Cost Surface Analysis

1. Introduction

This project aims to find out the least-cost path by using ArcGIS Pro software to determine the cost-effective path to lay the pipeline between the refineries and the oil field.

2. Data and methods used

The data used for this project are roads, pipelines, settlements, terrain, oil fields, refineries, and non-protected areas. These data were used to determine the optimal route and compared with other alternative routes. There were basic criteria that were used to develop the model. Firstly, some works were done manually using individual tools. But after that, these works were done using a model builder in ArcGIS Pro. The least-cost path analysis looks over the eight neighboring cells and finds out the least accumulated cost and defines its way to the lowest accumulated cell. The determination of the optimal route for pipeline placement should consider cost-effectiveness and least impact upon nature. Safety of nearby settlements, protected areas were considered important criteria for determining optimal pipeline routing. (Walter Collischonn, 2020) Multi-Criteria Evaluation was used to weigh and evaluate different factors necessary for deriving optimal routes. (Malczewski, 1999) The Least Cost Path Analysis was used to derive alternative paths that not compulsory to be the shortest distance but the cost-effective one. The project used the so-called cost units (CU), which refers to an area of 1 km^2 or 1 km in length, respectively.

3. Data preparation

The data used in the project are used as suggested in the exercise. Firstly, the model builder named 'modelnew' was created and in the environmental variable, the output coordinate system was set up as **WGS_1984_UTM_Zone_32N**, and the cell size was fixed to **50 m** for spatial consistency**.** It was done so because we do not have to change it frequently and these variables were set to the above values until we changed them in the model builder.

All vector files were converted to raster data except the starting point 'refineries'. It is important to convert the data to the raster because we use the pixel information in finding the least cost path. All the data in the raster model were resampled to a size of 50m * 50 m pixel. Different resampling methods were for both discrete and continuous data. The resampling method nearest neighborhood method was used for resampling the discrete data while the bilinear resampling method was in resampling the continuous data. For instance, for slope, the bilinear resampling techniques were used.

The terrain was converted to the slope and the degree measurement was used as an output because as per the criteria, these data has to be reclassified to the different cost unit (CU) according to the degree rather than the percent rise. For example, on flat terrain, the construction costs were considered with 1 CU. Each additional tenth part of a degree of terrain slope added costs of 0.3 CU. So, in this case, the value of the degree should be used for determining the cost unit. The whole description of classifying these data is discussed in the following section.

4. Reclassifying

The reclassifying techniques helped to reclassify or change the input cell values to another value. The alternative values can be manually inputted. It is also used to replace values based on new information and also helped to group certain values. It also helped to put the values to NoData cells or NoData cells to a value. (Esri, 2020)

Cost Surface Analysis

4.1 Reclassifying of Roads

The reclassification of the road took place only after the conversion of the vector feature to the raster. While converting to raster, the **level** is placed as a field. The attribute table of the vector feature was observed and the 3 levels of the road were seen, namely Level 1, Level 2, and Level 3. After converting to the raster, we gave the cost unit to each of these levels. The cost value of 0.3, 0.2, and 0.1 were inputted as 3, 2, and 1 for Level 0, Level 1, and Level 3 roads respectively. We inputted 3, 2, 1 instead of 0.3, 0.2, and 0.1 because the reclassify tool does not support the decimal values as input. After that, these rasters were multiplied by 0.1 using a raster calculator to restore the pixel to the decimal values (original value). The NoData cells were given the value of 0. The output of classified ranged from 0 to 0.3 CU.

4.2 Reclassification of settlement areas

After the conversion of the polygon feature to raster, this raster was classified with all the values equal to 5 CU and NoData cell to 0. In the beginning, there was only one classification, but after reclassification, the whole cell is divided into the settlement and non-settlement area. As it is very difficult in laying pipelines in the settlement areas and is harmful in terms of environmental and health aspects, it is better to give more emphasis to make a path to the non-settlement areas. So, the non-settlement areas were given 0 cost units.

4.3 Reclassification of terrain (slope)

The slope was classified as specified in the criteria. After the terrain was converted to the slope by using the degree as an output measurement, the new slope will have a different dataset which is different from the previous one: 13-510. The output raster (slope) has a value up to 90°. Now based on this data in degrees the cost unit was calculated as given in the exercise.

