Raster Analysis

Binary Overlays

A. Table of Contents

I. Basics of Raster Data Modeling 2

II. Introduction 5

III. Understanding the Given Datasets 6

IV. Data Preparation 7

- A. Land Cover 7
  - A.1 Extracting the Water Bodies 8
    - A.1.1 Preparing Water Suitability Layer 10
  - A.2 Extracting the Urban Areas 13
    - A.2.1 Preparing Urban Suitability Layer 14
  - A.3 Extracting Forest Suitability 14
- B. Preparing the Road Suitability 15
- C. Preparing the Slope Suitability 17

V. Analysis 18

- a. Combine Tool 18
- b. Map Algebra: Raster Calculator 19

VI. Assessing the Results 22

VII. Report the findings 23

VIII. Your Turn 25

- a. Report Content 25
- b. Grading Rubric 26

IX. Readings 27

I. Basics of Raster Data Modeling

Raster analysis works on the basis of the **cell** (pixel), hence the term "cell-based analysis". So the cell size is a crucial component in the process. Many of the operations are based on cell by cell comparison, and others would be considering the neighboring cells (Neighborhood Functions), or group of contiguous cells (zones and regions) (Zonal and Region Functions). Review the main concept of cell size and how it can affect the amount of data available in a raster data set here:

http://resources.arcgis.com/en/help/main/10.1/index.html#//009t00000004000000

Similar to the vector data model, raster datasets that represent classified or grouped data with a defined number of values also have a "Value Attribute Table". It is much simpler than the vector data model since it contains essentially 3 fields: OBJECTID, VALUE, and COUNT. The following figure shows a sample raster dataset on the left and its corresponding attribute table on the right. The OBJECTID field is a default internal field that is associated with each record; the VALUE field holds the entries for each unique value in the raster dataset, while the COUNT Field maintains the cell count of each unique cell value. The VALUE field is especially useful in order to identify areas of specific interest in the area under investigation while the COUNT is essential to quantify the cells in each category or value; knowing the cell size (e.g. 30 Meters x 30 Meters = 900 Square Meters), we can calculate the area each value occupies in a raster dataset. For example, for value "1" in the following illustration, it has 9 cells, assuming that the cell size is 900 Square Meters, then the area that value 1 occupies in this raster dataset is 9 x 900 = 8,100 Square Meters.

1	1	1	3	3
2	1	1	3	3
2	2	2	3	3
3	1	1	4	4
3	1	1	2	2

OBJECTID	VALUE	COUNT
0	1	9
1	2	6
2	3	8
3	4	2

A raster dataset is nothing but an array of numbers stored in a series of cells organized in rows and columns. Each cell has a location based on the first cell on the top left corner of the raster data set. And since the raster data set's main focus is not the feature (point, line or polygon) but rather a location (cell area), representation of features can be accomplished by using single

detached pixels for point data, while it can use zones and regions to represent linear and polygonal ones.

A "**Zone**" in the raster data set is composed on 2 or more cells of the same value, they can be adjacent, disconnected or both. The zones that are contiguous and are connected can represent area features such as a patch of trees, a water body, a building …etc. The following figure illustrates each cell value in a raster dataset and the corresponding zones.

1	1	1	3	3
2	1	1	3	3
2	2	2	3	3
3	1	1	4	4
3	1	1	2	2

Zone with value 1
Zone with value 2
Zone with value 3
Zone with value 4

Each group of contiguous cells that have the same value are considered a "**Region**". For example, in the following illustration, the zone that has a value of 1 is composed of 2 regions, while the zone that has a value of 4 is composed of only one region.

1	1	1	3	3
2	1	1	3	3
2	2	2	3	3
3	1	1	4	4
3	1	1	2	2

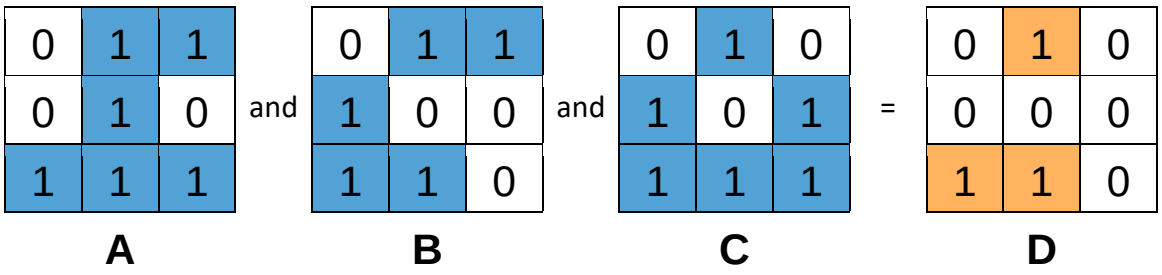
Region 1
Region 2
Region 3
Region 4
Region 5
Region 6
Region 7

Similar to vector, within the cell-based analysis you can derive new information from existing data (e.g. slopes from elevation (DEM)), combine layers to find suitable sites, distance and travel cost analysis, statistically analyze different information, and interpolation. All of those are accomplished within the realm of the cell, or group of cells, in your raster layer(s). Therefore, the concentration of the raster model is on the cell itself, which is a representation of a squared

(mostly) area of the space and not a specific feature (point, line or polygon) in contrast with the vector model.

The simplest form of raster suitability analysis is the binary site selection method (Boolean), which is one of the multiple ways of identifying the best locations for a specific activity or land use (suitability). The result will contain values of 1 and 0, where 1 is suitable and 0 is not.

The following figure illustrates the cell-by-cell based selection of a typical raster analysis procedure. It portrays three different data sets (A, B and C) each contains suitability information (0-Not suitable, and 1-Suitable), a conditional statement would reveal the output where each cell is assigned its suitability based on the values of the same cell in every input.



An example of the binary site selection where A, B, and C are all raster datasets where all the cells that are suitable have a value of 1. For example, if A is representing the slope suitability for, let's say, agriculture, all the cells with 1 are slopes that are suitable while the ones with 0 are not, similarly, if B is the suitability by soil type and C is the suitable direction of the terrain (aspect); we'd need to identify those cells that are suitable in all three inputs to identify the output as suitable.

For example, the first cell at the top left corner of all inputs has a value of "0", the output will be not suitable, since all of the inputs are not, while the following cell to the right is "1" in all of the inputs and therefore it'll be classified as "suitable" or "1". On the other hand, other cells, such as the one in the center is "1" (suitable) in the first input (A) and "0" (not suitable) in the rest (B and C), it'll be classified as "not suitable" or "0", since it is not suitable in one of the inputs. In other words, the output will be assigned a value of "1" (suitable) only if all corresponding input cells are suitable "1".

Although binary or Boolean suitability methods (1-Suitable or 0-Not Suitable) are best applied with vector analysis but can be used with raster analysis too. On the other hand, those suitability models that are weighted or ranked and have "fuzzy logic" operations associated with them are best suited for raster modeling and will be discussed later in this course.

Raster Analysis

Binary Overlays

SUMMARY

This exercise is to demonstrate how the raster model can be applied in identifying the areas that are best suited for natural wildlife habitats using a simple binary overlay method. In this assignment we investigate the areas in Clarion County, PA that could be considered as best suited for a general natural wildlife habitat. It is assumed that those areas would be at a close vicinity of water bodies while maintaining a distance from major roads and human environment. Natural forest or open land cover with a relatively flat slope would be the best.

II. Introduction

Download the exercise data from the corresponding link.

Most of the raster data handling and analysis will rely on the use of a software extension (i.e. plug in) called Spatial Analyst. It comes bundled with ArcGIS Pro, but the proper licensing should be acquired/assigned.

Raster analysis is different than vector in a way that it might involve many preparation steps and that their output might need tweaking in order to reveal or quantify the results. To be more specific, the concentration of the raster analysis model, unlike the vector methods, is the space more than the exact location or the individual feature (point, line or polygon). It definitely can be applied to study specific features (e.g. parcels), but it will not be as efficient or straightforward as the vector model.

This exercise is to demonstrate how the raster model can be applied in identifying the areas that are best suited for natural wildlife habitats. It is intended as a guideline of GIS analysis and modeling strategy and results should not be adopted without further investigation and research.

In this assignment we would like to investigate the areas in Clarion County, PA that could be considered as best suited for a general natural wildlife habitat. It is assumed that those areas would be at a close vicinity of water bodies while maintaining a distance from major roads and human environment. Natural forest or open land cover with a relatively flat slope would be the best.

