Lab Report

Title Lab 3
Notice: Dr. Bryan Runck
Author: Megan Marsolek
Date: 12/1/2021

Project Repository: https://github.com/mmarsole/GIS5571
Google Drive Link: NA
Time Spent: 7hrs

Abstract

This lab concentrated on making a Cost Surface map (image/raster) and finding Optimal Paths for an imaginary person (Dory) to use to walk from their home to another location, based on select preferences. These preferences included: avoiding crop fields, walking on 'gradual slopes', and avoiding streams altogether (she would only cross a stream if there was a bridge).
Based on my calculations, I have found several paths that could be used based on varying preferences.
Although my Cost Surfaces varied in values across space, the resulting Optimal Paths were all relatively, similar. The main differences being the choice of direction during the middle of the route, depending on preference (prefer land use over slope or vice versa) the path either veers, South, East, or in between. An interesting (and annoying) conclusion I found was the need to enlarge the buffered width for roads, originally, I was trying to replicate real world observation by setting roads width to be 12ft wide (thus the width of the bridge Dory would cross), but realized I needed a much larger value so as to match the spatial resolution of my raster data (15m and 30 m).

Problem Statement

"To create a surface that shows places where Dory would more or less prefer to walk in order to get to the park, provided that (1) Dory prefers to not walk through any farm fields, (2) She also **doesn't cross water bodies**, and (3) wants the path that is the most gradual in terms of slope." (from lab 2 assignment doc)

Table 1. Desired data

#	Requirement	Defined As	(Spatial) Data	Attribute Data	Dataset	Preparation
1	Land Cover classification	Numerical values assigned to each cell within the tiff, that designates land as crops, impervious, wetlands, water, or grasslands, etc.	Tiff file	NA	Mn GeoSpatial Commons	Cropped to a smaller spatial extant
2	Elevation data	DEM, a tiff file that records the elevation in each cell	Tiff file	NA	Mn GeoSpatial Commons	Cropped to a smaller

						spatial extant
3	Stream data	Line data that provides the locations of streams/bodies of water in MN	shapefile	NA (only needed the location of the streams)	Mn GeoSpatial Commons	Cropped to a smaller spatial extant
4	Start and End points	Had to make a shapefile that designated Dory's Start and End points for her walk	Shapefile	Description (field that indicated if it was the start of end point)	NA	Made in ArcPro interface
5	Road data	Line data that provides the locations of roads (and bridges that cross streams)	Shapefile	NA (only need the location of roads)	Mn Geospatial Commons	Cropped to a Smaller extant, and buffered to 12ft wide roads

Input Data

As stated, we want to create a cost surface for where Dory will prefer to walk to get from her home to her favorite fishing spot. Based on her preferences, I have downloaded data from MN Geospatial commons (Land Cover, Elevation, Streams, and Road data), and created data (Start and End points for Dory's walk). Both the Land cover data and elevation data are tiffs, while the other three files are shapefiles. The spatial resolution of the elevation data is 30 meters while the Land cover is 15 meters. Although not the same spatial resolution, both will be sufficient to conduct a cost surface map for Dory's walk.

Before the data could be used for the cost surface, it had to be 'reclassified' or 'rescaled' depending on whether it was discrete or continuous data. This step ensures that we can directly compare the values to different data. See Table 2 for the scale parameters.

The following optimal paths rely heavily on elevation data, land cover, and Avoiding Stream Crossing. In fact, I had to alter the stream data so that it reflected the locations of bridges (by erasing stream data wherever a road intersected the stream data, this processed relied on buffering the roads to 100ft, and then using the Esri 'Erase' Tool to remove any intersections between the two shapefiles).