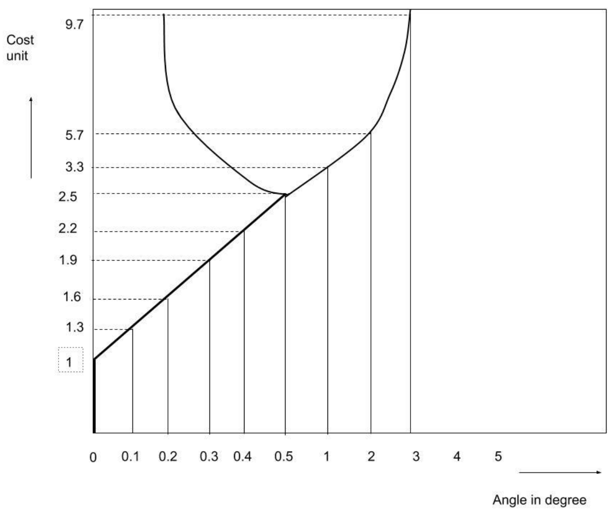

Figure 1: Cost unit on Slope

Table 1: Table showing the relationship between degree of slop and cost unit

Degree	Cost unit
0	1
0.1	1.3 (1+0.3)
0.2	1.6 (1.3+0.3)
0.3	1.9
0.4	2.2
0.5	2.5
1	3.3 $(2.5+(0.8)(1)^2)$
2	5.7 $(2.5+(0.8)(2)^2)$
3	9.7
4	15.3
5	22.5
6	31.3

The following steps were taken into consideration while doing this calculation:

1) On flat terrain, the values of degree were reclassified as 1 CU.
2) On each addition of 0.1° up to 0.5°, there is a rise of 0.3 CU in each. Equation for step 2: **1+ 3* (x)**, where x is the angle in degree
3) After 0.5 degree, the CU follows the quadratic path with the factor of 0.8 CU. Hence it is clear that as the degree of the slope increases there is increase in the cost. Hence it is advisable to use the least sloppy area for the least cost path analysis. Equation for step 3: **2.5 + 0.8* (x)**

4.4 Reclassification of the pipelines

The existing pipelines were resampled to 50 m * 50 m pixel size. There was a 70% reduction in the total cost if the new pipelines were made in the existing corridor. So, the pipelines were reclassified and given the cost unit as 3 CU. The other areas which have NoValue are reclassified as 10 CU. The reclassification tool does not support decimal values; hence these values were reduced to 0.3 and 1 CU respectively using a raster calculator.

4.5 Reclassification of the protected areas

The non-protected areas had some values while the area other than non-protected areas had no values. Hence after resampling to 50 m * 50 m pixel size, it was reclassified as a) the non-protected area was given 0 CU and b) the protected was given 1 CU. These were done because we cannot build the pipeline in the protected areas.

