Lab Report

Title: Lab 2
Notice: Dr. Bryan Runck
Author: Yaxuan Zhang
Date: 10/31/2023

Project Repository: https://github.com/YaxuanSeanZhang/MGIS_ARCGIS/tree/main/GIS%205571/Lab1
Google Drive Link:
Time Spent: 8 hrs

Abstract
ETL is used to gather, process, and transfer data from various sources to a data warehouse or other target systems for analysis and reporting. In this lab, we will build ETL pipeline to conduct API queries with raster, cube, TIN, and Terrain data transformation steps to create an extract, transfer, and load system for LiDAR data and .bil file. We will also convert the .bil data into spacetime cubes to show the spatiotemporal trends of precipitation.

Problem Statement
In this lab, we will implement ETL process for lidar and .bil data. For lidar data, we will convert it into DEM and TIN, and visualize them as pdf files. For .bil data, we will conver it into spatiotemporal cube and visualize it as an animation of the timeseries.

Table 1. Main Steps

#	Requirement	Defined As	(Spatial) Data	Attribute Data	Dataset	Preparation
1	Extract Data from API	Raw input dataset pulling from geospatial common API	raster	Elevation; landcover	geospatial common	Define API, and pull data
2	Create Indicators	Slope and landcover use	raster	Slope; landcover	geospatial common	Calculate slope from elevation dem, reclassify landcover
3	Cost surface	Calculate cost surface using different weights	Various	Various	geospatial common	Map algebra & cost surface calculation

Input Data
We will use data from the Minnesota Geospatial Common FTP server.
Table 2. Required Dataset

#	Title	Purpose in Analysis	Link to Source

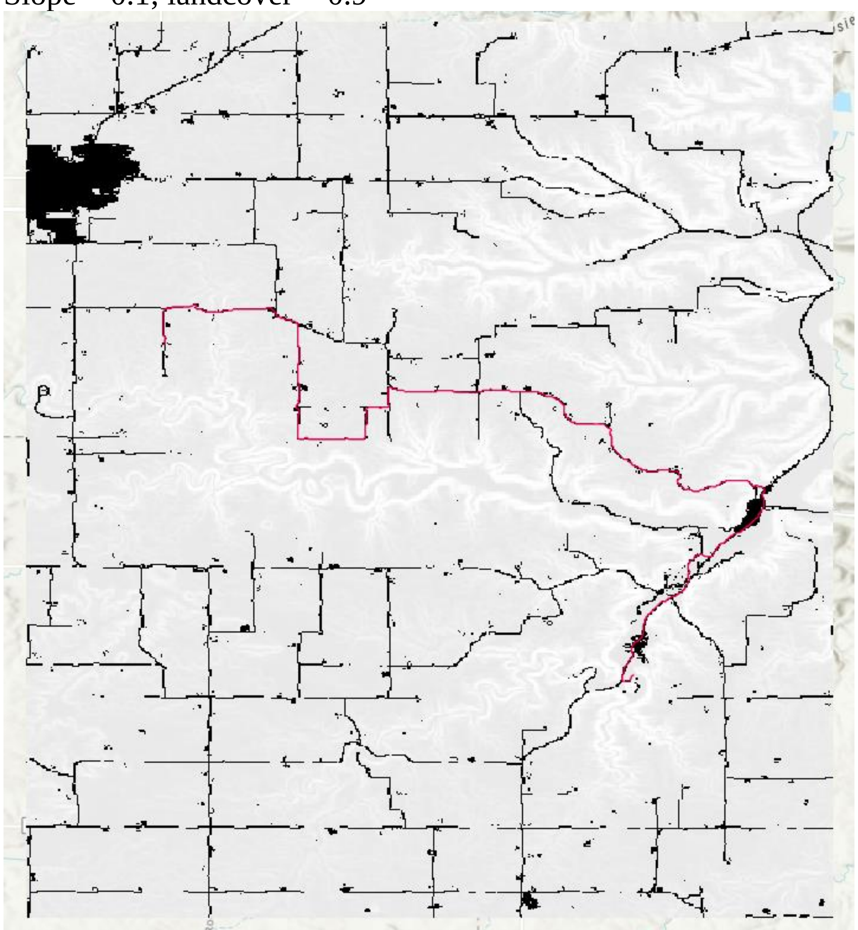
1	MN DEM elevation	Dem elevation data	https://resources.gisdata.mn.gov/pub/gdrs/data/pub/us_mn_state_dnr/elev_30m_digital_elevation_mode
2	MN landuse	Landcover use data	https://resources.gisdata.mn.gov/pub/gdrs/data/pub/us_mn_state_dnr/biota_landcover_nlcd_mn_2019

