Raster Analysis

Weighted Overlays

Table of Contents

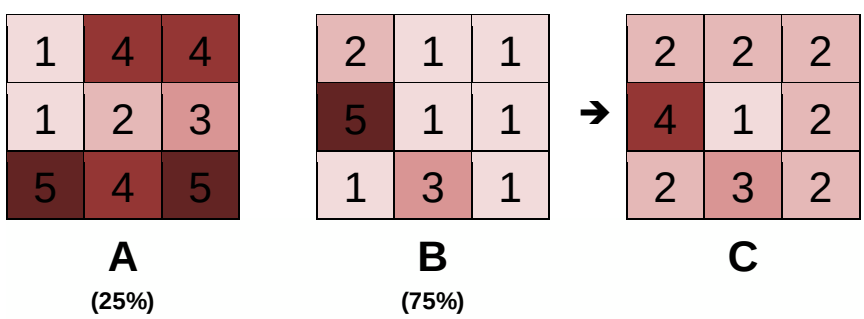
I. Basics of Raster Weighted Overlays 2
II. Planning Weighted Overlay Analysis 3
III. Introduction 4
A. Defining an Evaluation Scale 4
B. Adding Rasters 13
C. Setting Scale Values 13
D. Assigning Weights for Inputs 13
E. Setting up Mask Environment 13
F. Running the Tool 14
IV. Assessing the results 18
V. Your Turn 20
A. Report Content 21
B. Grading Rubric 22

I. Basics of Raster Weighted Overlays

In the previous section you were introduced to the most basic and straightforward raster overlay method: Binary. Now the binary overlay method is very simple and relies on the intersection of suitability values between different inputs; all inputs has to be equally suitable in order for the output to also be categorized as suitable. This is not necessarily true, or always true, in real world. In reality, some inputs could play a greater role than others in site selection and should be ranked higher in the analysis; i.e. given a higher priority. For this, all inputs should be weighted according to their level of importance to the selection process. For example, a natural wildlife habitat could vary depending on the species; each might have different requirements, for example, the distance from a built environment can be closer or further than a specific number, they might require less forest coverage and more of grassland, or are best located in slopes that are between 5% and 10% but can also be found in steeper or flatter landscapes.

Weighted Overlay Analysis deal with those types of situations, avoiding the exclusive property of the binary method, it attempts to take into consideration a weighted factor within each input as well as between the inputs too. In other words, each raster dataset values should be ranked according to their suitability (most suitable to not suitable), as well as the raster dataset itself should be weighted in the analysis compared to the other datasets (e.g. proximity to roads is more important than the soil types, or forest land cover is less important to proximity to streams …etc.).

Weighted Overlay analysis uses integer raster datasets as input. It expects that all inputs would be classified using the same scale where the higher value is the best suited and the lowest is the worse or not suitable. Then, according to the importance of each of the inputs, they have to be given a percentage weight (from 100%).



For the weighted overlay multiplies each cell value to its corresponding weight percentage then adds them together. The rounded result is placed as an integer value for the output. In this example, first cell on the top left corner: (1 * 0.25) + (2 * 0.75) = 1.75 (i.e. 2), and for the last cell on the lower right corner: (5 * 0.25) + (1 * .75) = 2

II. Planning Weighted Overlay Analysis

Weighted Overlays rate suitable locations when the level of suitability corresponds to well-defined value ranges on source layers, and when you want to assign more importance to some criteria than others. Therefore, defining the problem you're trying to solve would be the first step, which will drive not only the process but also what type of output and what are the expected results. Based on the problem definition, you start extracting the set of criteria that would make a location within the suitable (or not suitable) range. It is prudent to identify the different input layers needed for the analysis together with the necessary attributes as usual. But in the case of a weighted overlay analysis, you take a step further in defining suitability levels or classes within each of the source layers using a set scale, then define the importance of each of those layers as a percentage weight in the process.

According to ArcGIS documentation about weighted overlay, the steps for running the tools can be summarized as:

A. Select an evaluation scale: Where the highest value of the scale represents the best suitability and the lowest is the worse. It is important to note that the selected evaluation scale is going to be equally used for all value data classification of each input layer as well as for the final output representation.
B. Adding rasters: Only discrete integer rasters are accepted as input, all continuous data has to be reclassified prior to using them in the process
C. Set scale values for each input: You may elect to change the re-classify values or use them as they are according to the set evaluation scale.
D. Assign weights to input rasters: Each input could be assigned a weight (out of 100) according to its importance or relevance to the analysis subject.
E. Run the tool.

Raster Analysis

Weighted Overlay Method

SUMMARY

Similar to the Binary Overlay method, this assignment will repeat the process but with a more realistic vision of the nature of natural wildlife habitats. Those habitats cannot be defined strictly based on given distances, or land covers, but rather a set of levels where areas would be more or less suitable than others based on proximity or land cover types. For example, the closer to an urban area the less likely natural wildlife habitat would prevail or the flatter the landscape the better the habitat for many species, but a little bit of slope would not seriously affect the habitat.

III. 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.

To review the problem of the previous exercise, in this assignment we would like to investigate the areas in Clarion County 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. It should be located within 800 meters from water bodies;
2. Generally it would be about 2000 meters away from major roads ;
3. It should be located 1600 meters away from Urban/Built-up areas;
4. The land should be covered by open land or forested; and,
5. The landscape should be relatively flat (0-8% slope).

A. Defining an Evaluation Scale

Accordingly, each of those conditions has its own “ranking” of suitability. The following tables summarize the suitability ranks in each of the given layers and their corresponding range of values. I elected to use a 1-5 scale for evaluation; which would make the output less complex and easier to read and interpret.

Distance from Water (m)	Suitability
< 200	5
200 – 400	4
400 – 600	3
600 – 800	2
> 800	1

Distance from Urban (m)	Suitability
< 400	1
400 – 800	2
800 – 1,200	3
1,200 – 1,600	4
> 1,600	5

Distance from Roads (m)	Suitability
< 500	1
500 – 1,000	2
1,000 – 1,500	3
1,500 – 2,000	4
> 2,000	5

Slope Class (%)	Suitability
< 2	5
2 – 4	4
4 – 6	3
6 – 8	2
> 8	1

In addition, from the land cover information, it is necessary to note that the best suited areas would be forested or open grassland. Nevertheless, other land cover might be of potential suitability depending on their type. The following table is an attempt to assign suitability ranks to the different available land cover classes. Notice that the urban/built up covers such as the residential land was assigned a higher suitability rank than the Industrial ones, furthermore, the residential with 5-30% impervious surfaces and with more tree cover was assigned a higher suitability rank than its own class with more impervious and less tree cover. The other classes were re-assigned a suitability rank according to their potential of contributing to a natural wildlife habitat.