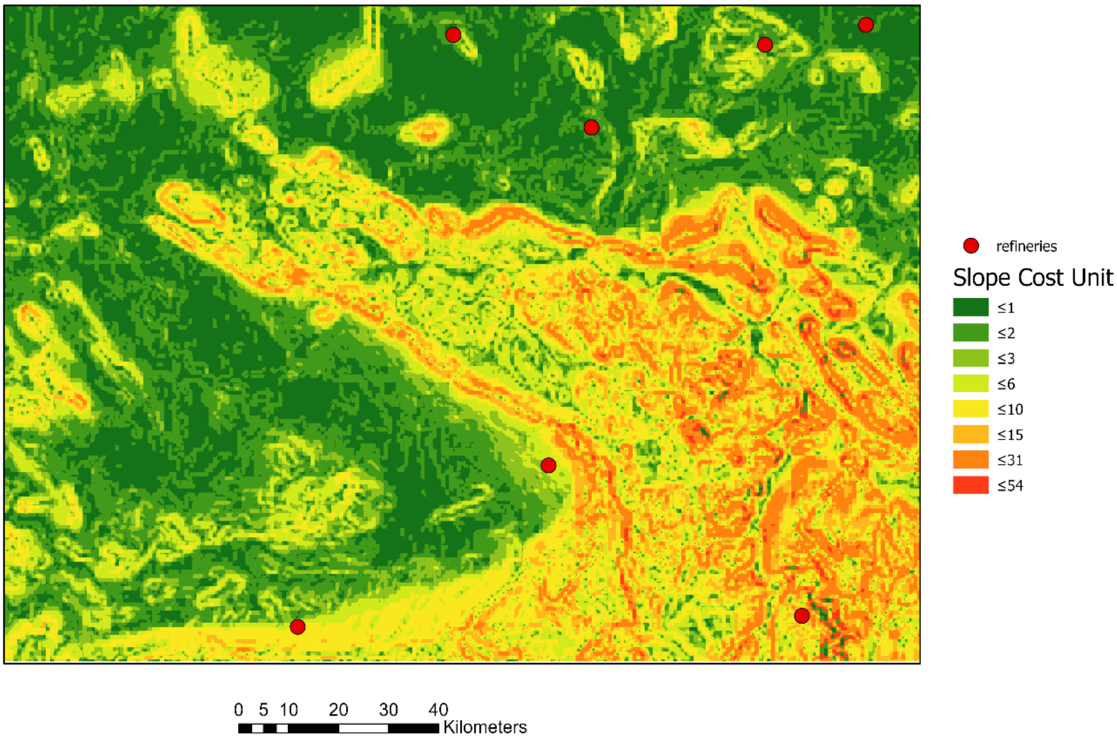

Figure 2: Map showing the Cost Unit in Slope terrain

5. Combine raster

We combined the rasters obtained from reclassifying roads, settlement areas, slopes, existing pipelines, and protected areas. For these, we consider them as equal weight. If it was not of equal weight, one can divide them in terms of their weight with the total sum of 1. For example, take two variables: slope and roads. The weight of the slope should be given more weight than the roads. So, one can give 0.7 to the slope and 0.3 to the road while doing the raster calculation. But it has already been given the cost unit to the individual variables in the reclassification. Hence the variables can be simply added to get the combined output raster. For this the following formula for the combined output were used:

Output raster = ("%settlement%" + "%slope (2)%"+"%protectedarea%"+ "%road%"+ "%pipe%")

The cost unit of the output raster ranges from 1 CU to 56 CU.

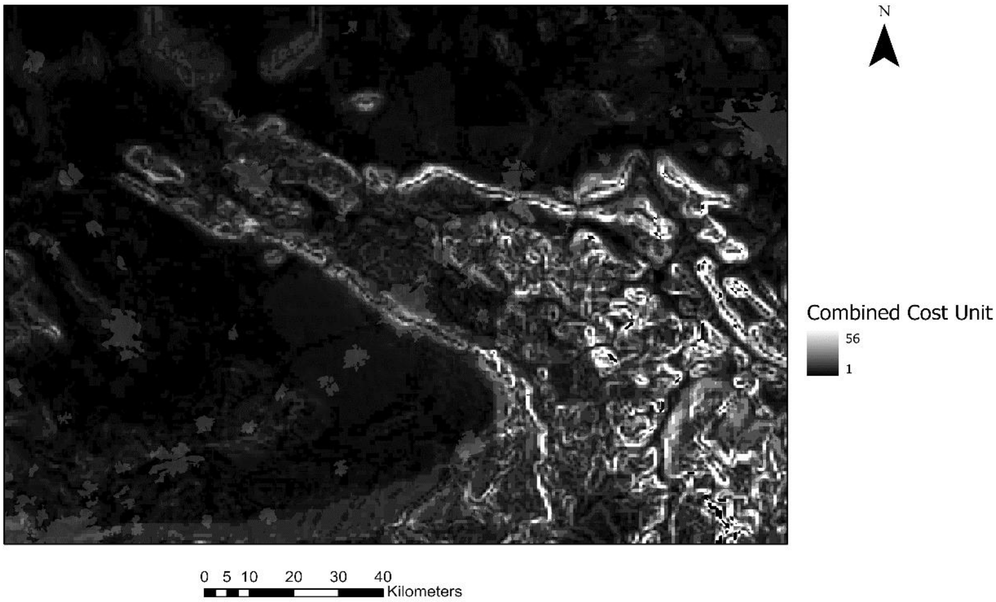

Figure 3: Map showing Cost Unit of the combined Raster

6. Path Analysis

A path analysis requires input raster source or feature source data, input cost raster, and an algorithm for deriving the least accumulative cost path (Chang, 2007), which is already developed in the toolbox in ArcGIS Pro. The input source was refineries and the input cost raster was the combined raster that had been developed before.