The suitable sites would generally match the following criteria:

1. They should be located within 800 meters from water bodies;
2. Generally, they would be about 2,000 meters away from major roads and 1,600 meters away from Urban/Built-up areas;
3. The land should be covered by open land or forested areas; and,

4. The landscape should be relatively flat (0-8% slope).

III. Understanding the Given Datasets

From the Catalog pane in ArcGIS Pro, you're given a Geodatabase that contains multiple data layers in vector and raster format. Before proceeding, previewing the given data and making sense of them is essential. Also, using notepad or similar application, open the given **palulc_05_readme.txt** text file that include a detailed explanation on the land cover data classification system.

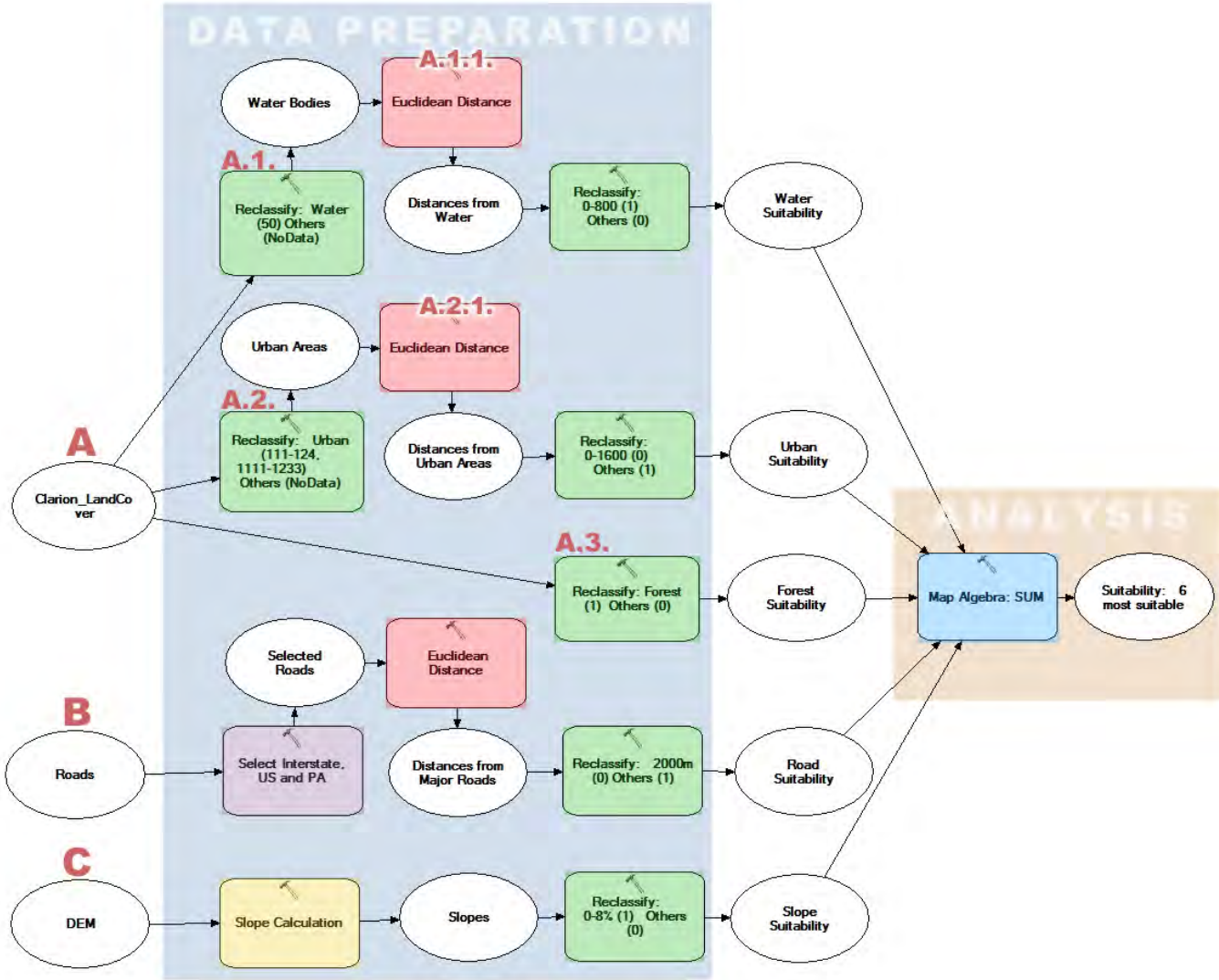
You'll notice that in raster modeling in general, many data will have to be derived from other datasets. Many data preparation steps will be involved in order to ensure the integrity of all of the data used in the analysis process. The following table summarizes the different criteria, what type of source layer it would be coming from, and what type of attributes and attribute values are needed in order to perform this analysis.

	Criteria	Source Layer	Derived Data	Attributes	Attribute Value
1.	800 meters from water bodies	Clarion_LandCover	Distance from Water	VALUE	50
2.	1600 meters away from Urban/Built-up	Clarion_LandCover	Distance from Urban Areas	VALUE	111-124 (urban) 1111-1233 (urban with mixed tree cover)
3.	open or forested land cover	Clarion_LandCover	Forest Suitability	VALUE	24 (pasture/grassland) 41-43 (forest)
4.	2000 meters away from major roads	Roads	Distance from Major Roads	ROAD_TYPE	Interstate, PA, US
5.	0-8% slope	Digital Elevation Model (DEM)	Slopes	VALUE	0-8%

As you notice in this table, three of the input source layers are derivatives of the 'Land Cover' raster data (1, 2, and 3), one is computed from exiting subset of the vector source layer 'roads' (4), and the last is a derivative of a 'Digital Elevation Model (DEM)' which is the slope (5).

In this exercise you'll learn how to prepare this data from the original source layer as well as performing the analysis. The preparation tasks will involve the extraction of a specific set of land cover classes from a larger dataset, reclassify its values (aggregate different attribute values to one class) to suitable and non-suitable cells (1 & 0), creating a road distance raster surface layer out of the selected subset of roads, reclassify them to suitable distances and non-suitable ones, then finally derive a slope suitability layer from a DEM.

Following is a brief descriptive flowchart that depicts the preparation as well as the analysis tasks and order for this assignment:



General diagram of all Data Preparation steps (blue background) and Analysis (light brown background). You'll notice that many of the tasks in raster modeling will involve data preparation in order to achieve an easy analysis method. Reclassifying raster data is a common task while calculating distances, slopes or using the Map Algebra for different input calculations are also common.
This model flowchart can be dissected to the 3 major inputs (A, B and C), 4 initial processes (A.1., A.2., B.1., and C.1.), that lead to the 5 main data outputs (1-5) leading to the final analysis calculation.

IV. Data Preparation

Following the "Data Preparation" part (light blue background) from the above diagram, 3 major inputs (A, B and C) will be used to generated the suitability raster layer for each of the given criteria.

A. Land Cover

The landcover dataset contains several classes that can be extracted based on the project requirements. By opening the associated text file **palulc_05_readme.txt,** you can conclude that the water bodies, urban areas, and forested land each has one or more classification code. In the following subsections, you're going to work to extract each of those land covers into its

separate layer and reclassify them into a binary (0 and 1) representation for suitability purposes (0 = unsuitable and 1= suitable).

The land cover dataset will be used to extract three different layers for the analysis using the Reclassify tool (Analysis tab >> Geoprocessing group >> Tools, then in the Geoprocessing pane, under Toolboxes tab >> Spatial Analyst >> Reclass >> Reclassify, you may also search for "reclassify" in the search field in order to locate the tool) to isolate the Water, Urban areas, and Forest in three separate raster datasets. A distance layer is going to be generated for the Water and Urban areas, then the same tool is going to be used repetitively to reclassify all of the input raster dataset cells to either 1-Suitable or 0-Not Suitable for the Water and Urban areas.

Before proceeding with the extraction steps, we would like to setup the geoprocessing environment to limit any process to the extents of Clarion County. This practice would not only reduce the amount of processed data in every step but would also clean up the outputs instead of clipping the data for every output to the County's borders. Clipping and cleaning up would occur but possibly at the end of the analysis stage in order to better present the results.

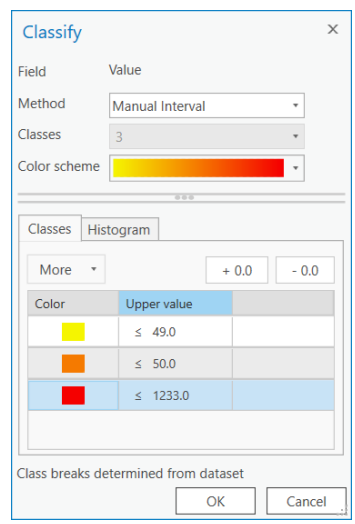
Clarion County processing extent (Envelope) showing in red

The Geoprocessing Environment can be set for each individual process or can be set globally for all subsequent process. Here we are going to set it globally so that we do not have to worry about it during this assignment.