Table 2. Obtained Data

#	Title	Purpose in Analysis	Link to Source
1	Minnesota Land Cover Classification and Impervious Surface Area by Landsat and Lidar: 2013 update - Version 2	Reclassified to a scale of 1-10, where 10 was the highest cost (less desirable), and 1 the lowest cost (more desirable). I decided the values/scale for each land type as follows: **Urban/Developed**: (org. value) 1-100: (new value) **2** **Wetlands:** 101,102: (new value) **8** **Open Water:** 103: (new value) **10** **Extraction**: 104: (new value**) 9** **Forest**: 105,106,107: (new value) **1**	Mn GeoSpatial Commons

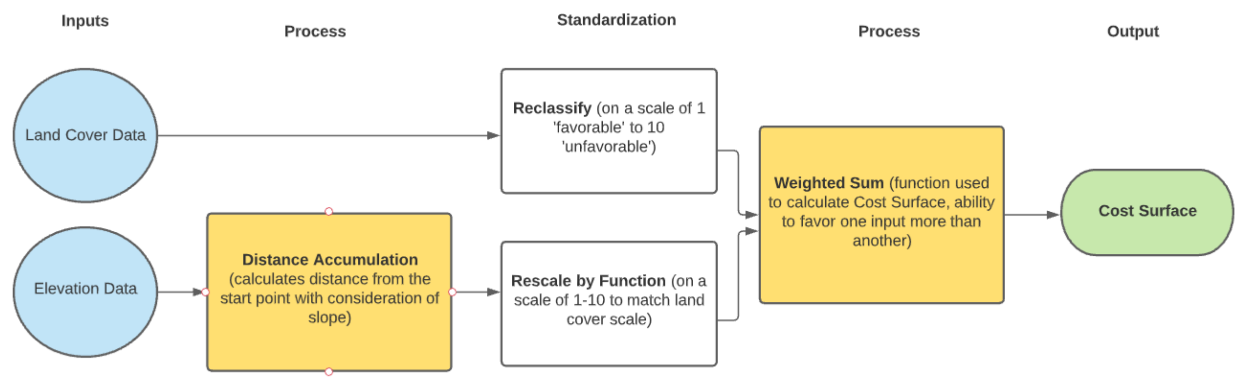
		Managed Grass/Natural Grass: 108: (new value) **1** **Agriculture:** 109,110: (new value) **9**	
2	Minnesota Digital Elevation Model - 30 Meter Resolution	Used to calculate slope. I chose that Dory would only walk on surfaces between -20 to 20 degrees (to ensure a more gradual path).	Mn GeoSpatial Commons
3	Stream Routes with Kittle Numbers and Mile Measures	Used as a 'barrier' (an impediment that Dory is not allowed to cross), but first had to alter the original stream data by erasing stream segments where there were roads (performed an 'Erase' function on the intersection of buffered roads with streams).	Mn GeoSpatial Commons
4	Roads, Minnesota, 2012	I buffered the road line data, so that all roads had a width of 12ft (an average guesstimate of road widths). I then intersected this with the stream data to identify where there was bridge crossing	MN Geospatial Commons

Methods

Once the appropriate data has been obtained, we can build a Cost Surface. As displayed in Figure 2, the cost surfaces for this lab were based upon land cover and elevation. The elevation was first processed using the "Distance Accumulation" Tool. This tool gave me the opportunity to calculate a raster whose cells' values were based on the distance from the starting and ending point and the slope. In this tool I chose to use a Binary function to calculate values for areas where slope was between -20 to 20 degrees (thus removing the possibility of walking over steep terrain). As a result, any cells where the slope was steeper, were then classified as 'No Data' values and could not be crossed later during Dory's walk. This method is less forgiving and favors Dory's preference for gradual slope. The values for the binary function could be altered to be more inclusive, or another function could be used to calculate slope preference.