Methods

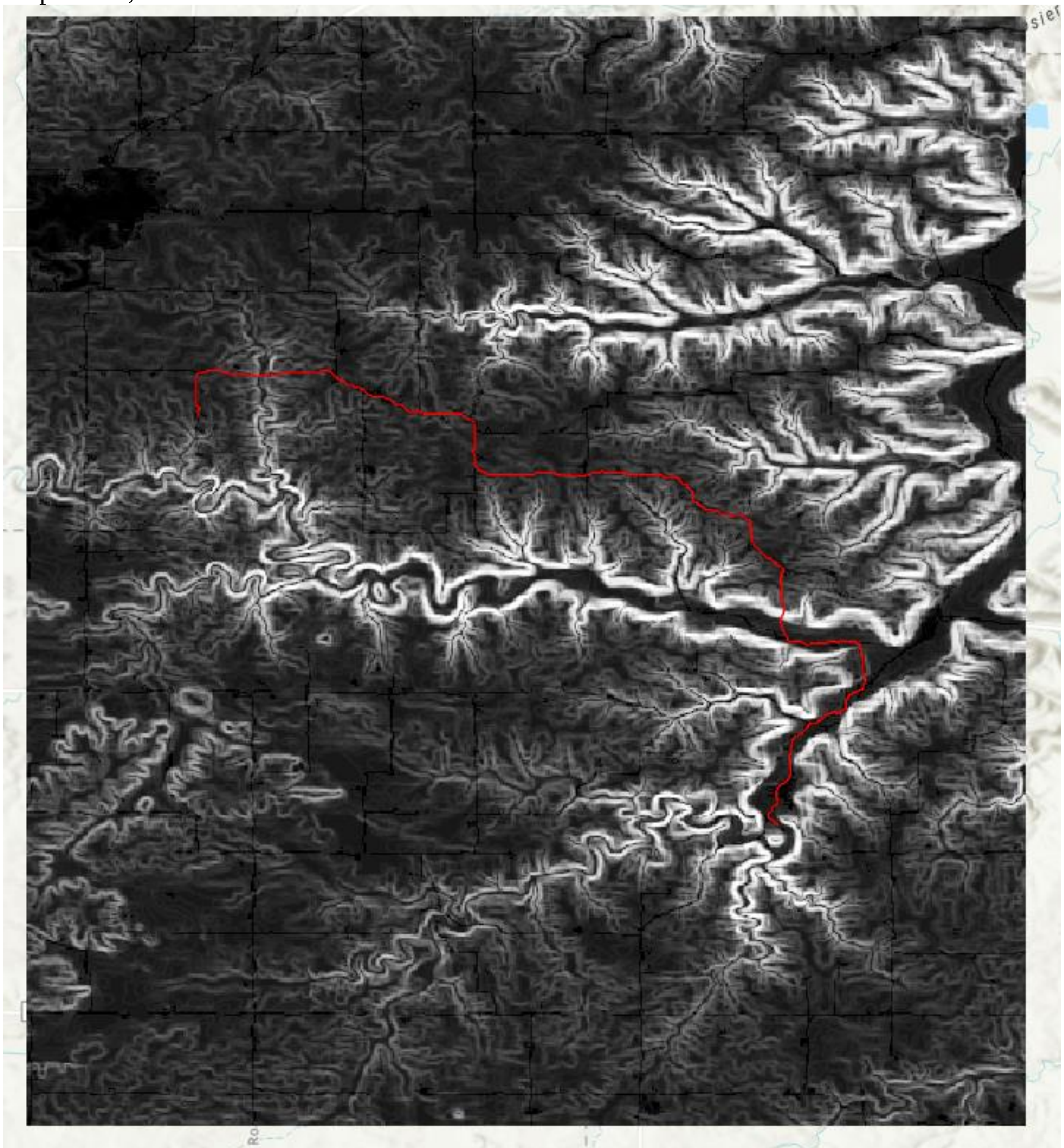
After pulling DEM and landcover use data from API, I calculated and standardized the slope and reclassified the landcover for further cost surface calculations. Next, I used map algebra to create various cost surfaces by assigning different weights of slope and land cover. Lastly, OptimalRegionConnections functions was used to get the optimal route between origin and destination.

Results

The cost surfaces and optimal routes are different if you assign different weights. For example, Slope = 0.1; landcover = 0.9



Slope = 0.9; landcover = 0.1



Results Verification
I ran the whole workflow and there was no error. Also, I visualized the data and analyzed the results to make sure the code was working properly.

Discussion and Conclusion
I have reached the goal of this lab:

- Create an ETL for data to go into a cost surface model
- Create a cost surface model and justify how you created your cost surface
- Map the range of cost surfaces given uncertain preferences and model weights

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**