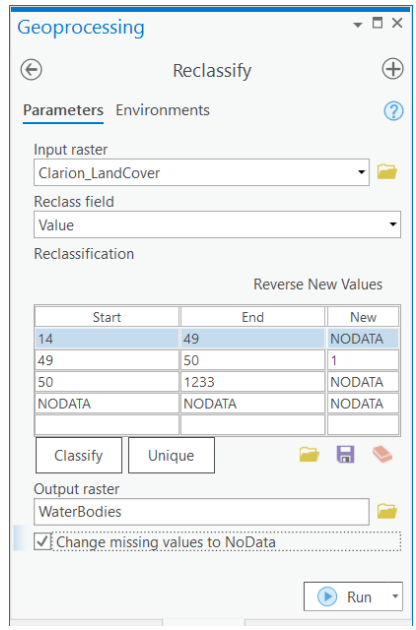
From the Analysis tab >> Geoprocessing group, click on Environments. In the Environments dialogue, extend the "Processing Extent" and change the Extent to the ClarionCounty feature class in the given geodatabase (or the layer if you added it to the Contents). This will limit any process output to the rectangular boundary of Clarion County borders (also called feature "Envelope").

A.1 Extracting the Water Bodies

- Run the Reclassify tool (Analysis tab >> Geoprocessing group >> Tools search for "reclassify". Either one of the top two results should work: Reclassify (Spatial Analyst Tools) or Reclassify (3D Analyst Tools)).
- Use the Clarion_LandCover as input raster;
- The Reclass field should be VALUE.
- Then, for the Reclassification, click "Classify" and change the number of classes to 3. Change the Break values to 49, 50 and whatever is left over (maximum in dataset is 1233), then click OK.



- In the Reclassification area, change the values of the "New values" to the following:

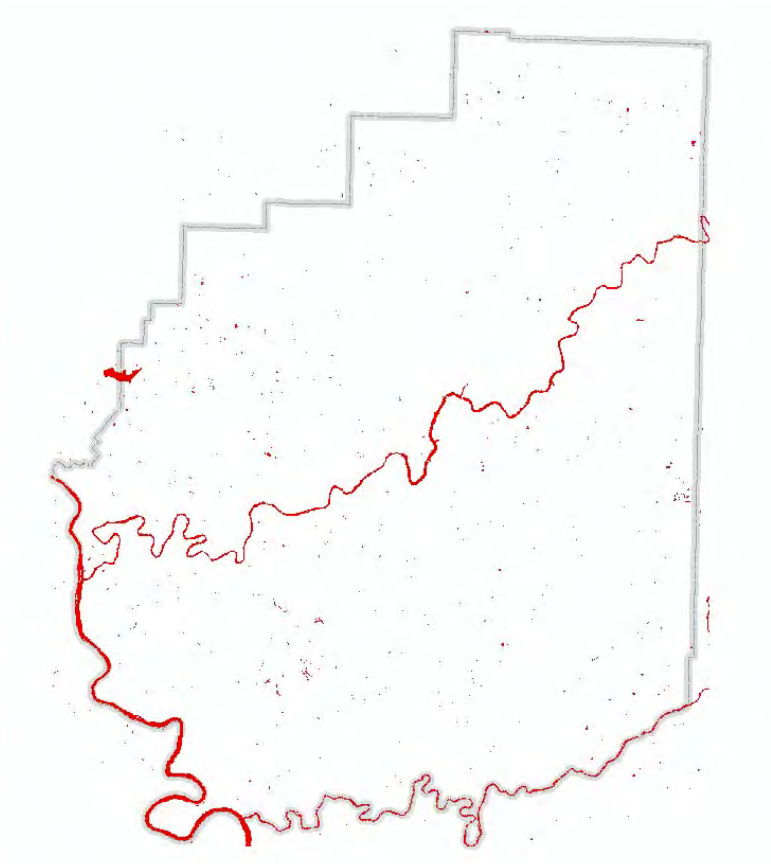


Old values	New values	Notes
14-49	NODATA	All caps
49-50	1	50 is Water bodies
50-1233	NODATA	All caps
NODATA	NODATA	All caps

Notice that we are setting all the values below 50 (Water bodies) and above it to "NODATA". NODATA is the way to go for not available or missing information in raster analysis. A value of "0" does not mean the absence of data, but that the data value is 0 instead. In a suitability model, for example, a "0" would mean "not suitable" while "NODATA" would mean that the data is not available, or those cells are background cells that should not be considered in the analysis.

- Save the Output raster into the designated assignment geodatabase and rename it to ***WaterBodies***;
- Check the "*Change missing values to NoData*" box;
- Click *Run*.

The resulting raster should be similar to the following figure:



For illustration purposes, the Clarion County boundary lines are added to the figure and placed underneath the result. The extracted water bodies will be scattered cells all over the County as well as the clusters of rivers and lakes. Since we already set our geoprocessing environment to the extents of Clarion County, only those cells that fall within the County's "Envelope" are output.

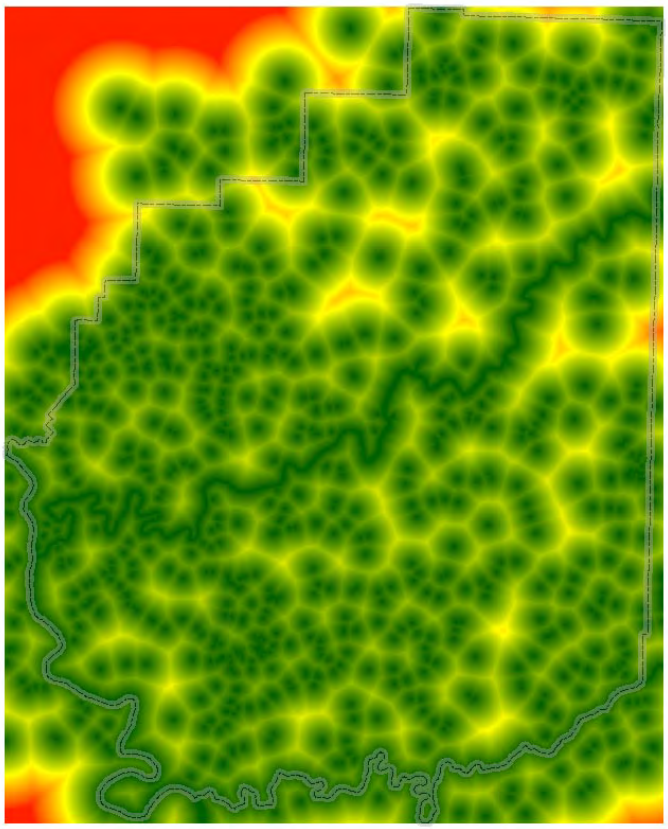
A.1.1 Preparing Water Suitability Layer

Now we'll prepare the suitability layers for the Water layer. We are going to first create a distance layer, then, using the same previously applied reclassification procedures, we are going to reclassify the distances to either 1-Suitable or 0-Not Suitable depending on the given criteria (800 meters).

The Euclidean Distance tool under Analysis >> Geoprocessing >> Tools, search for Euclidean Distance (it is located under the Legacy group under Spatial Analyst tools >> Distance). The tool generates a continuous raster layer of distances from the existing features and beyond within the defined study area extents (geoprocessing environment).

Use the WaterBodies layer in the table of contents as the Input raster or feature source data, save the Output distance raster as **WaterDistance** inside the assignment geodatabase. You'll notice that the Output cell size is "28.987721" which is the same as the original input layer of Clarion_LandCover. Click Run to create the distance raster layer.