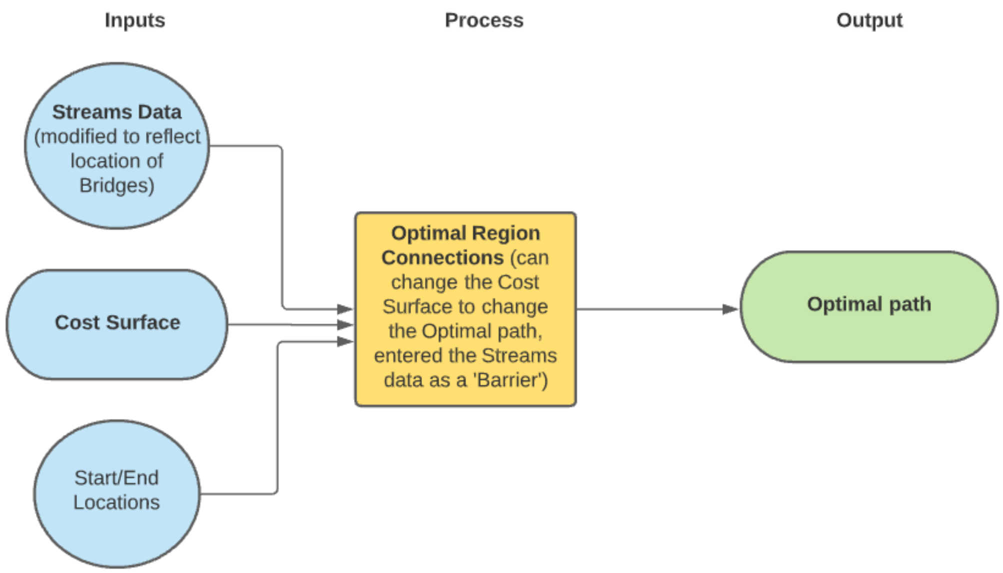
Once the raster from the "Distance Accumulation" has been calculated, this raster needs to be rescaled to values between 1 and 10, so it can be comparable with the land cover preferences. Then we can create a Cost Surface that considers all of our preferences (Land Cover reclassified and Elevation Rescaled) to calculate the walkability of surfaces for Dory. Using the "Weighted Sum" Tool, you can favor a preference over another by increasing its weight (I did three versions, one where each input's weight was equal and another where the land cover weight was higher than the elevation input, and the last which favored the elevation input over the land cover see the Results section to read more about these outcomes).

Figure 1. Creating a Cost Surface



From the Cost Surfaces we can compute Optimal Paths. These paths will vary if you change the cost surface (which are altered by either your choice of weights or inputs). In this lab I created 3 optimal paths, based on three different inputs. Using the "Optimal Regional Connections" tool you can select your start and end point data as well as your cost surface (this is what influenced the most change in the resulting Optimal paths) to calculate the best path for Dory. Additionally, all Optimal Paths used the altered Streams data as Barriers (enforcing Dory's desire to not cross a stream unless there was a bridge). See figure 2 for better understanding of the inputs for the "Optimal Regional Connections" tool. Alternatively, you can use "Optimal Path as Line" tool, but you will only be able to consider the slope as the input for your path (this method is less preferable since it will ignore our desire to avoid crop fields).

Figure 2. Creating Optimal Paths



Results

Based on the methods described I created three Cost Surface rasters. Image 1 displays a Cost Surface where the inputs elevation and land cover are equal (their weights were both 1), and thus the scale ranges from approximately 2 to 20 (since each of the inputs were scaled to values between 1-10). While Image 2 is a Cost Surface with a stronger preference for Land Cover (the spatial weight of 5 was assigned to land cover while 1 was assigned to the elevation. Image 3 displays the Cost Surface that favors the Elevation based input over the Land Cover (a spatial weight of 1 was assigned to land cover while 5 was assigned to the elevation input data). In all cases the higher the values, the less favorable the walking surface.

Based on these Cost surfaces, I calculated optimal paths (see image 4). Note that all the resulting paths are rather similar. The biggest distinction is observed in the middle of the path (image 4). It can be observed that the path that favors land cover type (see image 2) seems to follow the road more (veering East then South) while the Distance favored Path (see image 3) cuts straight through a crop field (shortening the distance to the destination).

Image 1: Cost Surface 1 (equal weights)