6.1 Source Raster or feature source data

A source raster defines the source cell(s). There were 8 refineries which were point values. Only the source cell has a cell value in the source raster, all other cells are assigned no data. Cost distance measures spread from the source cell. But in the context of path analysis, one can consider the source cell as an endpoint of the path or the origin. A path analysis derives the least cost path to the source cell, or the closest source cell if two or more source cells are present. (Chang, 2007)

6.2 Cost Raster

A cost raster is a cost to move through each cell. A site analysis of a pipeline project must consider the construction and operational costs. Some of the cost includes distance from the source to destination, topography, roads, settlement, and protected areas. The cost may be relative in the pipeline project. Relative costs are ranked values. For example, costs may be ranked from 1 to 5, with 5 being the highest cost value. As it involves a wide variety of cost factors, some factors can be measured in actual costs but others such as wildlife habitats, and protected areas are difficult to measure in actual costs. Relative costs are therefore a means of standardizing different cost factors. (Chang, 2007)

6.3 Cost distance measures

The cost distance measure in a path analysis is based on the node-link cell representation (figure 4). A node represents the center of a cell, and a link-either a lateral link or a diagonal link- connects the node to its adjacent cells. A lateral link connects to a cell to one of its four immediate neighbors, and a diagonal link connects the cell to one of the corner neighbors. The distance is 1.0 cell for a lateral link and 1.414 cells for a diagonal link. The cost distance to travel from one cell to another through a lateral link is 1.0 cell times the average of the two cost values:

$$1 * [(C_i + C_j)/2]$$

Where C_i is the cost value at cell i and C_j the cost value at neighboring cell j. The cost distance to travel from one cell to another through a diagonal link, on the other hand, is **1.414** cells times the average of the two cost values.

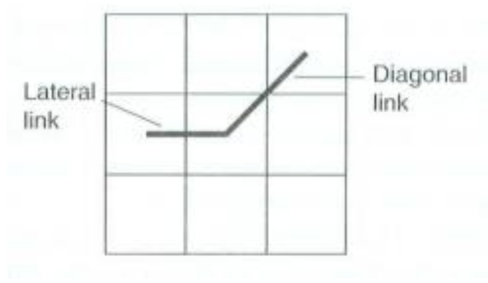

Figure 4: Cost distance measures follow the node-link cell representation: a lateral link connects two direct neighbors, and a diagonal link connects two diagonal neighbors. (Chang, 2007)

6.4 Cost Backlink

The cost back link convert the input source data (i.e. refineries) to raster internally before the analysis. The resolution is set up as same as input cost raster. The backlink raster contains value from 0 to 8, which defines the direction of the next neighboring cell along the least cost path from a cell to the source. Backlink raster is required to create the cost-path to determine the route to return to a source via the most-effective route. (Esri, 2020)