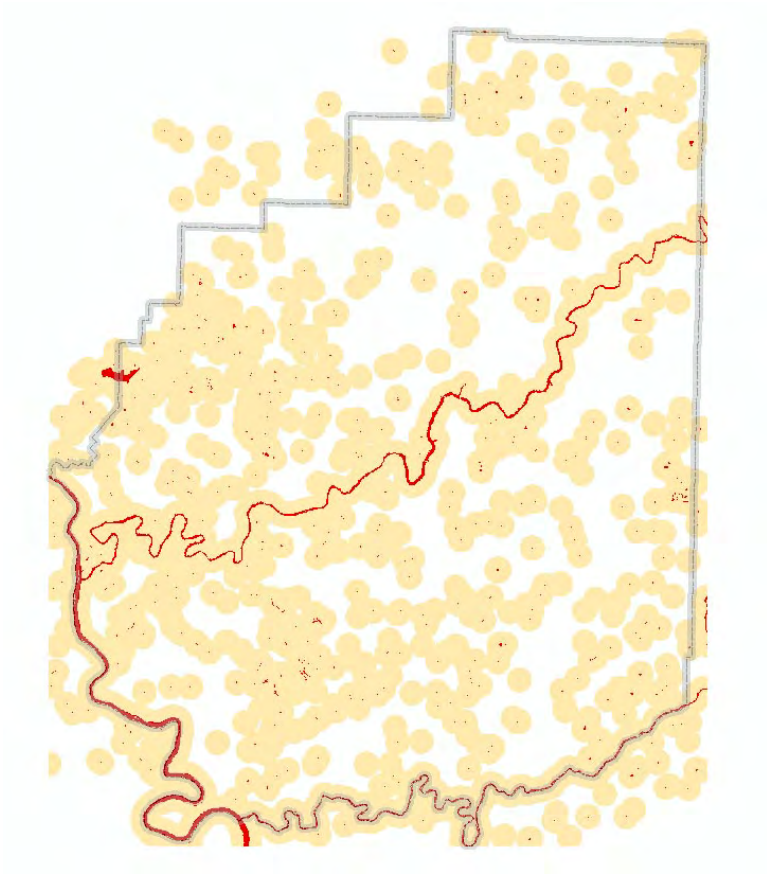
You may change the symbology of the WaterDistance layer in the table of contents to show a stretched symbols (green to red) to view all of the created data, also you can change the symbols to show only those values that are 0-800 meters.



Distance layer calculated around the water bodies, all distances are presented in this figure. Darker greens are closer to the extracted water bodies and red colors are further away.

Run the Reclassification tool in order to produce a raster layer with 1-Suitable for the 0-800 meters and 0-Not Suitable for all others. Call the produced raster layer **Suitability_Water**.

The resulting Suitability_Water raster layer should be similar to the following figure:

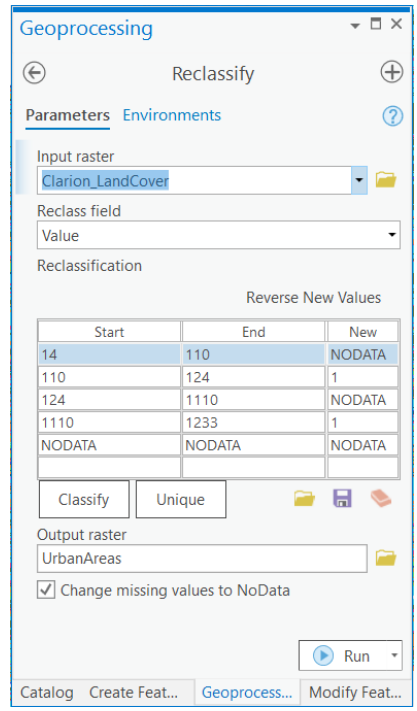


Only the 800 meters distance (beige) around the existing water bodies (red) are presented in this layer. This would be exactly the expected output from the reclassification of the continuous distance layer.

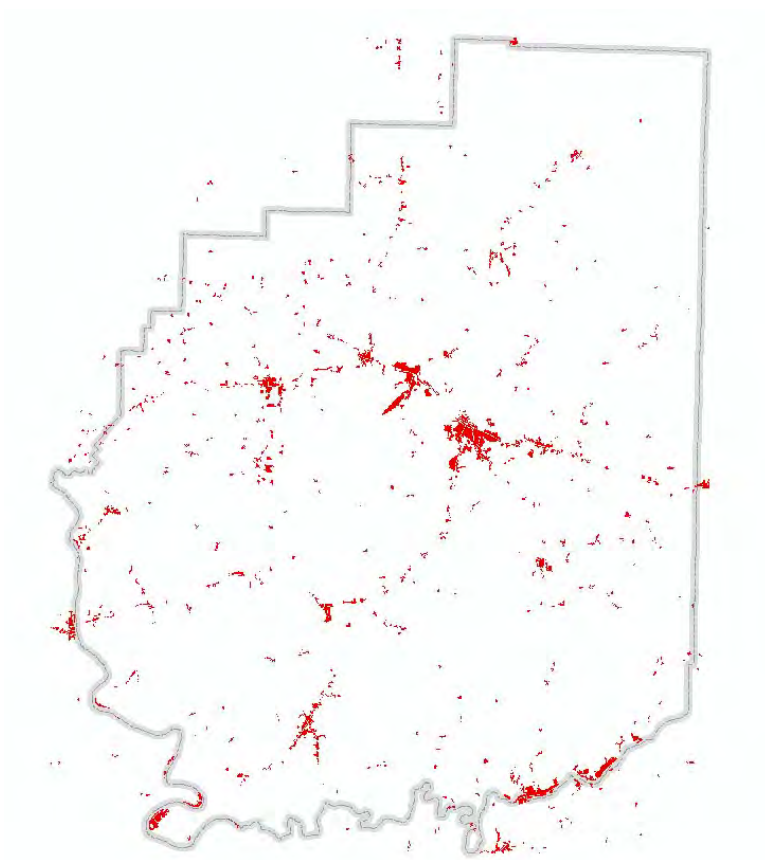
A.2 Extracting the Urban Areas

Repeat the previous steps but with the following Reclassification schema (four classes this time) and rename the output to **UrbanAreas**:

Old values	New values	Notes
14-110	NODATA	
110-124	1	111-124 are classified as Urban land: Residential, Institutional, Industrial, Commercial, and Airports.
124-1110	NODATA	We do not need this class since 241 are golf course and 750 is Active Mines
1110-1233	1	Remaining sub-classes of the Urban areas
NODATA	NODATA	



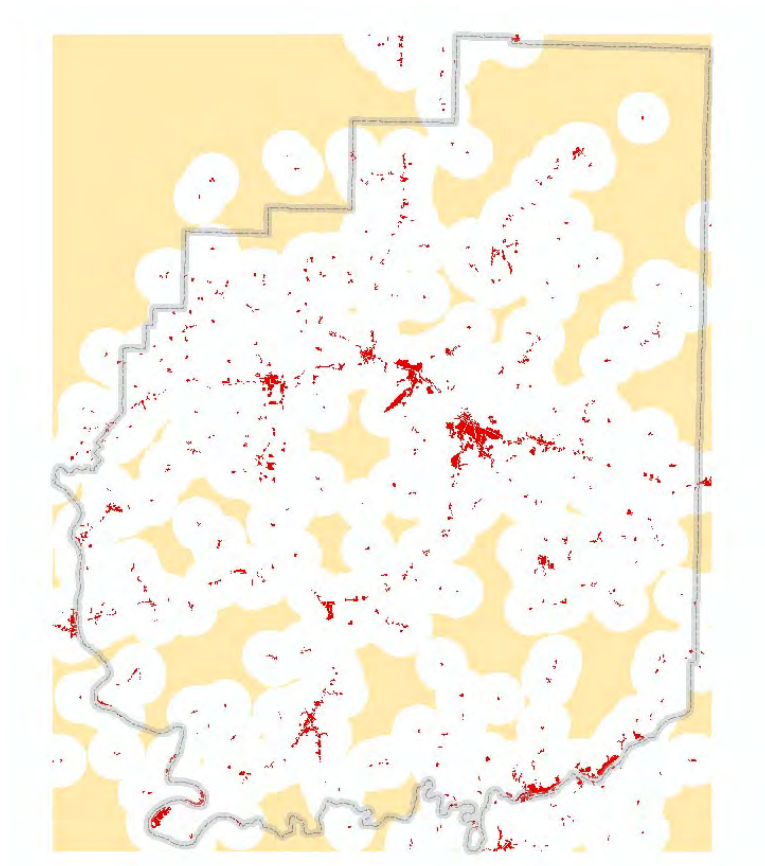
The resulting raster should be similar to the following figure:



For illustration purposes, the Clarion County boundary lines are added to the figure. The figure shows the extracted Urban and Urban-related areas that were reclassified from the original land cover raster dataset.