	Land Cover Code Description	Original Code	Suitability
1.	Roads	14	1
2.	Residential Land; 5-30% impervious	111	3
3.	Residential Land; 5-30% impervious; Deciduous tree cover	1111	3
4.	Residential Land; 5-30% impervious; Evergreen tree cover	1112	3
5.	Residential Land; 5-30% impervious; Mixed tree cover	1113	3
6.	Residential Land; 31-74% impervious	112	1
7.	Residential Land; 31-74% impervious; Deciduous tree cover	1121	1
8.	Residential Land; 31-74% impervious; Evergreen tree cover	1122	1
9.	Residential Land; 31-74% impervious; Mixed tree cover	1123	1
10.	Residential Land; 74% < impervious	113	1
11.	Residential Land; 74% < impervious; Deciduous tree cover	1131	1
12.	Residential Land; 74% < impervious; Evergreen tree cover	1132	1
13.	Residential Land; 74% < impervious; Mixed tree cover	1133	1
14.	Institutional/Industrial/Commercial Land; 5-30% impervious	121	2
15.	Institutional/Industrial/Commercial Land; 5-30% impervious; Deciduous tree cover	1211	2
16.	Institutional/Industrial/Commercial Land; 5-30% impervious; Evergreen tree cover	1212	2

	Land Cover Code Description	Original Code	Suitability
17.	Institutional/Industrial/Commercial Land; 5-30% impervious; Mixed tree cover	1213	2
18.	Institutional/Industrial/Commercial Land; 31-74% impervious	122	1
19.	Institutional/Industrial/Commercial Land; 31-74% impervious; Deciduous tree cover	1221	1
20.	Institutional/Industrial/Commercial Land; 31-74% impervious; Evergreen tree cover	1222	1
21.	Institutional/Industrial/Commercial Land; 31-74% impervious; Mixed tree cover	1223	1
22.	Institutional/Industrial/Commercial Land; 74% < impervious	123	1
23.	Institutional/Industrial/Commercial Land; 74% < impervious; Deciduous tree cover	1231	1
24.	Institutional/Industrial/Commercial Land; 74% < impervious; Evergreen tree cover	1232	1
25.	Institutional/Industrial/Commercial Land; 74% < impervious; Mixed tree cover	1233	1
26.	Airports	124	1
27.	Row Crops	21	4
28.	Pasture/Grass	24	5
29.	Golf Courses	241	4
30.	Deciduous Forest	41	5
31.	Evergreen Forest	42	5
32.	Mixed Deciduous and Evergreen	43	5
33.	Water	50	1
34.	Forested Wetlands	61	5
35.	Emergent Wetlands	62	5
36.	Bare; Unclassified Urban/Mines, Exposed Rock, Other Unvegetated Surfaces	70	2
37.	Active Mines/Significantly Disturbed Mined Areas	750	1

So, in order to achieve those classifications of suitability, you'd need to run the Reclassify tool on each of those five input rasters (Land cover, Roads, Urban areas, Water bodies, and Slopes). Watch the need to switch the suitability ranking according to the input layer (5-Most suitable and 1-Least suitable). For example, notice that the closer from water bodies the more suitable a site would be, therefore, from the previous tables, a Suitability rank of 5 would be assigned to distances of 200 or less from water bodies. On the other hand, the further the distance from Urban areas and Roads the best suited, therefore, a distance greater than 1,600 meters from Urban areas and greater than 2,000 meters from roads would be most suitable (5). Same concept applies for the Slope suitability (steepest areas are less suitable than flatter ones).

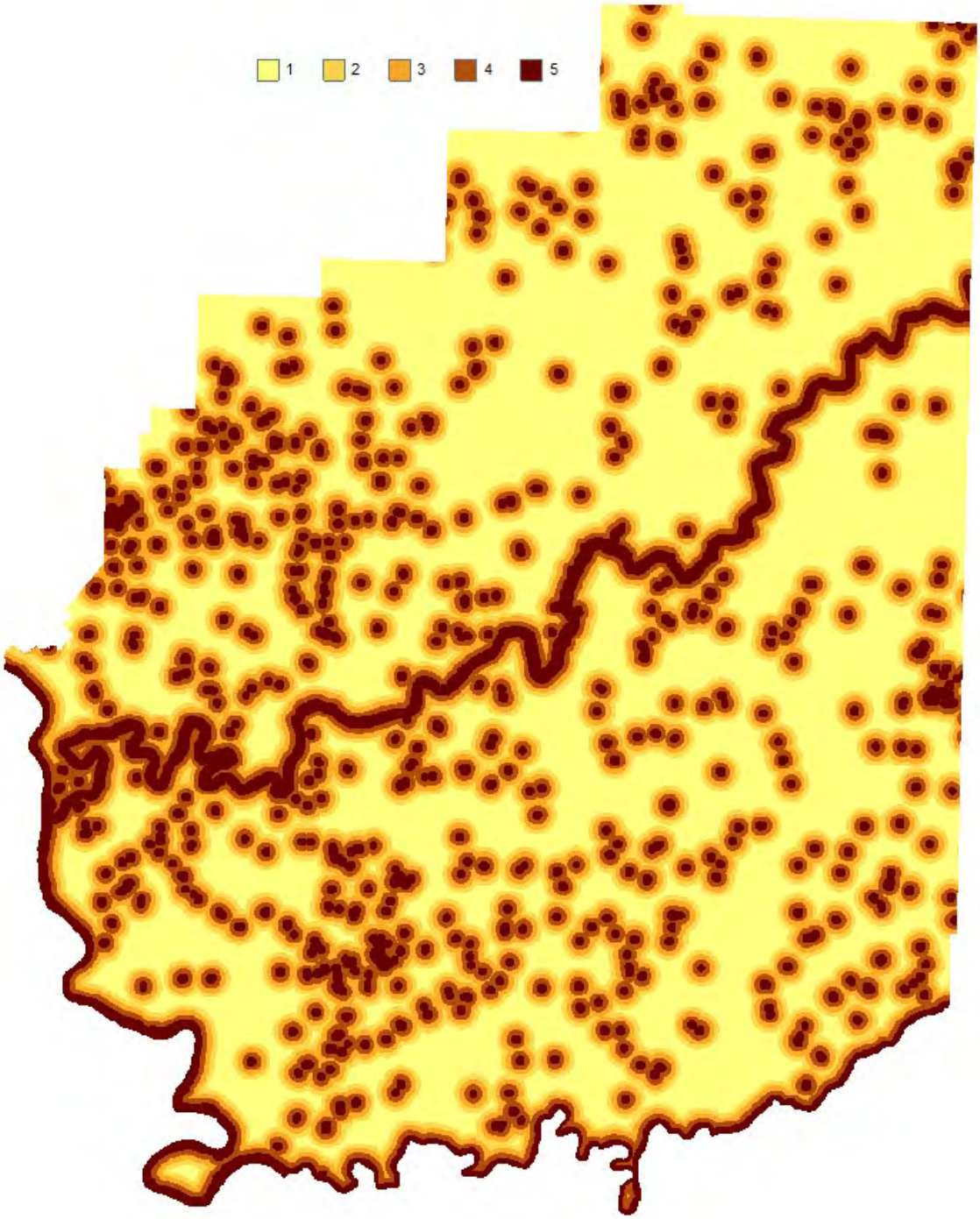
You may use the same Euclidean Distance results from the previous assignments as a basis for this one, or re-calculate the Euclidean Distances for each of the given layers. You may also

extract the water bodies and the urban areas in separate layers prior to the distance calculation.