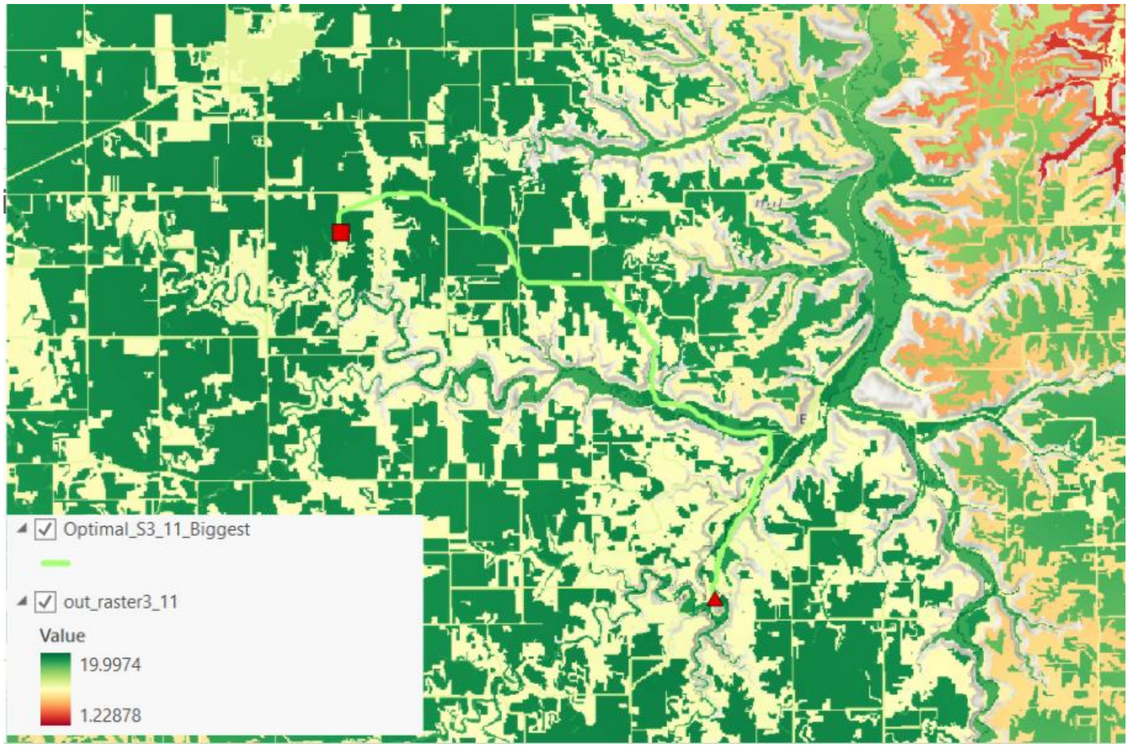


Image 2: Cost Surface 2 (favors Land cover type)