A.2.1 Preparing Urban Suitability Layer

Repeat the same steps as **A.1.1 Preparing Water Suitability Layer** (above) to produce the **Suitability_Urban** raster dataset using the Euclidean Distance tool first (**UrbanDistance**), then reclassifying the resulting distance layer to 1-Suitable for the classes that are 1600 meters and more, and 0-Not Suitable for the 0-1600 meters distances. The resulting Suitability_Urban raster layer should be similar to the following figure:



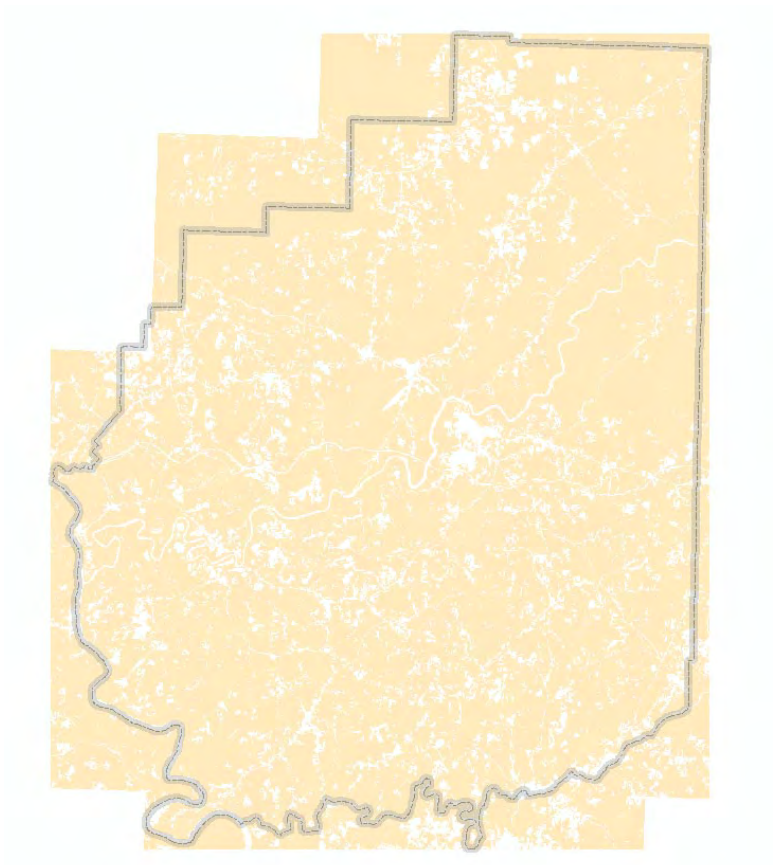
The resulting **Suitability_Urban** raster layer where all the beige cells are suitable: beyond 1600 meters away from the red urban areas (value = 1), and the non-suitable cells (1600 meters around urban area) are left as transparent/white (value = 0)

A.3 Extracting Forest Suitability

Using the same reclassification procedures, classify the Forest areas as 1-Suitable, and assign all other classes a value of 0-Not Suitable. Remember that the Forest classes are given a code of 41, 42 and 43 (Deciduous, Coniferous and mix), and the Grass/Pasture land are given a code of 24 (include in the suitable class). Follow the following table for reclassification scheme, also rename the output to **Suitability_Forest**:

Old values	New values	Notes
14-23	0	
23-43	1	Only 24, 41, 42, and 43 are within this range
43-1233	0	
NODATA	NODATA	

The resulting raster should be similar to the following figure:



For illustration purposes, the Clarion County boundary lines are added to the figure. The suitable forested areas and grass/pasture lands shown in beige are extracted from the land cover raster dataset.

B. Preparing the Road Suitability

For the Road suitability layer, we need to use only those major roads (all roads without the "Other", which include "Interstate", "US" and "PA" from the vector layer.

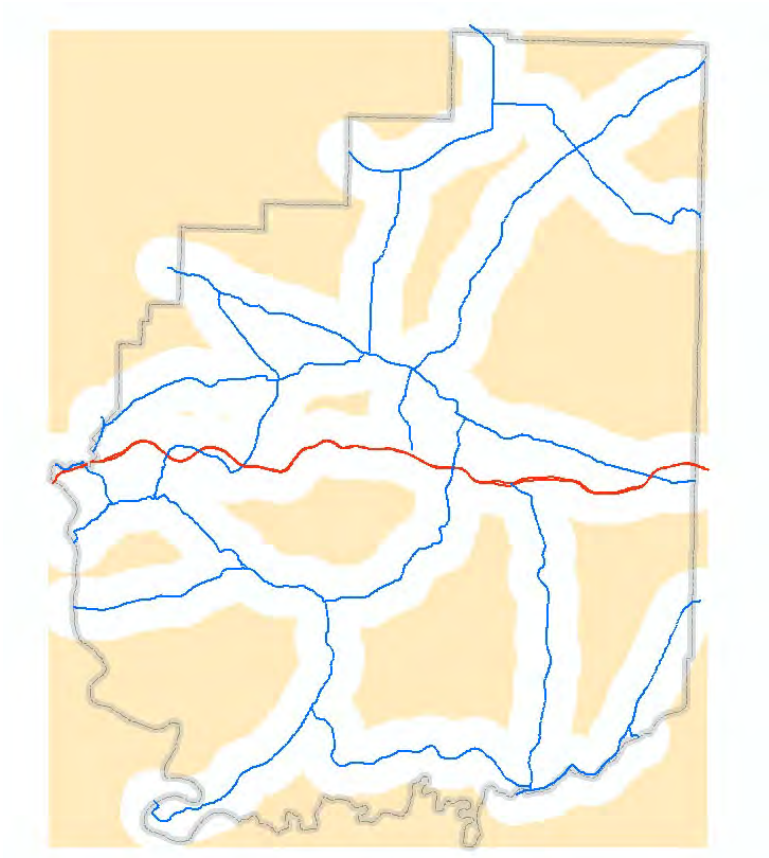
- Add the Roads feature class to the Contents.
- From the Contents pane, right-click the Roads layer and select Properties >> Definition Query >> New Definition Query… and build the following statement:

Where
Road Type
is not equal to
Other
Click Apply
Click OK

This will keep everything but the ones that are labeled as “Other”.

- Run the Euclidean Distance tool to create **RoadsDistance** raster layer. Make sure to change the Output cell size to match the Clarion_LandCover (either browse for Clarion_LandCover raster, or change it to match its resolution).
- Reclassify the RoadsDistance layer to represent all those that are 2000 meters or less to 0-Not Suitable, and all others beyond the 2000 meters to 1-Suitable; name the resulting raster dataset **Suitability_Roads**.

The resulting Suitability_Roads should be similar to the following figure:



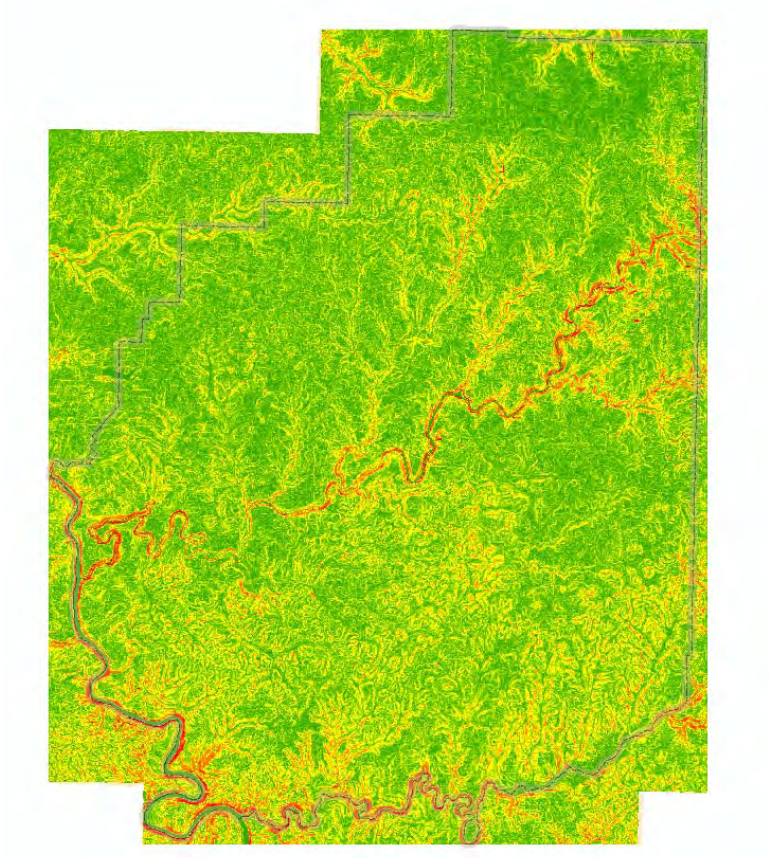
This figure shows the suitable areas (beige, value = 1) based on the 2000 meters away from roads criteria. White areas are not suitable and are assigned a value of 0.