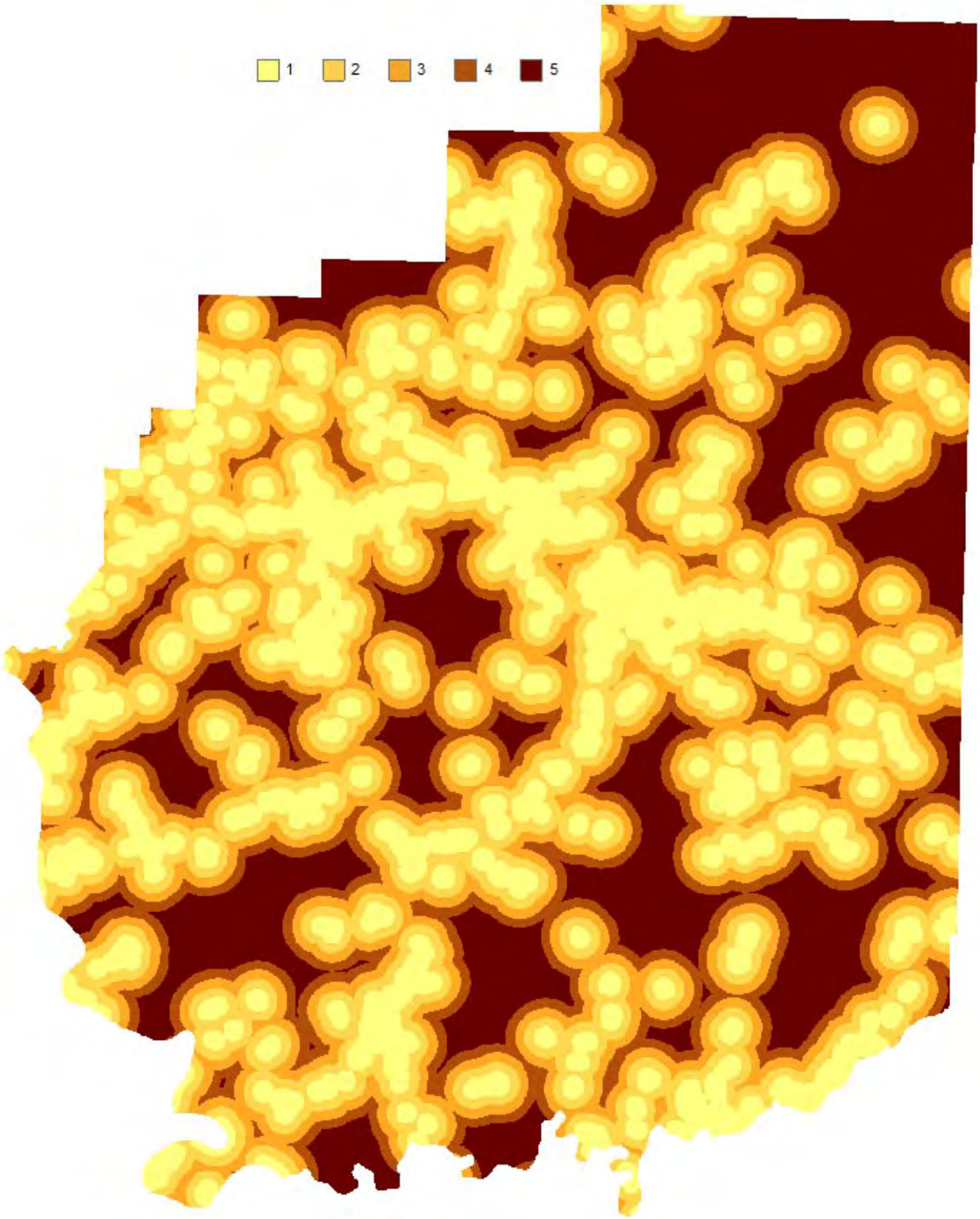
Name the reclassified rasters as:

- **Water_Reclass**;
- **Urban_Reclass**;
- **Roads_Reclass**;
- **Slopes_Reclass**; and,
- **LandCover_Reclass**

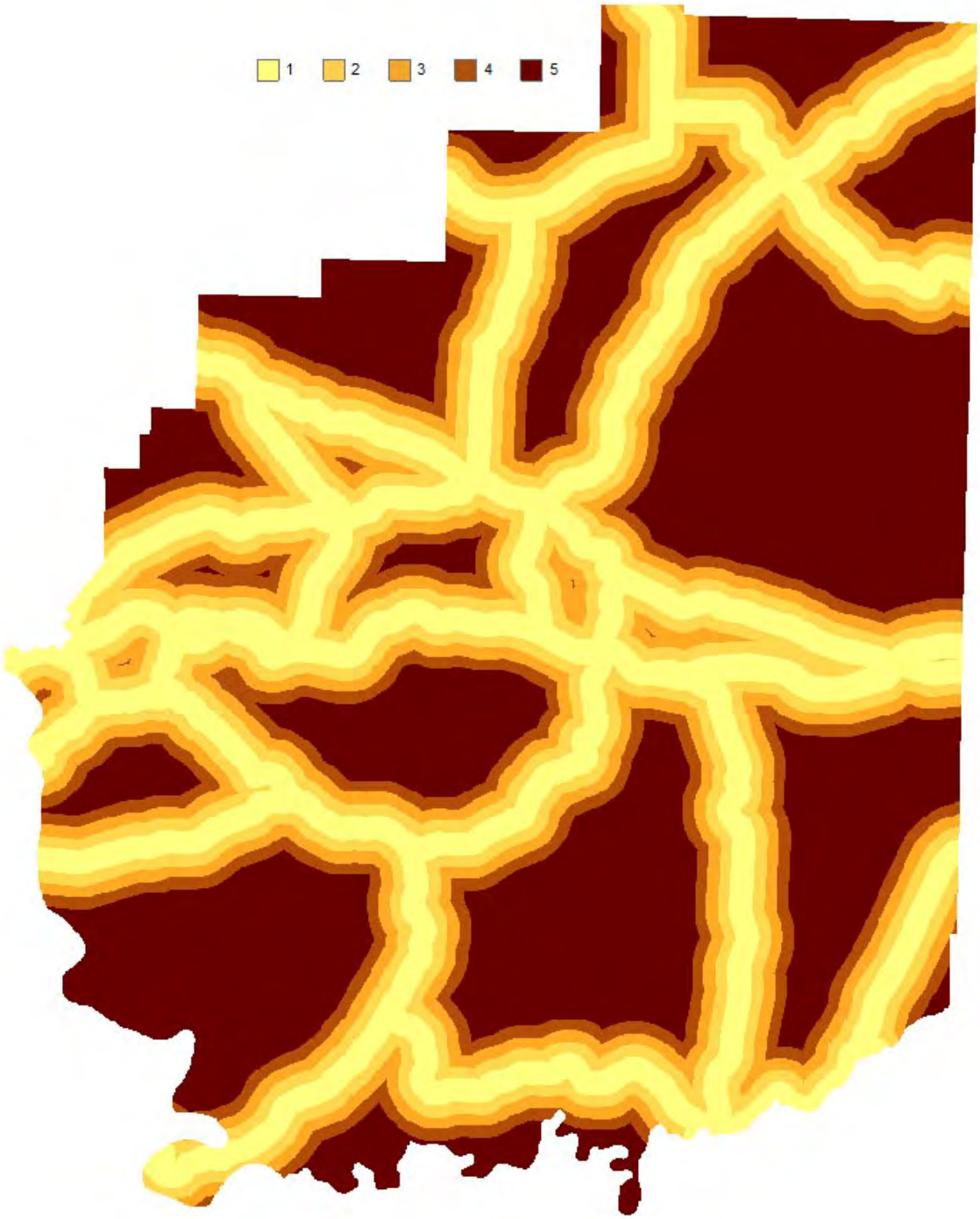
The following series of figures show the resulting reclassified layers and their corresponding suitability codes.