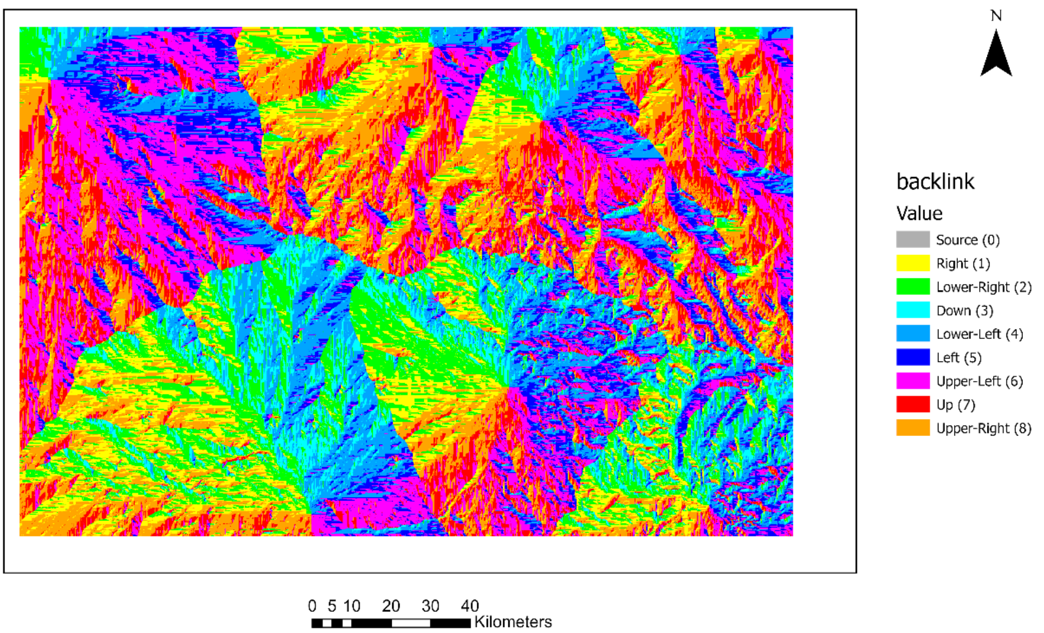

Figure 5: Backlink Raster showing the direction of the next neighboring cell

6.5 Deriving the Least Accumulative Cost Path

If we have a cost raster, we can calculate the accumulative cost between two cells by summing the costs associated with each link that connects the two cells. There can be many different paths that can connect two cells that are not immediate neighbors. The least accumulative cost path can only be derived after each possible path has been evaluated. (Chang, 2007)

7. Result

A cost path tool was used to give a cost-effective cost path. It used the Oilfield raster as an input raster. The cost distance raster (computed in the operation) is inputted in the input cost distance raster and use the backlink raster as input cost backlink raster. After these values were inputted, the cost path operation was operated. The results are in the raster form. The cells of the least-cost path were converted to vector from raster using a tool *raster to polyline* for better visualization. From the result, the most cost-effective routes for pipelines from the oil field to refineries were found out.