C. Preparing the Slope Suitability

As the final step in data preparation, the slope suitability layer is going to be derived from the given Digital Elevation Model (DEM).

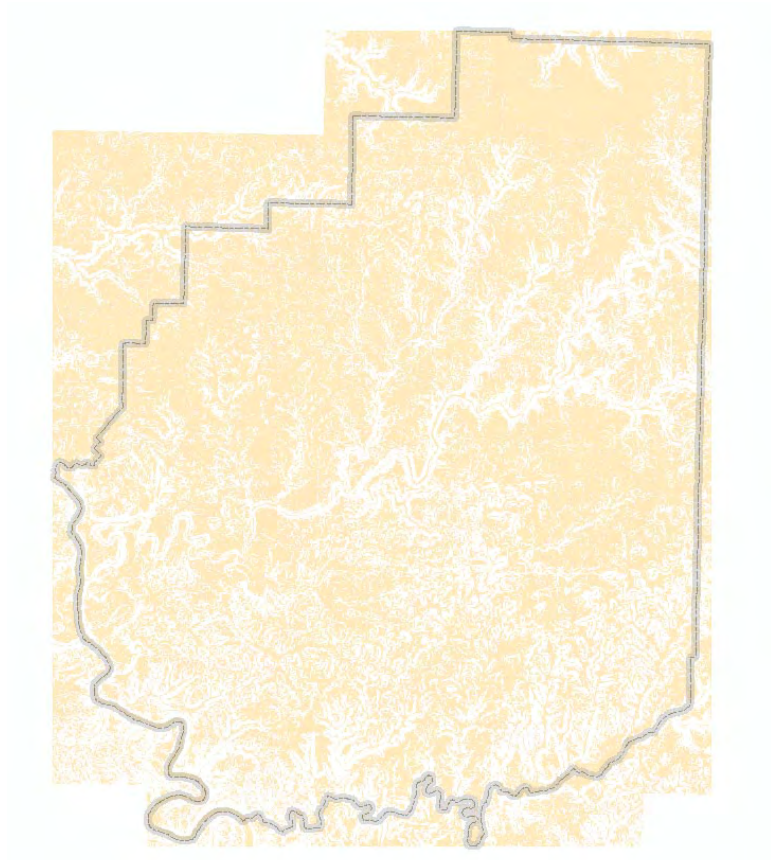
Slope is a surface function that is available in both ArcGIS extensions: Spatial Analyst and 3D Analyst. You can search for the Slope tool in the Analysis tab >> Geoprocessing group >> Tools or find it under ArcToolboxes >> Spatial Analyst Tools >> Surface >> Slope.

- For the Input raster, use the ClarionDEM mosaic dataset, and change the name of the Output raster **ClarionSlopes**.
- Make sure that the Output measurement is set to DEGREE and the Z factor is set to 1, and then click OK.
- The resulting layer should look similar to the following figure:



Slope degrees are calculated from the given DEM raster mosaic. Green areas are relatively flat while the red cells are the steepest.

- Reclassify the ClarionSlopes to 1-Suitable for those slopes between 0 and 8 degrees, and 0-Not Suitable for those beyond those values, name the output **Suitability_Slopes**.
- The Suitability_Slopes layer should look similar to the following figure:



The slopes that are 0-8% were reclassified to be suitable (value = 1, shown in beige) while all other cells were assigned a value of 0 (not suitable)

V. Analysis

There are several methods in the raster modeling processes that would enable the overlay all of the suitability layers together in order to reveal the intersecting cells that are suitable based on all of the 5 given criteria. I'll be discussing two of those methods here, but keep in mind that there would be other ways to accomplish the analysis in addition to those methods presented here.

a. Combine Tool

The Combine tool is part of the Spatial Analyst's Local tools (Analysis tab >> Geoprocessing group >> Tools >> Toolboxes >> Spatial Analyst tools >> Local >> Combine), it'll combine up to 20 input raster datasets into one. The resulting attribute table would have all inputs values as separate fields and therefore would allow for flexible queries of different input combinations. So for 5 input raster layers where each has 2 different values (1-Suitable and 0-Not Suitable), the resulting combined raster would have 32 possible unique combinations, in which the all suitable (all 1) would be one of them.

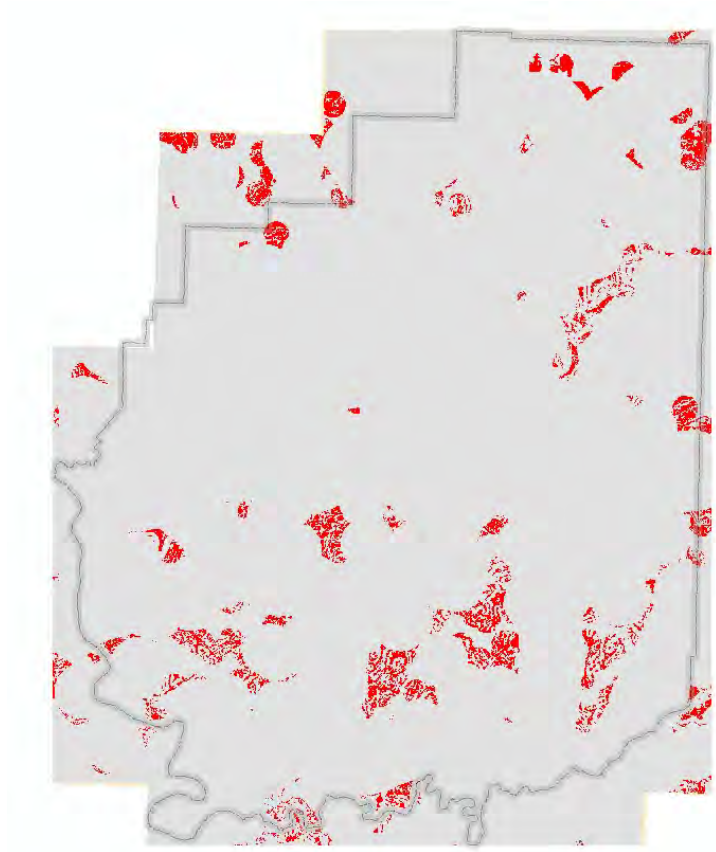
Run the Combine tool; name the output **Suitablle_Combine**.

b. Map Algebra: Raster Calculator

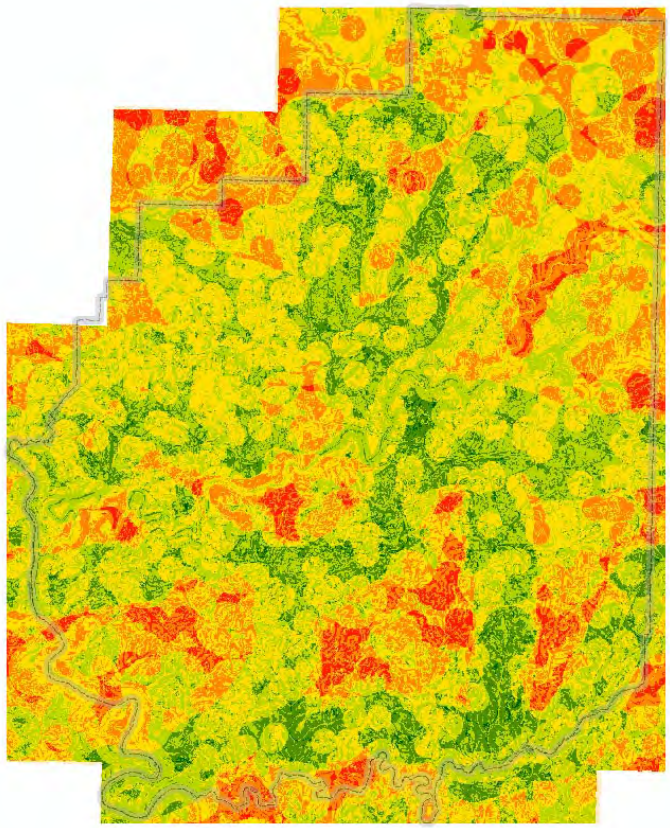
Using the Raster Calculator under the Spatial Analyst Tools >> Map Algebra to arithmetically add all of the suitability raster layers together would produce a cell by cell addition result. In summary, it is expected for 5 inputs with the minimum value is 0 and the maximum value of each is 1, the sum would be between 0 and 5. All cells marked as 5 would be those that fit all 5 given criteria, while all the other values would be variations on how many criteria are met (not identified which one). The drawback of this method is that it is impossible to separate the cells on the basis of individual criteria.