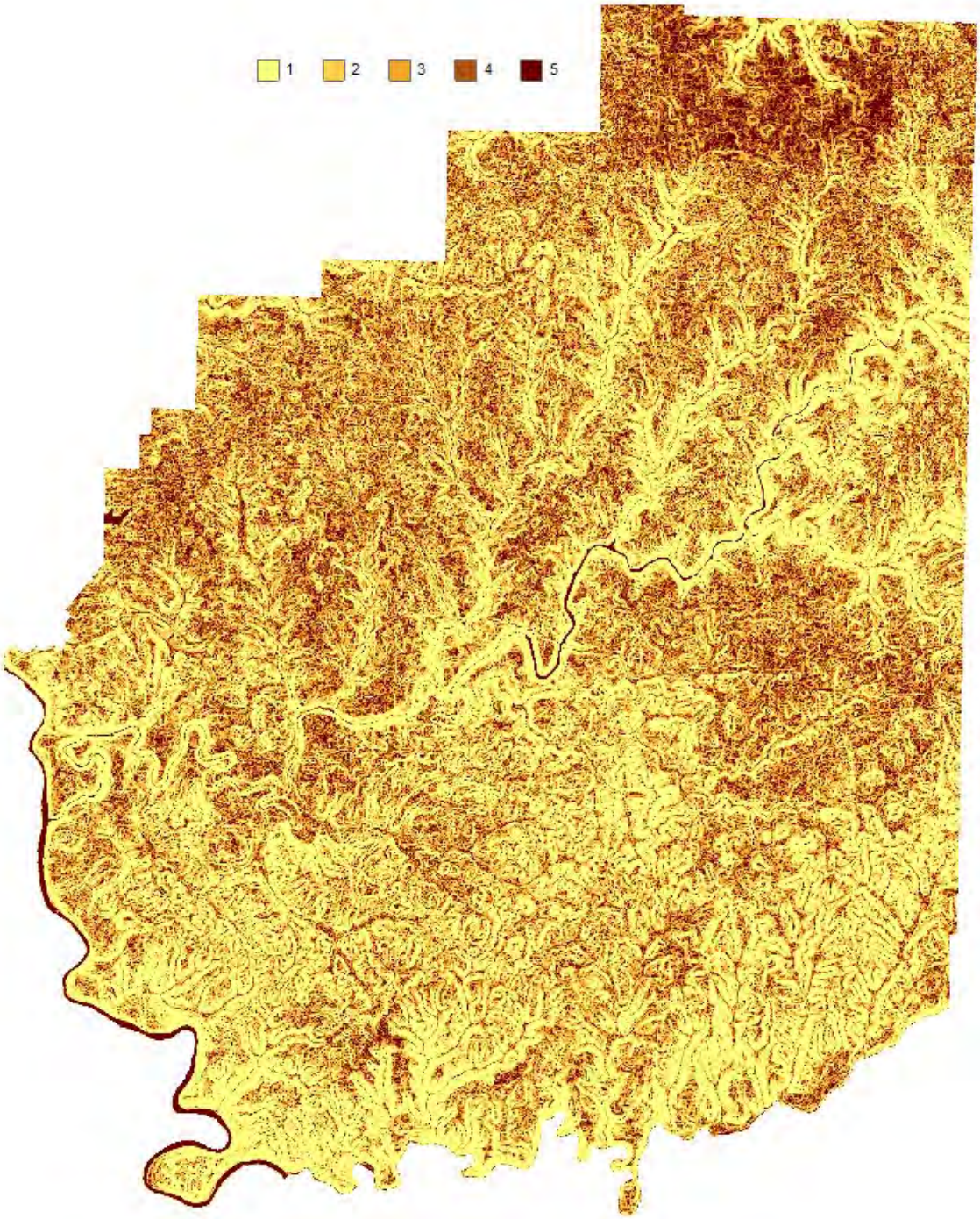
Water Reclassification: The water reclassification shows that the further the distance from a water body the less suitable a location would be. Areas closer to water bodies are assigned the most suitable value of 5, while those that are further away are assigned the least suitable value of 1.



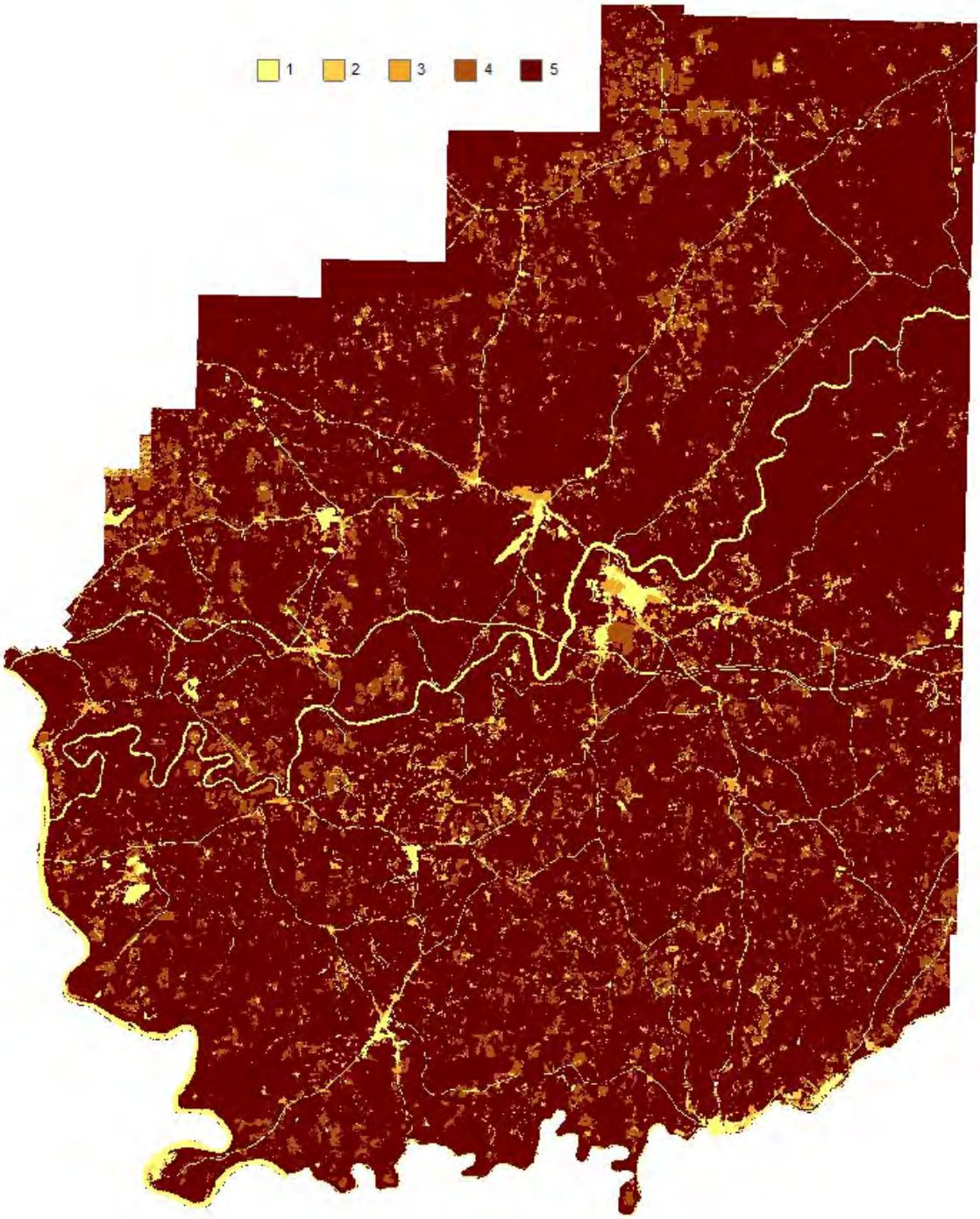
Urban Reclassification: The urban reclassification shows that the further the distance from a built-up area the more suitable a location would be. Areas closer to urban areas are assigned the least suitable value of 1, while those that are further away are assigned the most suitable value of 5.



