**GEOG489 Lab #2 – Suitability Modeling**

(Due date: Feb 17, 2017)

Before you begin with the exercise make sure you have downloaded all the data from:

[https://github.com/qiang-yi/GEOG489/blob/master/labs/lab2\_data/lab2\_data.zip](%20https://github.com/qiang-yi/GEOG489/blob/master/labs/lab2_data/lab2_data.zip)

Unzip the files into your U drive. Data for the three parts are stored in folder Part A, B and C. You can create a Lab2 folder to store these data. Remember to change file paths accordingly in the exercises

Part A is a tutorial on using Spatial Analyst in ArcGIS 10, with detailed step-by-step instructions. You will need to follow the instructions in **Spatial\_Analyst\_Tutorial.pdf** for this.

Download the instruction from:

<https://github.com/qiang-yi/GEOG489/blob/master/labs/lab2_data/misc/ex1_tutorial.pdf>

Part B is a fairly basic suitability modeling exercise using only vector data and vector analysis tools. No detailed step-by-step instructions are provided, assuming you should be familiar with these tools through previous exercises. The purpose of this part of the lab assignment is to explore some of the differences between vector and raster-based suitability modeling.

In Part C you will demonstrate your newly acquired understanding of suitability modeling by developing a wolf habitat suitability model for Yellowstone National Park. Several articles in the readings cover this subject matter, although the instructions provided here should be sufficient to develop a meaningful model.

**Part A - Suitability Modeling in Raster**

This first part serves both as an introduction or refresher of Spatial Analyst as well as an introduction to raster-based suitability modeling.

You are given a PDF version of Using ArcGIS Spatial Analyst. You will need to complete Exercises 1, 2 and 3 (4 is optional) of this tutorial and then answer the following questions. The data for this tutorial are stored in the Part A folder in the zip file, so do not follow the instructions to navigate to the ArcTutor folder (pages 3 and 4) but navigate to wherever you downloaded or copied the data to.

The general analysis process you will be going through is shown in Figure 1. You are free to choose to use ModelBuilder as demonstrated in the tutorial or write python scripts to accomplish the same workflow.



Figure 1.

**To include in your report:**

1. Screen capture of your ArcMap interface showing the final suitability grid (the result on the bottom of page 42), and a screen capture of your model’s dialog box (like on page 41) or python scripts you used to accomplish the same workflow (2pts).
2. Answer the following questions (answer < 150 words to each question, 1pts each):

* Towards the end of Exercise 2, the Con statement is used to select those areas with high suitability scores, and then Majority Filter is used to remove the isolated cells with high scores. Are there any issues of using this approach? If there are, please explain and suggest alternative techniques for the same purpose.
* All the variables (except land use) are reclassified using a 10 class equal interval classification, which rescales the variables into the same range from 1 through 10. Are there other techniques to rescale the variables? If so, please explain.
* What is the underlying assumption of the equal interval reclassification? Is the equal interval reclassification good for all variables? If not, please explain why and suggest solutions.

**Part B - Suitability Modeling in Vector**

In this part you will practice suitability modeling using vector data and vector analysis techniques – **do not convert your data to raster format** and **do not use any raster tools**! The example is a little bit simple, but should be sufficient to get a good feeling for two types of GIS data: raster v.s. vector.

A group of conservationists meets yearly to set conservation priorities in Northern Kentucky. Using the results of their preliminary voting, you are going to prepare a map for the group to consult during the meeting. The results of this map may affect their final voting decision.

The linear equation you will use to rank areas was determined by the conservation group's opinions. They prioritized protection of wetland areas first, agricultural lands second, forest stands third, and urban areas fourth. The preliminary voting targeted areas near water bodies for conservation efforts because of their importance to migrating birds. The group also wants to target areas that have received previous support.

The equation is: **Rating = landuse\_rating \* water\_proximity \* prior\_support**.

The data layers used in this exercise contain:

* Water Bodies — lakes and rivers in the study area
* 1 Mile Water Buffer — 1-mile buffer zones around the lakes and rivers
* Landuse — land use within the study area
* Financial Support — donations by census tract provided over the last five years for conservation

Take a few moments to familiarize yourself with all the layers and their attributes. Note especially the attributes needed for calculating the rating: the LRATING field in the Landuse layer, the BUFFERDIS field in the buffer layer (waterbuf\_1mile) and the ZSUPPORT field in financial support layer. The LRATING field contains conservation priority ratings. These ratings range from one to nine, with nine indicating the highest priority. For this analysis, wetlands, forests, and transitional lands were assigned the highest priority.