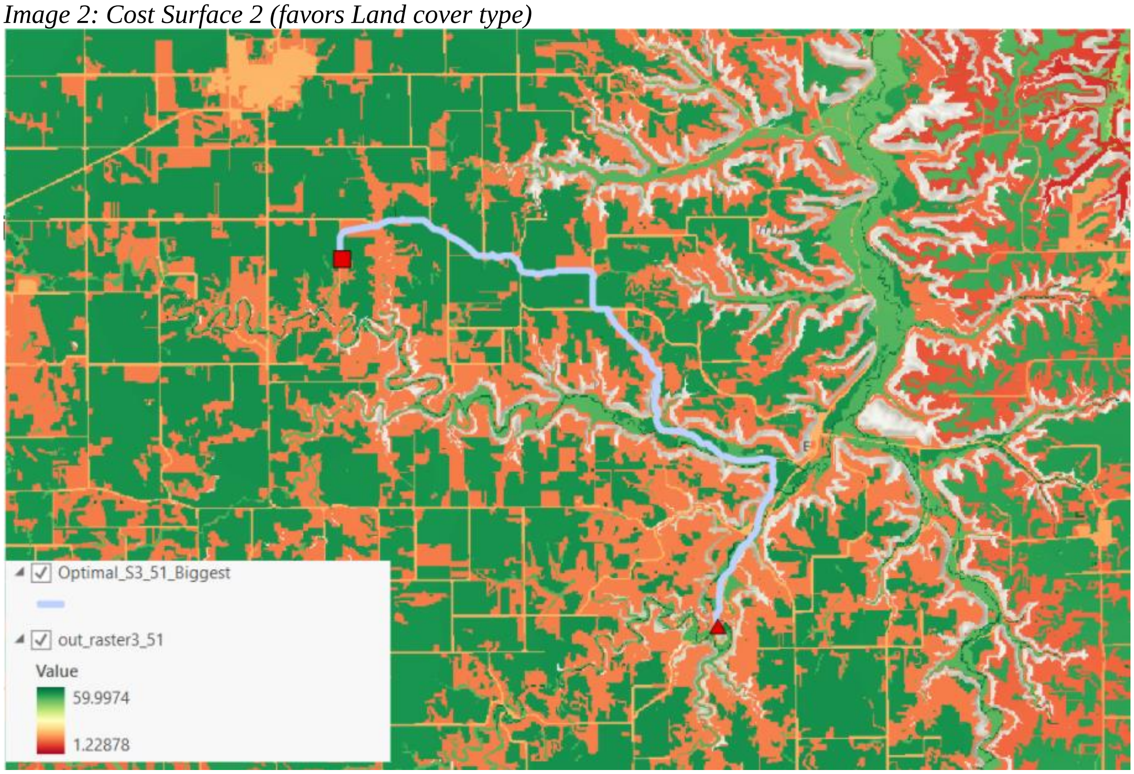


Image 3: Cost Surface 3 (favors the distance/slope from and to the End and Start points)

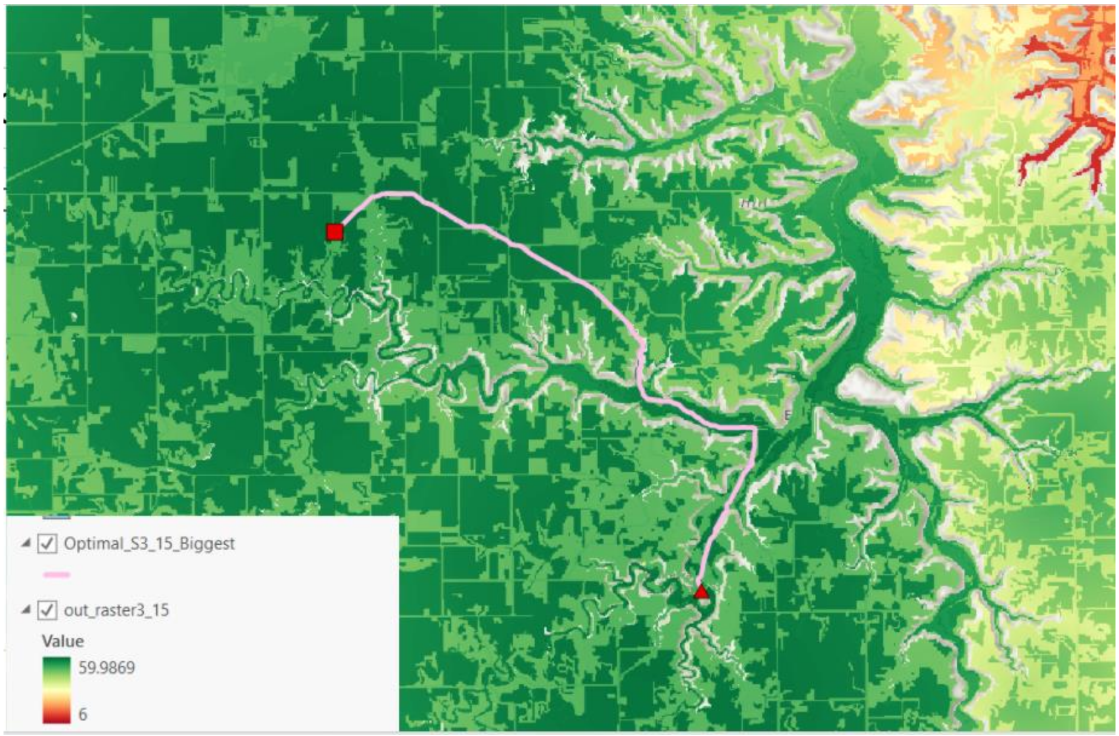
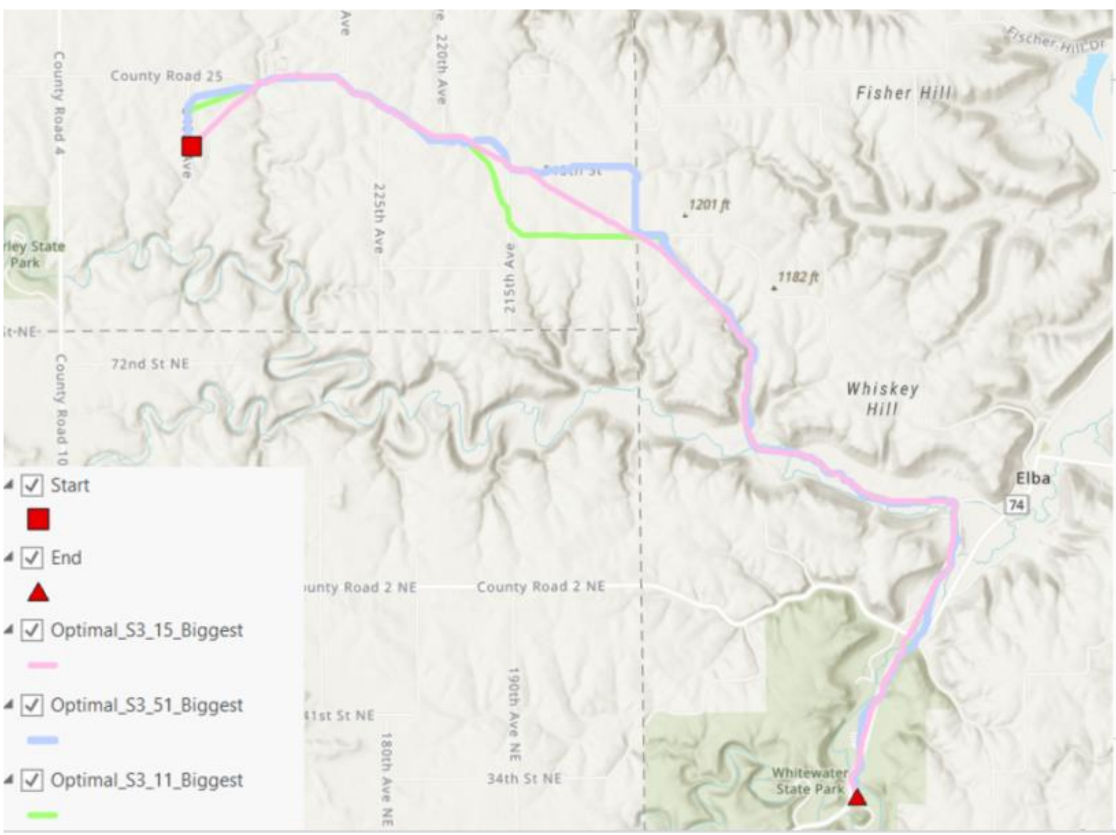