Roads Reclassification: The roads reclassification shows that the further the distance from a major road the more suitable a location would be. Areas closer to the main roads are assigned the least suitable value of 1, while those that are further away are assigned the most suitable value of 5.



Slopes Reclassification: The slopes reclassification shows that those flatter areas are the most suitable for a natural wildlife habitat (5) while the steepest landscapes are the least suitable (1).



Land Cover Reclassification: More natural land covers are assumed most suitable to a natural wildlife habitat (value = 5) while others more man-made and disturbed land covers are considered the least suitable ones (value = 1).

B. Adding Rasters

Run the Weighted Overlay tool from Analysis tab >> Geoprocessing Group >> Tools, search for Weighted Overlay, it should be located under Toolboxes >> Spatial Analysis Tools >> Overlay >> Weighted Overlay

Start adding the reclassified raster layers from the previous steps (**Water_Reclass**, **Urban_Reclass**, **Roads_Reclass**, **Slopes_Reclass**, and **LandCover_Reclass**).

C. Setting Scale Values

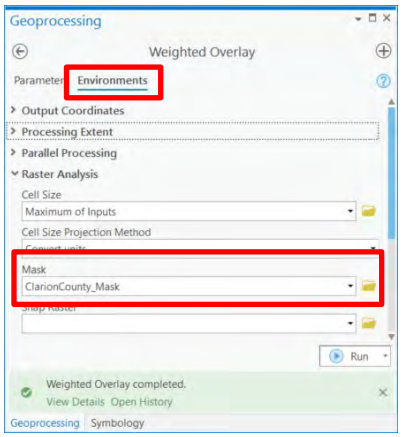
For every input, the tool lists the list of values in the VALUE field as well as a default assignment of the Scale column. Change the Scales to 1 - 5 ; then re-assign, if necessary, the Scale values of each input to match those of its VALUE field.

D. Assigning Weights for Inputs

Now, it is kind of difficult to come up with a sole schema for weighting those input layers according of importance, for that reason, and in order to take a deeper look at how changing those weights would affect the final results, we are going to run this tool 4 times, each time we are going to change the % influence or input weights to a set of different values according to the following table:

	Percent Influence			
Input	**Equal**	**LC_Urban**	**LC_Water**	**LC_Slopes**
Water	20	10	20	20
Urban	20	25	15	10
Roads	20	15	10	15
Land Cover	20	40	40	30
Slopes	20	10	15	25

In this table, you'll find that the first run assumes an equal influence (20% each) from every input layer, all other three runs assume a domination from the Land Cover as the major factor, while each would assign a higher influence to another input: the second run assumes a higher influence from the Urban distance layer, the third Water, and the fourth the Slopes.



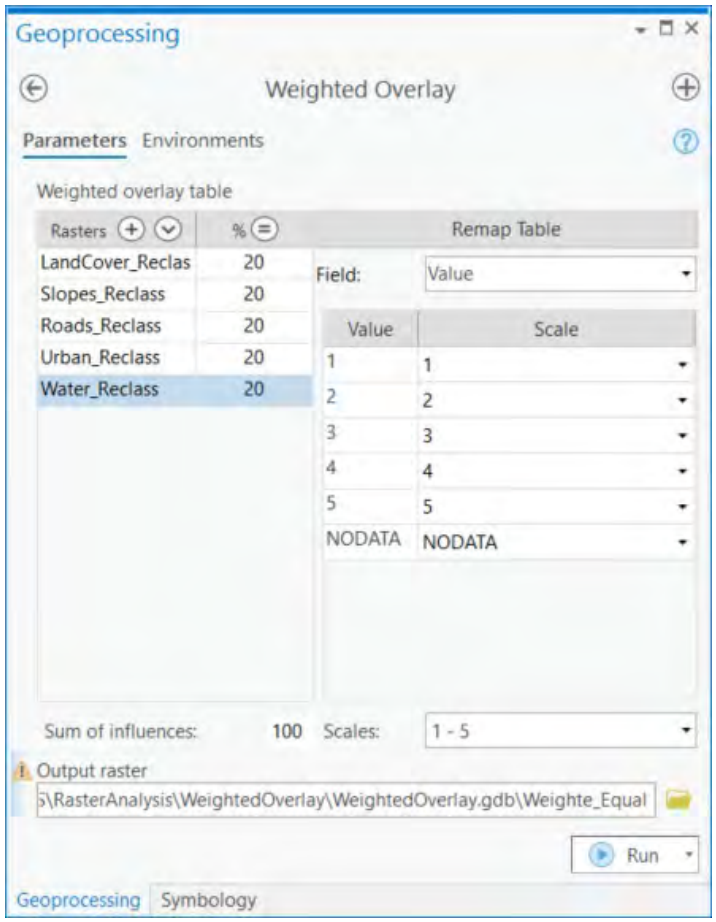
E. Setting up Mask Environment

In order to make sure that the processing occurs within the limits of Clarion County, you need to add the ClarionCounty_Mask layer to the

Mask field of the Weighted Overlay tool. It is located under the Environments tab >> Raster Analysis >> Mask.