You may run the Map Algebra's Raster Calculator to add all inputs, name the output **Suitable_Sum**.

Using any of those tools should reveal the suitable cells based on the 5 given criteria. The result should all look similar to the following figure.

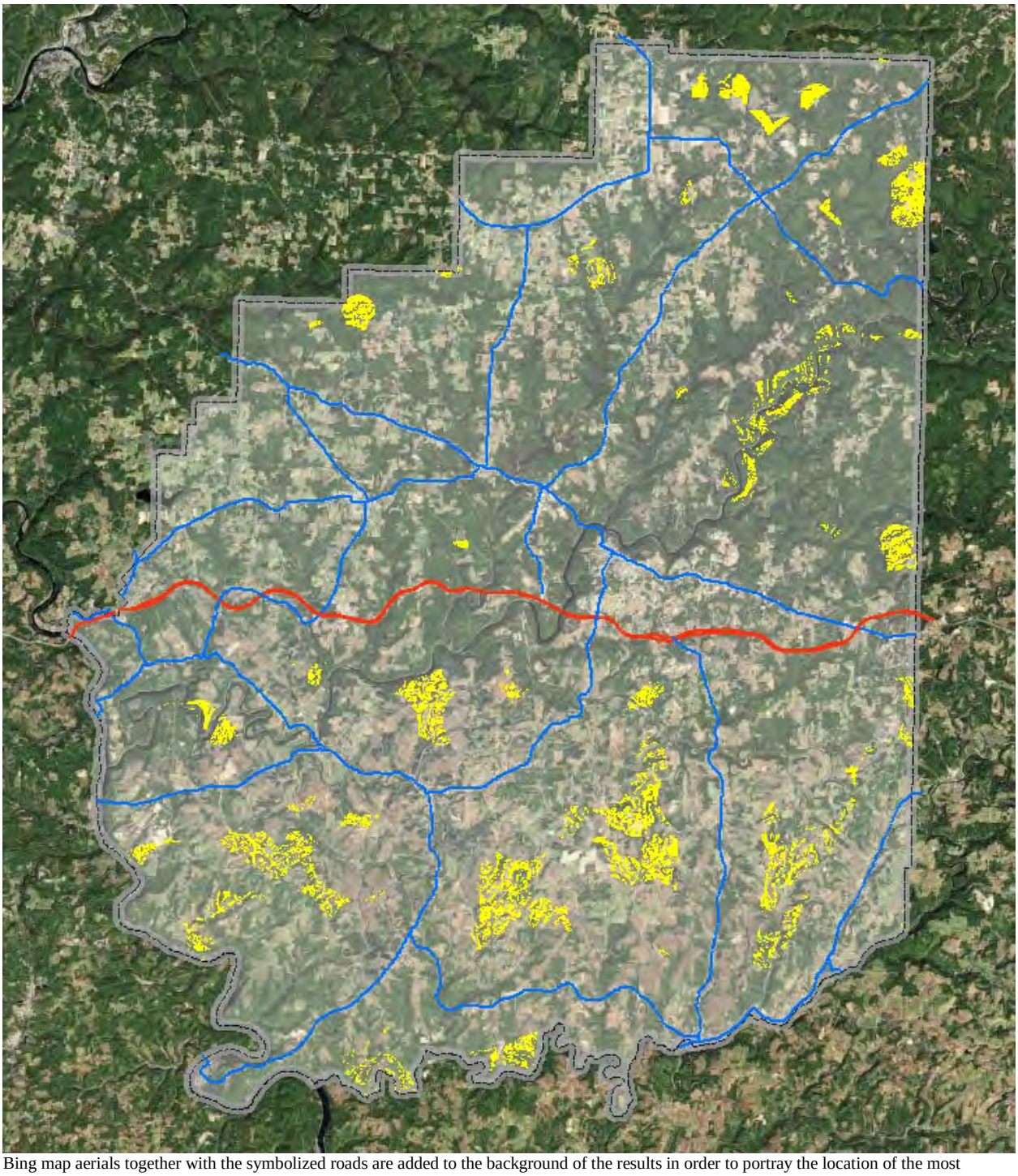


The result of all of the methods regardless of the tool you use should look similar to this figure. The location of the suitable sites that match all 5 criteria are highlighted in red while all other cells are gray.



The result of adding all the input suitability layers using the Map Algebra: Raster Calculator would yield a result that would enable us to classify the cells to ones that meet all criteria or some of them (0 to 5 of the given criteria). Here the red cells are the best suited with all 5 criteria met, while the darker green ones are the ones that did not meet any of the criteria.

- Open the attribute table of either of the results (*Suitable_Combine* or *Suitable_Sum*), check the number of fully suitable pixels (Value = 12 for *Suitable_Combine*, and value = 5 for *Suitable_Sum*), both are 74,498 as cell count. The cell size of both of the output is 30x30 meters, so the total suitable area would be 900x74,498 = 67,048,200 Square Meters, or 67.0482 Square Kilometers.
- Refine the output by clipping the resulting raster layer to the boundaries of Clarion County (Search for "Clip Raster" – ArcToolbox >> Data Management Tools >> Raster >> Raster Processing >> Clip Raster) then recalculate the total suitable area.
- You may also add a background layer (basemap) and the roads in order to add a sense of location to the results:



Bing map aerials together with the symbolized roads are added to the background of the results in order to portray the location of the most suitable cells within the county of Clarion. Those yellow cells have a slope of less than 8%, are 2000 meters away from roads and 1600 from urban and built up areas, within 800 meters of water bodies and are mostly forested, pasture or grass land.

VI. Assessing the Results

After achieving the resulting raster data, the attribute table would contain the value (either 0-5 for the calculation method, or 0-32 for the combine method with those 5 inputs), and the cell count for each of the given values. We can view those sites that are 100% matching the given criteria, but also we can take a closer look at those that did not necessarily meet all of them: the ***Suitable_Combine*** dataset would contain the best data for this type of selection.

Open the attribute table of the ***Suitable_Combine*** raster layer, add a short integer field to the table and calculate it so that it sums up all Suitability fields. The highest value should be 5 with only one record that matches this value. Notice that there are 5 records that have a value of 4; a value of 4 would mean that those cells has met all but one criteria. If you take a closer look at a quick statistics, those constitute about 18% of the total area of the results. You may check the statistics using the following method:

- From the attribute table, select all records with a value of 4
- Right-click the COUNT field and click Statistics
- Take a note of the SUM of those records (384,221)
- Clear the selection and get the statistics for the COUNT field, the total SUM of all records in the table is 2,160,059
- With a quick calculation using excel or a calculator, divide the SUM of the selected records by the total SUM of all records and multiply the result by 100. The result would be 17.79%

You can use the statistics calculations to assess the results of any criteria combination at any time.

You might consider adding one or more of those group of cells that met all but one criteria in your final report if you see that they are relevant. Here I'm going to show you how to proceed with looking deeper at those that met all of the criteria only (all 1s or SUM field = 5).

For this demonstration, I'm going to discard any cells that do not have a SUM = 5, or in other way, I can use the ***Suitable_Sum*** raster dataset instead. Using a Reclassify on the *Suitable_Sum*, I assigned a value of 1 to the old value 5, and assigned a value of "NODATA" to everything else. Doing this, I'm creating only one zone (value =1) for the data.

In order to get information about cell clusters and their areas, I need to use Regions information (review what are the zones and regions in the introduction of this section). Regions are going to be contiguous clusters of cells, each patch would be assigned a new value and cell counts. We use the Region Group tool under Spatial Analyst Tools >> Generalization >> Region Group tool to accomplish this. Use the ***Suitability_Sum*** as the Input raster, name the Output raster to ***Suitability_Sum_Regions***, while leaving everything else as default and click OK.

Open the attribute table of the *Suitability_Sum_Regions* and sort the COUNT field descending. Now we are able to see all cell groups that are contiguous while also being able to quantify them using the COUNT field. We may add field called SQKM (Type: Double) and calculate the area in square kilometers by multiplying the COUNT filed by the cell area (30x30 meters or 0.0009 Square Kilometers).

It is important before calculating the final areas and percentages for reporting to clip the dataset to the extents of Clarion County (using the Clip Raster tool) name the output ***Suitability_Sum_Regions_Clarion***.