Image 3(4): Optimal Paths (see three examples based on three varying inputs)



Results Verification

These results were a result of code from my Jupyter Notebook and were compared to their GUI calculated counterparts. I can also make sense of these results through some critical thinking and deduce that seem plausible with visual inspection.

Discussion and Conclusion

The most frustrating experience in this Lab was trying to find the appropriate buffer width for the roads. Because I tried to assume a real world approximates (forcing the roads to be 12ft wide), none of my initial paths were executable (Esri insisted it was impossible to connect the Start and End destinations). I surmised this was likely due the spatial resolution of my raster-based data (15 m and 30m), Although I cannot confirm this is the case. When I finally increased the roads buffer size to 100feet the Optimal Path were calculable.

I made a few different adjustments to my Optimal Path computations since the last lab, first, I strictly enforced Dory's aversion to crossing streams without a bridge and I also increased Dory's tolerance for slope (last time I used slope between -10 and 10 degrees and doubled it to -20 and 20 degrees this time). Changing her tolerance for slope I believe gave my paths a little more flexibility than before but enforcing her use of bridges did result in less variability (as evident in the very similar optimal path outputs).

References

Distance Analysis: Identifying Optimal Paths Using Rasters. (2021, March 29). Www.esri.com; esri. https://www.esri.com/videos/watch?videoid=qO1LIFwbgDI&title=distance-analysis-identifying-optimal-paths-using-rasters

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 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	22
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	26
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	18
		100	94