F. Running the Tool

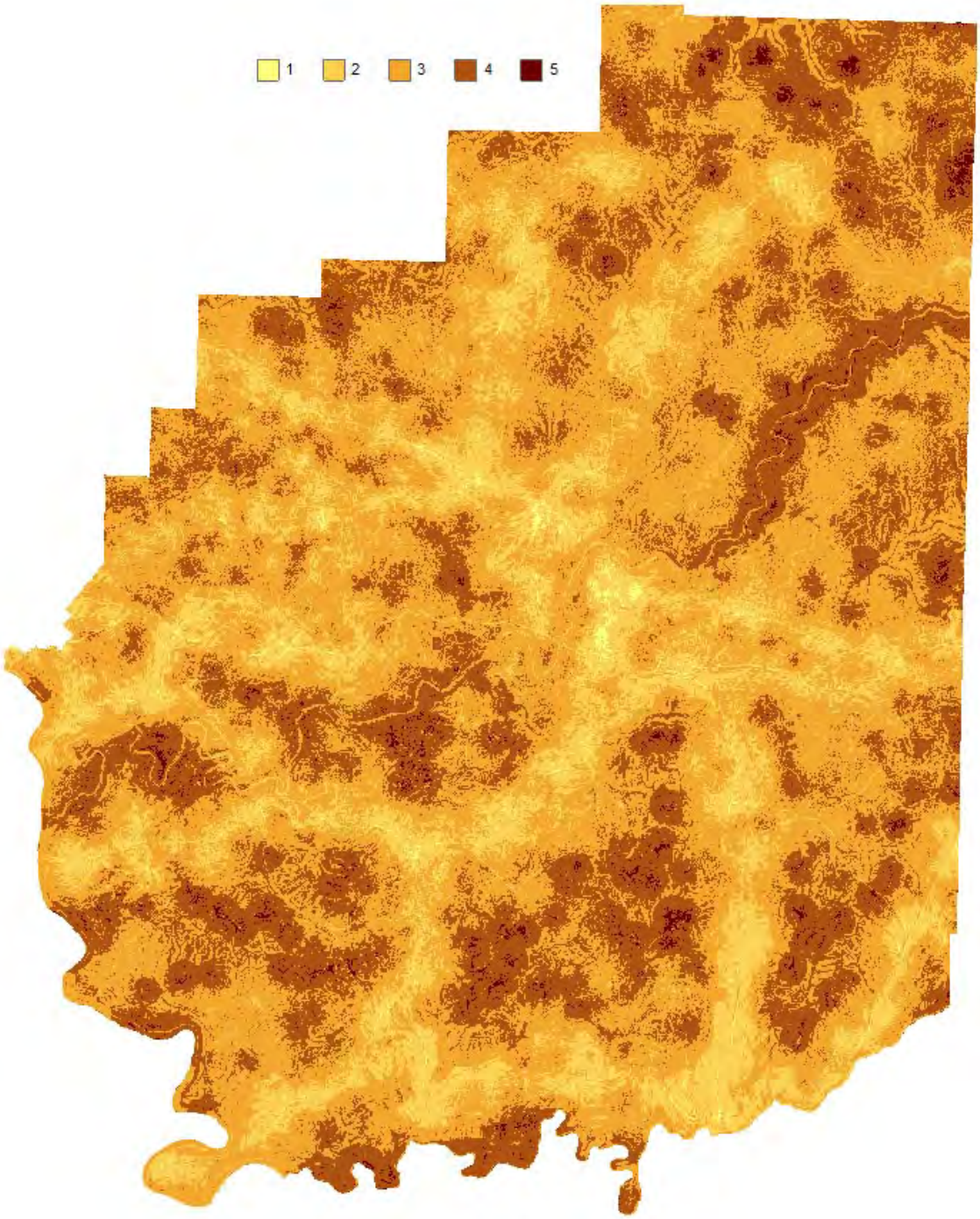
Run the tool 4 times changing the % influence accordingly. Name the resulting outputs as follows: **Weighted_Equal**, **Weighted_LCUrban**, **Weighted_LCWater**, and **Weighted_LCSlopes**. The following figure is a summary of how the weighted overlay dialogue box should look like for equally weighted inputs.



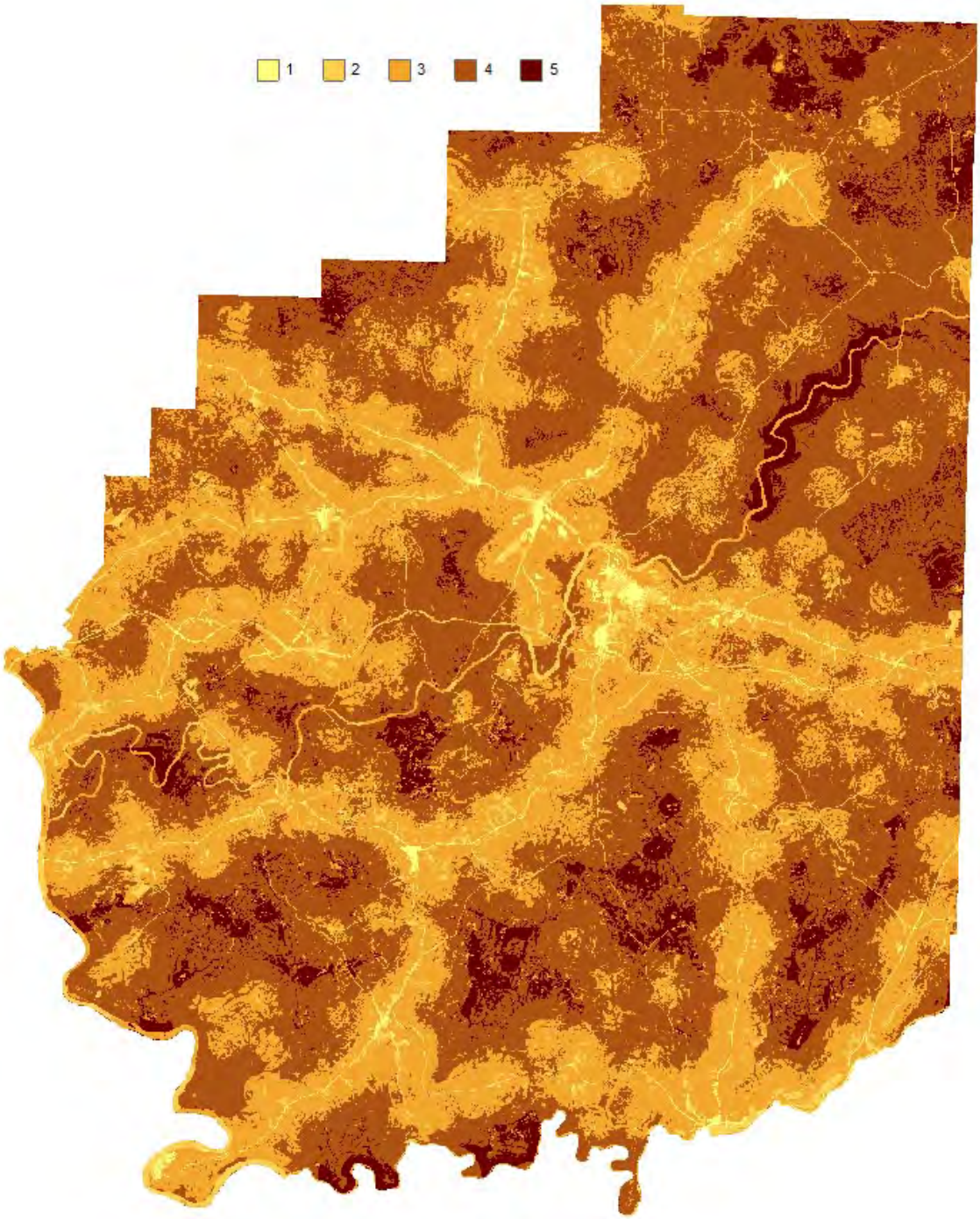
A sample dialogue box of the Weighted Overlay for equally weighted inputs.

After running each of those Weighted Overlay tasks, 4 layers are going to be added to the Contents pane. They might be showing as a stretched symbology, you'd need to change this by showing each unique value by approving of building the attribute table, since we are going to need it for further evaluating our results.