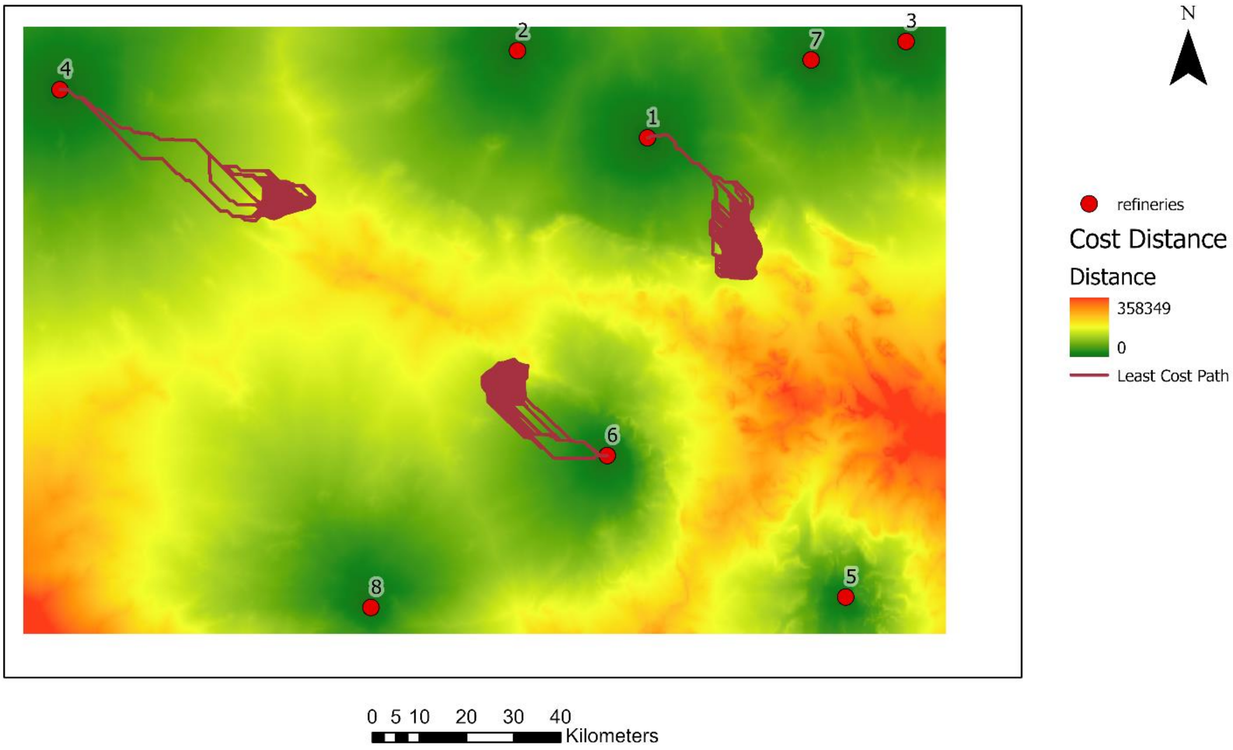

Figure 6: Map showing the least cost distance and Least cost Path

References

Chang, K.-t. (2007). *Introduction to Geographic Information Systems.* Tata McGraw-Hill Education Pvt. Ltd.

Esri. (2020). *https://pro.arcgis.com/*. Retrieved from https://pro.arcgis.com/: https://pro.arcgis.com/en/pro-app/tool-reference/spatial-analyst/how-the-cost-distance-tools-work.htm

Malczewski, J. (1999). *GIS and Multicriteria Decision Analysis.* New York: Wiley.

Walter Collischonn, J. P. (2020). A direction dependent least-cost-path algorithm for roads and canals. *International Journal of Geographical Information Science*.

Cost Surface Analysis

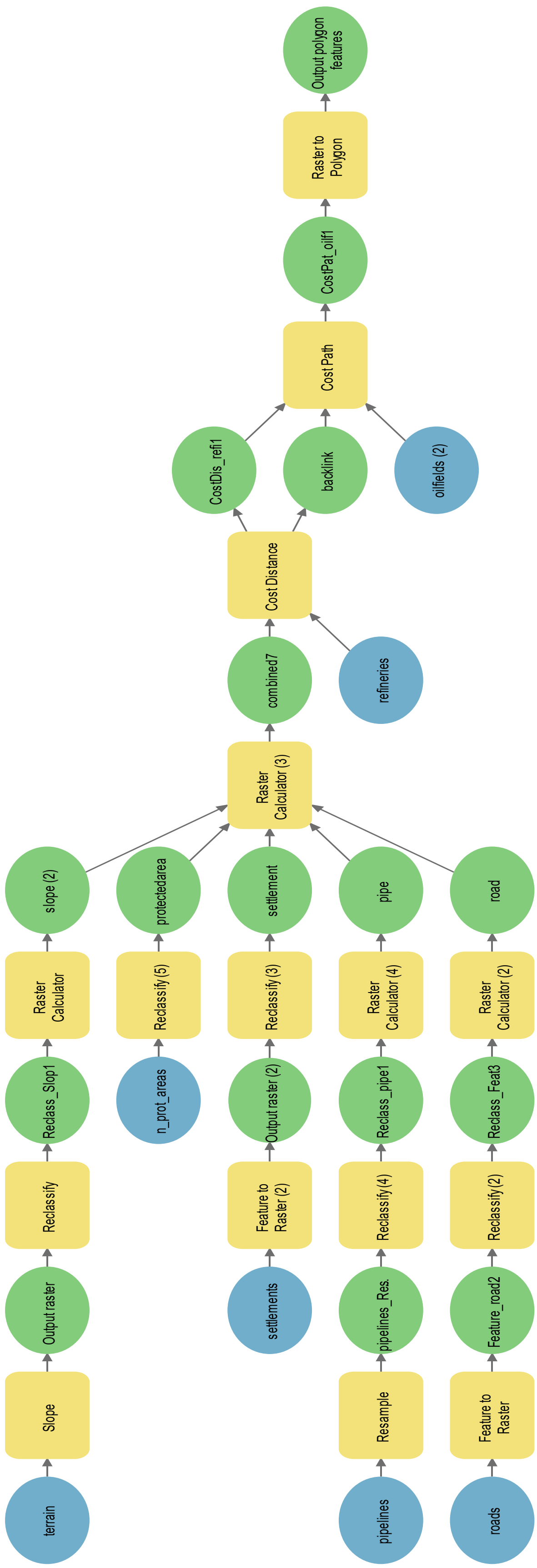

Figure 7: Model Builder for Cost Path Analysis