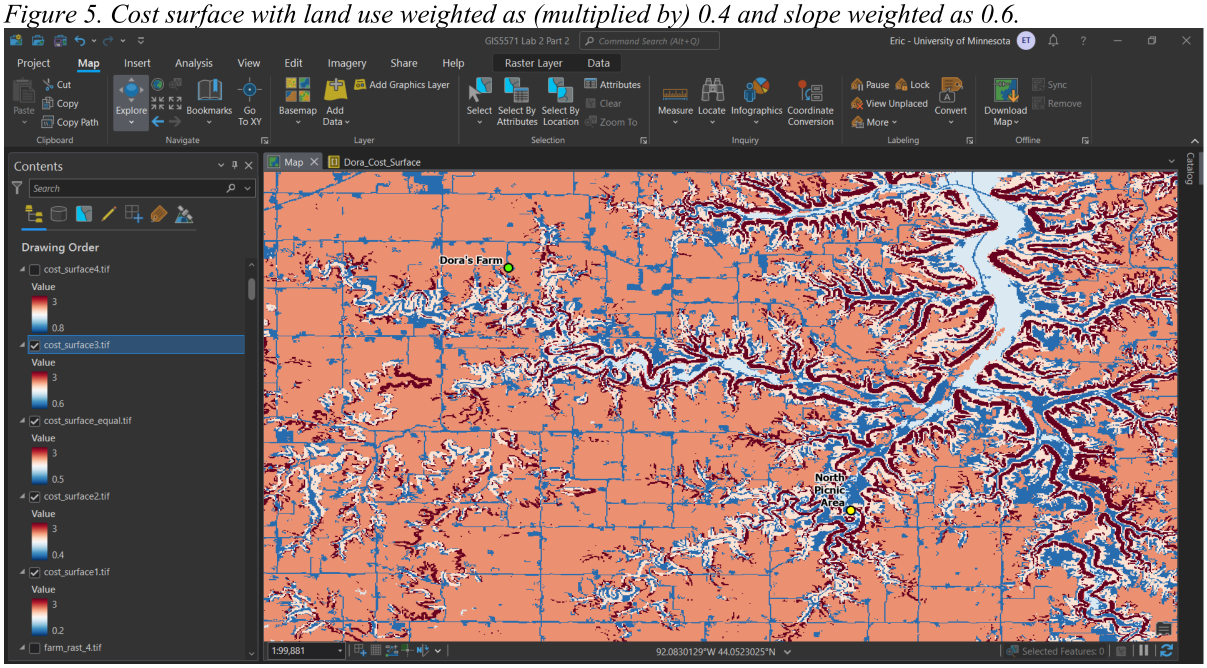
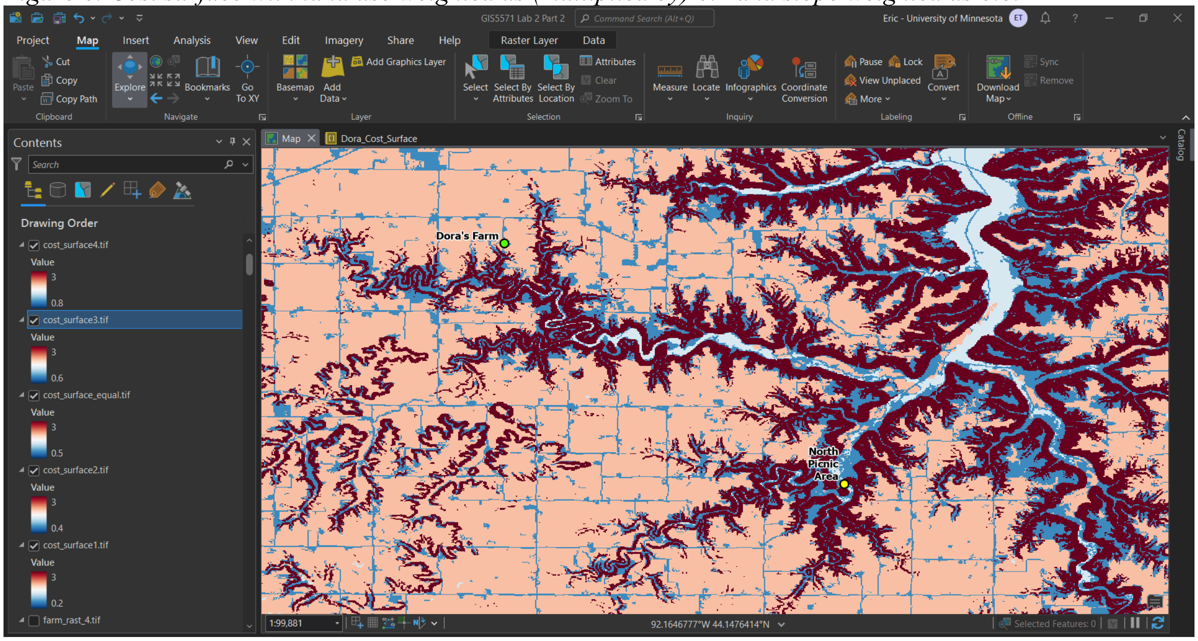