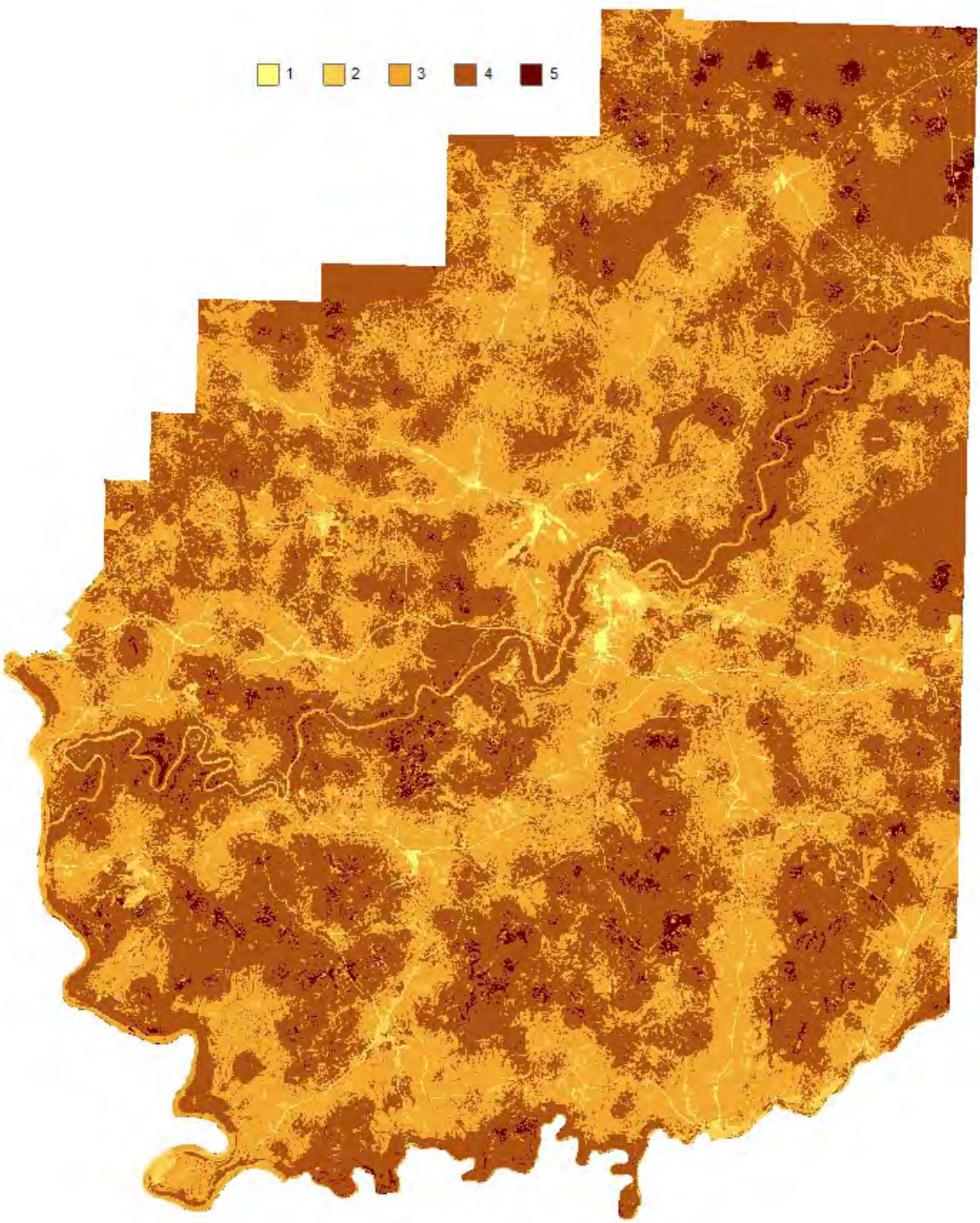
The following figures show the four weighted overlay outputs. Notice the different results that would be returned when you change the percent influence according to the previous table.



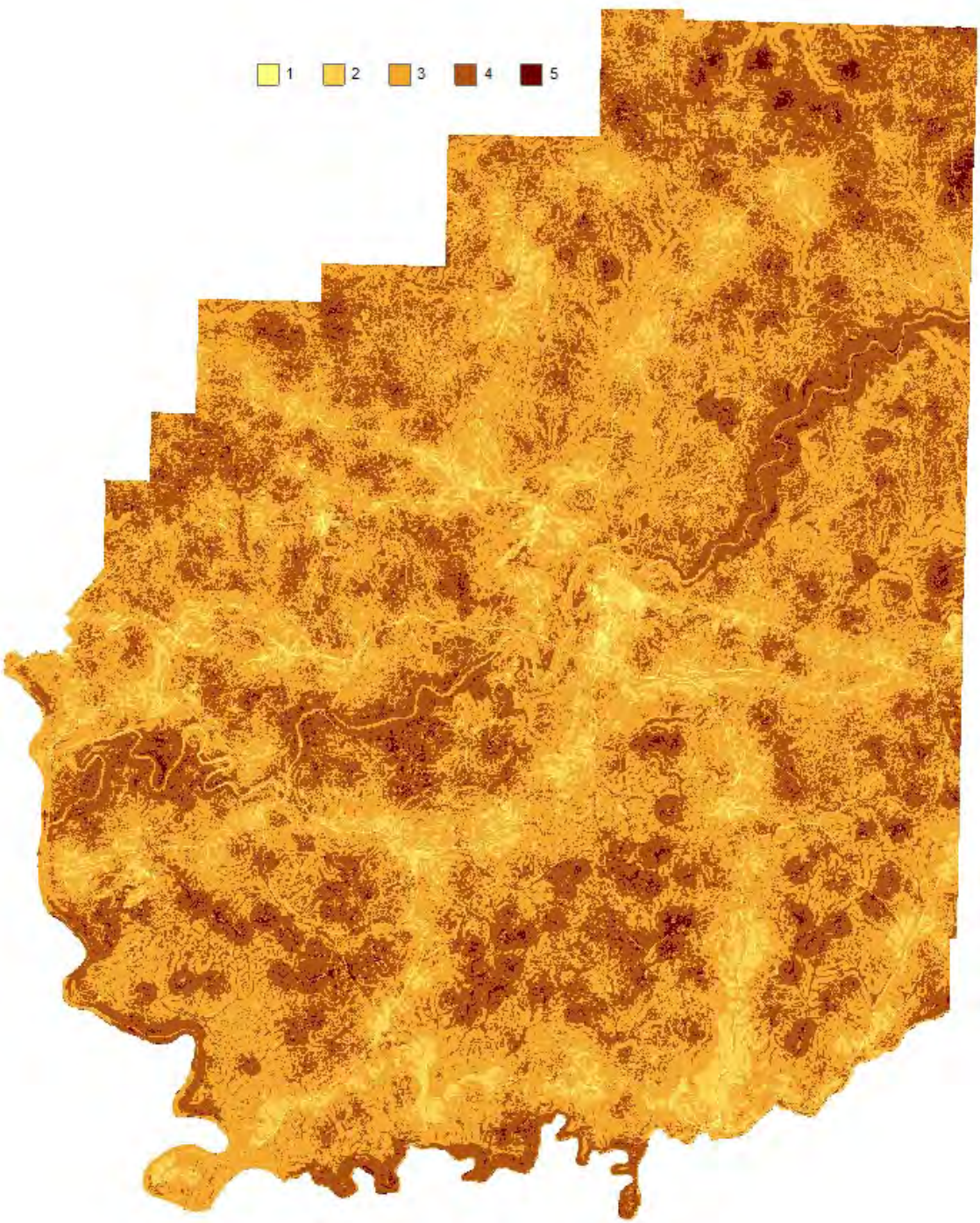
Equally distributed percent influence: By assigning an equal percent influence very few sites are identified as most suitable for a natural wildlife habitat.