The values in the LUCODE field represent specific types of land use. Here are the land use values and their associated descriptions. They are provided here for completeness, but you don’t really need them

11 = Residential

12 = Commercial and Services

13 = Industrial

14 = Transportation, Communications, and Services

15 = Industrial and Commercial Complexes

17 = Other Urban or Built-up Land

21 = Cropland and Pasture

22 = Orchards, groves, vineyards, nurseries, and ornamental horticultural

23 = Confined Feeding Operations

24 = Other Agricultural Land

41 = Deciduous Forest Land

42 = Evergreen Forest Land

43 = Mixed Forest Land

51 = Rivers

52 = Lakes

53 = Reservoirs

75 = Strip Mines, Quarries, and Gravel Pits

76 = Unclassified

The BUFFERDIS field contains two values, 0 and 1. A value of 1 represents areas that are within a mile of a water body. A value of 0 represents areas that are more than a mile from a water body. The values in the ZSUPPORT field represent the donations (in thousands) received from each census tract over the last five years.

The three fields just described are important as they will be used to calculate the final conservation ratings.

You will now need to union your three input layers: land use, buffer and financial support. It doesn’t matter exactly in which order you process the layers, as long as you end up with one layer containing the combined geometry and attributes of all three.

Now you are ready to calculate the ratings. Add a new field called Rating to your new layer – use a precision of at least 10 and a scale of 2. Now calculate the rating values as follows:

**Rating = landuse\_rating \* water\_proximity \* prior\_support**

Use a graduated color legend to show your results for the Rating you just determined.

**To include in your report:** Screen capture of your results so far identifying conservation priorities (1pts).

Consider the math used to create the ratings and how it led to a contrast between areas inside and outside the 1-mile buffer zone. The water distance variable has only two values, 0 and 1. As the formula was designed, if a variable's value is 0, then the result of the calculation is 0. Because all areas more than 1 mile from a water body had a water distance ranking 0, all areas were assigned a 0 after the calculation was performed, thus negating the land use rating of all areas outside the 1-mile buffer zone. Thus, a 0/1 ranking represents a Boolean (Yes/No) criteria – including areas meet the criteria and excluding areas do not meet the criteria.

There are also a number of limitations of the approach used so far. Now you have a chance to make some improvements. Consider the following **options**:

* The ratings equation is a multiplication of all variables, meaning a variable with larger absolute value is more decisive to the result. Is it reasonable? If not, how do you modify the approach to solve this issue?
* This analysis uses a multiplication equation to calculate the suitability score. Are there any drawbacks of the multiplication equation? How do you solve it?
* The 0 and 1 rating for the buffer might be a bit too decisive. Sometimes people have gradually changing preference to distance to water. How do you represent such gradual changing suitability in ArcGIS?

*Note: the objective here is to practice suitability modeling using only vector data, so you get a feeling for how it’s different from working in raster. You need to make some improvements on the model presented here, but don’t make a huge effort out of this one – save that for the following part.*

*Note: there is no need to build a model for this analysis using the Modelbuilder, although it is of course possible and a good opportunity to practice. If you do want to use the Modelbuilder, you will need the tools Add Field and Calculate Field, under Data Management Tools > Fields.*

**To include in your report:**

1. Brief description of your revised methodology and a screen capture of your final conservation priorities (2pts).
2. Now that you have practiced suitability in both raster and vector, provide a solid description of the difference between the two types of suitability modeling. Your answer can include the following aspects but is not limited to them: (2pts)

* Pro and cons of the two approaches
* Analysis tools used;
* Distance functions;
* Overlay analysis;
* Ranking and weighting.

Note: this is not a short-answer question (200 – 500 words)

**Part C – Wolf Habitat Suitability Analysis**

Now that you are familiar with suitability modeling using vector and raster techniques, you will carry out a wolf habitat suitability analysis using whatever analysis techniques you think will work best. You are encouraged to use the ModelBuilder to create and revise your suitability model.

In this exercise you will need to estimate where an alpha pair of wolves, soon to be released in Yellowstone National Park, will likely establish territory. The release will be a late-winter soft release. A soft release means that the wolves will be penned in the release area for two months to allow them to become accustomed to the area in the hope they will not travel too far from the original release point. The release point has not been established yet, and will be selected using the analysis results.

The dataset to be used includes the following layers. All data is in a UTM coordinate system; map units are therefore in meters.

* **Boundary** – Boundary of Yellowstone National Park.
* **Campsites** – Major campsites, which can be reach by vehicle.
* **Bc\_campsites** – Backcountry campsites, which can only be reached on foot.
* **Trails** – Hiking trails
* **Roads** – Major roads in the park for vehicles.
* **Lcov\_yel** – Grid of vegetation types.
* **Elev\_yel** – Digital Elevation Model
* **Wolfpacks.shp** – Approximate boundaries of existing wolf pack territories.