*Figure 6. Cost surface with land use weighted as (multiplied by) 0.2 and slope weighted as 0.8.*




These rasters feature a look at cost surfaces from where land uses are the most weighted for avoidance to where slope is the most weighted for avoidance comparatively. The point labelled "Dora's Farm" is Dora's starting point and the point labelled "North Picnic Area" is Dora's endpoint created via the arcpy.PointGeometry(arcpy.Point(*coordinates*)) function. The gradual change from farmland having the greatest avoidance value to slope having the greatest can be seen via the change in the location of the red coloration.

## Results Verification

The output needed to provide Dora with information regarding a path to take to the North Picnic Area at Whitewater State Park was relatively open. Outside of parameters mentioning avoiding farmland, potentially water features, and a steep slope, there was no determination of what should be weighted more: land use or slope. This led to the creation of several different cost surfaces with varying weights for both and the outputs in the **Results** section seem to replicate that. **Figure 2** depicts the cost surface where land use is the most weighted and **Figure 6** depicts the cost surface where slope is the most weighted with **Figures 3** to **5** gradually moving from land use being the most weighted to slope being the most weighted. When viewing these figures, there is a clear pattern showing this gradual change, confirming the validity of the cost surfaces.

Additionally, it is important to note there is also varying weights within the land use itself. Since Dora has waders, while she prefers to avoid water features, she can still traverse through them. This is why the water features were reclassified with a value of 2 instead of 3 and the cost surfaces do seem to validate this as well. When viewing **Figure 2**, while the land use is the most weighted, the water features only have a moderate avoidance value. This clearly describes how there is a difference between the weight of water features and farmland within the land use dataset.

## Discussion and Conclusion

This lab was a great introduction to creating a cost surface and will help immensely moving forward, especially considering my Final Project will likely utilize a cost surface to create trails for high pointing. I learned there can be multiple weights within the same cost surface depending on circumstances and needs of individuals, companies, organizations, etc. This is important to understand and create cost surfaces with various parameters which may not be able to be enforced otherwise. This relates to the problem with this lab because Dora not only has parameters and preferences between land use and slope, she also has certain preferences regarding specific land use types. Land uses are all within the same dataset, at least for this study, so understanding how to provide weights for certain attributes is important to provide the correct deliverables.

## References


"Where Developers Learn, Share, & Build Careers." *Stack Overflow*, 2008, stackoverflow.com/. Accessed 31 Oct. 2023.

Multiple prompts. *ChatGPT*, October 2023 version, OpenAI, 31 October 2023, chat.openai.com.

Chu, Justin. "Overview." *Google*, developers.google.com/maps/documentation/places/web-service/overview. Accessed 09 Oct. 2023.

National Land Cover Database. "NLCD 2019 Land Cover, Minnesota." *Minnesota Geospatial Commons*, 2019, gisdata.mn.gov/dataset/biota-landcover-nlcd-mn-2019.

USGS. "Minnesota Digital Elevation Model - 30 Meter Resolution." *Minnesota Geospatial Commons*, 2004, gisdata.mn.gov/dataset/elev-30m-digital-elevation-model.

ESRI. "ArcPy: Documentation: Arcgis Developers." *Documentation | ArcGIS Developers*, 2023, developers.arcgis.com/documentation/arcgis-add-ins-and-automation/arcpy/.

ESRI. "ArcGIS pro Geoprocessing Tool Reference." *ArcGIS Pro Geoprocessing Tool Reference-ArcGIS Pro | Documentation*, 2023, pro.arcgis.com/en/pro-app/latest/tool-reference/main/arcgis-pro-tool-reference.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** | Each element above is executed at a professional level so that | 24 | **20** |

| | | | |
|---|---|---|---|
| **Content** | 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)**. | | |
| **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 | **24** |
| **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 | **16** |
| | | 100 | **88** |