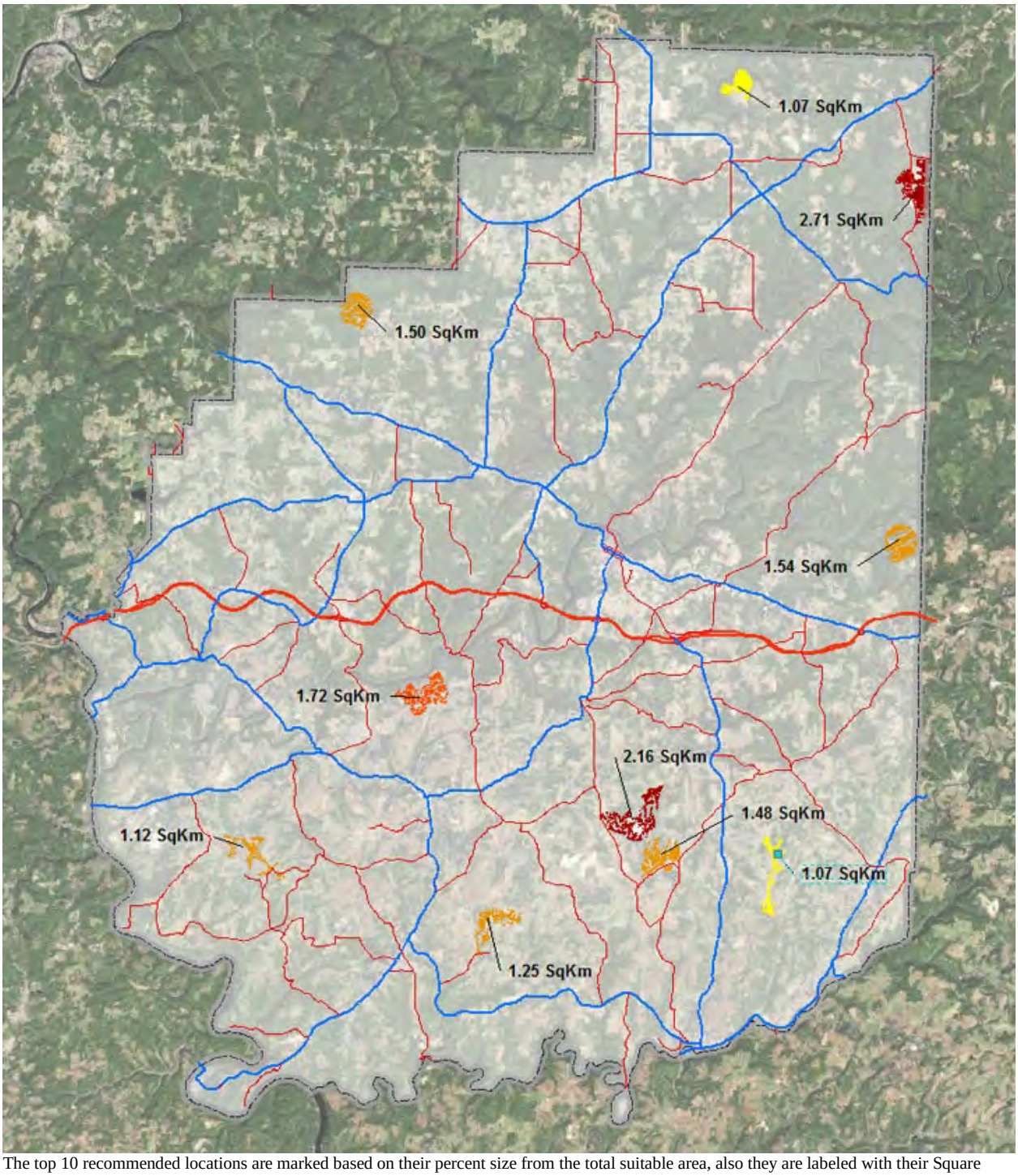
Add another field in the resulting raster dataset attribute table (*Suitability_Sum_Regions_Clarion*) call it **PERCENTAREA**. Calculate this field by dividing the SQKM by the sum of SQKM statistics for all records (right click SQKM >> Statistics, record the Sum value) then multiply it by 100:

```
100 * !SQKM! / 52.5294
```

Sort the PERCENTAREA field descending, you'll notice that we have 10 records that are more than 2% of the total Square Kilometers or above. For the sake of the demonstration I elect to report only those 10 records, those are more than 1 Square Kilometer in area (SQKM), and would constitute about 30% from the total suitable sites. But you may assess in the Your Turn section the locations of those sites and report the best sites you see fit to the given application.

VII. Report the findings

The final map should include those elements that would help a decision maker to spot the proper location based on your findings. The following figure summarizes the results and includes quick information about the site location as well as its size. You may also include a tabular report that sums the data as a whole and have a quick statistics about the average size of selected locations in general as well as about those recommended by your map.



The top 10 recommended locations are marked based on their percent size from the total suitable area, also they are labeled with their Square Kilometers.

VIII. Your Turn

Now, using similar procedures, and using the same given datasets, you are asked to perform a similar analysis to select those locations that match the following criteria:

1- Consider sites that have a slope between 0% and 10%;
2- Fall within a distance of 100 Meters from existing streams;
3- Located in a Forested, Grassland or pasture land;
4- 2 Kms (2000 Meters) away from Interstate 80; and,
5- Is within 250 Meters of all other roads for accessibility.

Those conditions are suitable for a stream remediation plan where proximity to the existing stream network is necessary for site reclamation, and proximity to existing roads is important for accessibility. The process would possibly include the construction of settling ponds, so the landscape needs to be relatively flat. We also need to avoid the proximity to heavy and high speed traffic Interstate for the safety of crew while maintaining the site away from any possible heavy road treatment that might affect the water chemistry in the treatment site.

You are required to create a **PDF report** (start in any word processing software (e.g. Word), then save it as PDF) of the procedures and results. Also, **include your geodatabase with only the main steps** that you followed in order to reach the final result. **Create a ZIP file for your exercise folder** and submit the zipped report and geodatabase to the corresponding assignment.

a. Report Content

- Include a cover page with a selected title "Raster Analysis: Binary Overlay for Stream Remediation Site Selection" or something similar, your name, class number and date of the report.
- No text is necessarily required, only **the figure captions** of every step you make during the analysis.
- Make sure to save all of the essential steps using the same raster dataset file naming convention used in the demonstration exercise.
- Start with a flowchart similar to the one I provided you at the beginning of this exercise. Your flow of work might differ than the one presented, as long as it will reveal the same results.
    - You build a flowchart using model builder, save the model inside the geodatabase/project
    - You may use a function such as "Reclassify" and define/rename its inputs and outputs accordingly.
    - Export this model to Graphic, or take a screen shot. You may then insert the resulting image in your final report.

- Summarize the criteria, the required datasets and the necessary attributes and attribute values in a table such as the one presented in the previous demonstration.
- Take screen snapshots (Alt + Prnt Scrn) of your Map window every time you make a step or an action that changes the result on the screen and paste the resulting clipboard image in your word processing software (e.g. Word), **DO NOT** crop or resize the pasted image, I would like to be able to see the Contents organization of layers. You may create a 1 column 2 rows table; past the image inside the top cell and insert your text in the lower one. if you can't see the details of the snapshot screen capture images very well .. I won't either!
- Under each pasted image, type in a brief description on how did you reach this step and what does it show.
- Show any possible locations that you might recommend that can be considered for the site selection even if they do not meet all of the criteria.

b. Grading Rubric

Grading for this exercise will be according to the following rubric:

Criteria	Score From
Cover Page	5
Flowchart	15
Summary Steps Table	10
Perform the Data Preparation & Analysis Tasks	35
Result Assessment	20
Final Results Reporting (Map/Table/Graph)	10
Overall Report Presentation Quality	5
TOTAL	**100**

Partial completion of any of the grading criteria will be graded according to the level of completeness/quality based on the following scheme for every criterion:

Level	%
Excellent	100
Very Good	75
Good	50
Fair	25
Absent	0

IX. Readings

- The Types of Operations in Spatial Analyst
  https://desktop.arcgis.com/en/arcmap/latest/extensions/spatial-analyst/performing-analysis/the-types-of-operations-in-spatial-analyst.htm

- NoData and How it Affects Analysis
  https://desktop.arcgis.com/en/arcmap/latest/extensions/spatial-analyst/performing-analysis/nodata-and-how-it-affects-analysis.htm

- The Analysis Environment of Spatial Analyst
  https://desktop.arcgis.com/en/arcmap/latest/extensions/spatial-analyst/performing-analysis/the-analysis-environment-of-spatial-analyst.htm#:~:text=The%20analysis%20environment%20includes%20the,the%20environments%20at%20various%20levels.