Wolf habitat preference can be predicted with the following variables:

*Elk distribution*. The best predictor of wolf habitat preference is elk distribution, which can be estimated based on vegetation types preferred by elk. These preferred vegetation types include Aspen, Douglas Fir, and all non-forested areas (which in the case of Yellowstone National Park are primarily **grasslands**). Areas in fairly close proximity to these vegetation types can also be considered as preferred habitat areas, since wolves can easily travel over a large distance to hunt after prey.

*Human avoidance*. Wolves are more likely to be found in areas away from people. The human activities to be considered are the campsite locations and the road and trail networks. Major campsites and major roads are likely to have more of on impact on avoidance by wolves than backcountry campsites and trails.

*Elevation.* In general an important factor is that wolves are preferring mid to high elevations. All of the park area is considered above the lowest elevation where wolves are likely to occur, but elevations above 3000 meters are likely to be very low in food (i.e. elk).

*Existing wolf packs*. New wolf packs will likely avoid existing wolf packs; obviously, these are not sharp fixed boundaries, since existing wolf packs might change their territory.

**Model 1: A simple model using boolean logic**

The first model you will try is a simple multi-criteria model using boolean logic. Boolean logic in the context of multiple criteria analysis means that you will use a series of “yes/no” maps and determine the area where all criteria are met. Use the following criteria:

* Vegetation consists of Aspen, Douglas Fir, and non-forested areas;
* Distance to campsites and roads is at least 2500 meters;
* Distance to backcountry campsites and trails is at least 1000 meters;
* Elevation is between 1500 and 3000 meters.
* No existing wolf packs are present.

Produce of map of the preferred areas for wolf re-establishment.

**To include in your report**:

1. Screen capture of the preferred areas for wolf re-establishment (2pts).
2. Answer the question: How much (total area in km2) of Yellowstone National Park is considered “suitable” using this simple model? Briefly describe your analysis steps (1pts).

**Model 2: A more sophisticated model using suitability scores.**

Your task is now to improve upon the simple model by creating a habitat suitability model using scores for the variables listed above. In the final wolf habitat suitability map higher values should indicate a higher probability of suitable habitat.

When creating the wolf habitat suitability map, keep the following in mind:

* Make sure you consider whether a variable has a positive or negative influence on wolf habitat suitability;
* Component maps can be Boolean (e.g. yes/no, absent/present), categorical (e.g. high/medium/low) or continuous (values from 0 or 1 to a certain maximum with all values in between); your procedures should deal with these data types appropriately, and keep in mind they need to be rescaled into comparable value range;
* You are free to use weight variables – in general the presence of food (i.e. elk) is considered a very dominant variable;
* All lakes (water features in the land cover map) should be excluded or ranked 0 in the final wolf habitat suitability map;
* There are several distance variables – consider that beyond a certain distance the influence of distance variables is probably very low; in other words, an area 10 kilometers away from a road is likely to be as suitable as an area 25 kilometers away;
* The extent of the wolf habitat preference map should be set to the boundary of Yellowstone National Park;
* You can use the map of existing wolf packs in a couple of different ways; for example, the presence of wolf packs can be used to validate the other components of your suitability model.

The wolf habitat suitability map is not the final result of the analysis; you will also need to consider the size of potential settlement areas. Wolf packs need at least 50 square kilometers of territory. Identify the 3 most likely areas where the wolves might settle based using this area requirement in addition to the wolf habitat suitability map.

*Notes:*

* *You will need to use a combination of raster and vector analysis tools; many analysis steps are possible in both data models, and you are free to decide for yourself what data model to use for which analysis steps.*
* *You can really make this into a detailed model with lots of small variations – make simplifying assumptions as needed to get this assignment done within a reasonable effort.*
* *There isn’t a single “perfect” answer to this challenge; the objective of the exercise is to demonstrate your ability to develop a spatial model, using vector and/or raster analysis tools appropriately, and to document your analysis process and result.*
* *You are free to manipulate ArcGIS manually, use ModelBuilder or Python scripts to finish this exercises. But ModelBuilder is recommended. If you do it manually, please clearly document the workflow of your analysis, a flowchart is recommended.*

**To include in your report:**

1. Workflow of your analysis – there should be sufficient detail so that there is no ambiguity over how you completed the analysis; add comments where needed for clarification; a screen capture of your Modelbuilder interface of your final model is helpful. (3pts)
2. Map layout of the wolf habitat suitability map and the 3 most likely areas where the wolves will settle. (2pts)
3. Brief discussion of your analysis process (2pts).