Second weighting option: the result from assigning the land cover a 40% influence, the urban suitability a 25% influence and varying values for all other layers would result in relatively more sites identified as most suitable (value = 5), it also clearly defined the urban/built up areas as not suitable. The value of 2 is almost missing in this result and is substituted with the value of 4 which is leaning towards a more suitable identification.



Third weighting option: the result from assigning the land cover a 40% influence, the water suitability a 20% influence and varying values for all other layers would result in relatively less sites identified as most suitable (value = 5). The value of 2 is almost missing in this result (similar to the second weighting option) and is mainly substituted with the value of 4 which is leaning towards a more suitable identification. The value of 4 dominates this result.

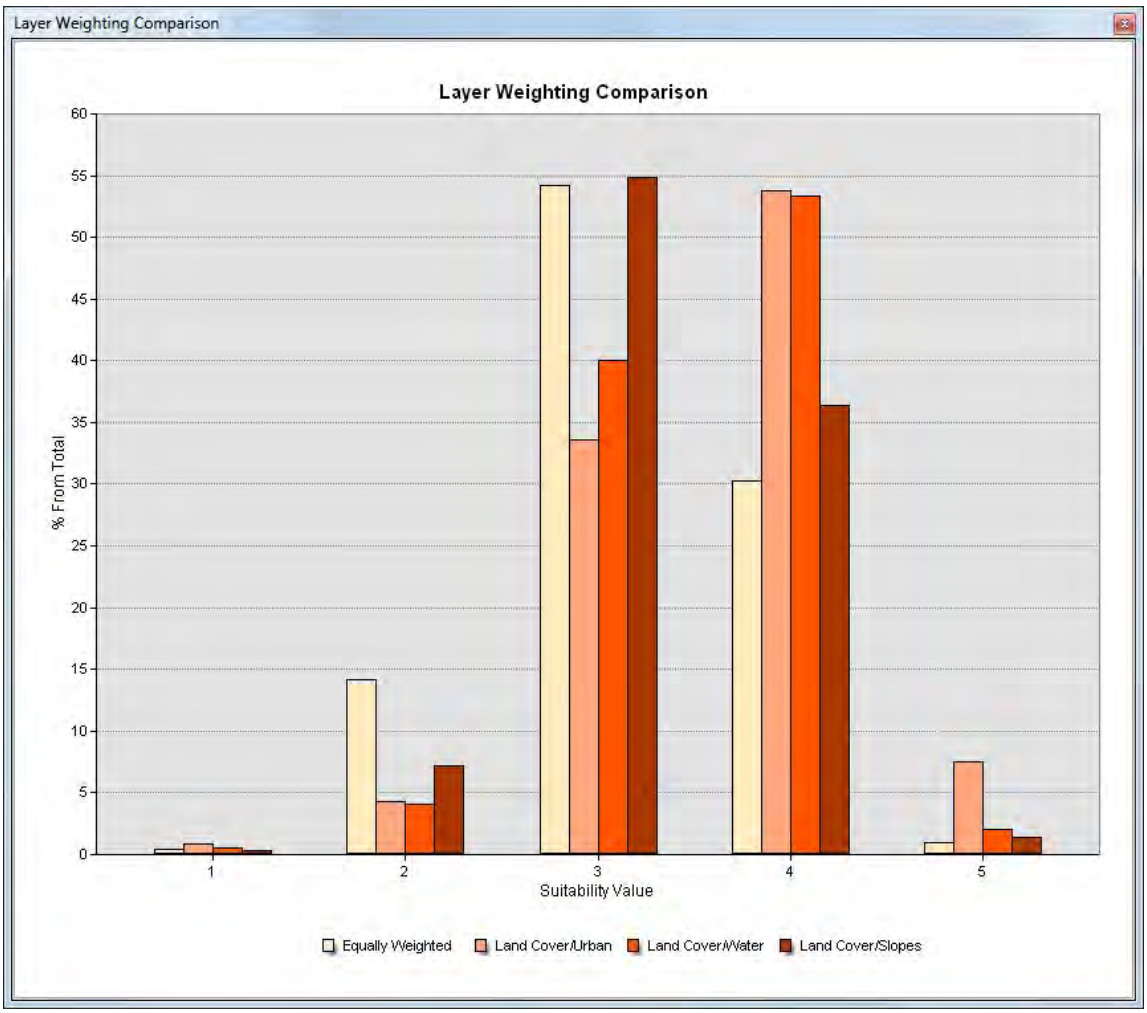


Fourth weighting option: the result from assigning the land cover a 30% influence, the slopes suitability a 25% influence and varying values for all other layers would result in much less sites identified as most suitable (value = 5), similar to the first weighting option. The value of 2 is relatively recovered and is mainly substituted with the value of 3 which is generally dominating the result. This result is very similar to the first one.

IV. Assessing the results

In order to assess the result, open each of the attribute tables of the resulting Weighted Overlays and calculate the percentage from the total pixel count in a new field; name it

PERCENT. Then add a graph using MS Excel. You may export the attribute tables of the results, organize them and create the comparative graph. A comparison graph would be very helpful to interpret the differences between the influence percentages. The following graph is an example of how to summarize the results according to the percentage from the total number of pixels for each weighted overlay.



This graph, created in ArcMap, shows the different distributions of the suitability values (1-5) according to the total number of cells in each output. It clearly shows that the least suitable cells are the least in all of the analysis with far less than 5% in each output. The suitability value 3 dominates the first and last weighting layers that give more weight to all classes as well as Slopes conditions in addition to the Land Cover suitability. The Fourth suitability value occupies most of the second and third weighted overlay results (Urban and Roads are weighted more than other suitability conditions).

You may also calculate the total areas for each of those suitability levels by adding and calculating an area field in the attribute table. The following figure is an example of such a calculation that was performed on the last weighted overlay result (Weighted_LCSlopes). Using

this table you could easily quantify the total area for each of the resulting suitability classes and report them accordingly.

Table

Land Cover/Slopes

OBJECTID *	Value	Count	Percent	SQKM
1	1	7073	0.327433	6.3657
2	2	154197	7.138307	138.7773
3	3	118435	54.827941	1065.9213
4	4	785192	36.349226	706.6728
5	5	29315	1.357092	26.3835

1 (0 out of 5 Selected)

Land Cover/Slopes

An example of an attribute table that includes the percentage from total area as well as the total area in Square Kilometers.

V. 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 (use **Clarion_Streams** feature class);
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.

Use 5 classification values for each of the input layers as well as for the evaluation scale. It'll be up to you to decide on the proper class breaks that you'd use for each of the criteria.

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 dropbox.

A. Report Content

- Include a cover page with a selected title "Raster Analysis: Weighted 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.
- Add a flowchart to summarize the steps, or you may set them as a bulleted list with explanation of details for each step (similar to this assignment).
- Extract all of the necessary information from the given data.
- Apply the different necessary re-classifications for all inputs.
- Take screen snapshots (Alt + Prnt Scrn or using the Snipping tool) of your ArcGIS Pro 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 pane with the 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 at least 3 different layer weightings for your result, compare them and add your argument for the reason of selecting such weights/percent influcences. Summarize the results in the form of tables and graphs as well as maps.

B. Grading Rubric

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

Criteria	**Score From**
Cover Page	5
Define a classification scale (tables)	15
Apply the classification to the inputs (Inputs)	10
Setting weights/percents influence and applying the tool	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