## Understanding overlay analysis

Overlay analysis in Spatial Analyst is a group of methodologies applied in optimal site selection or suitability modeling. It is a technique for applying a common scale of values to diverse and dissimilar inputs to create an integrated analysis.

Suitability models identify the best or most preferred locations for a specific phenomenon. Types of problems addressed by suitability analysis include:

- Where to site a new housing development
- Which sites are better for deer habitat
- Where economic growth is most likely to occur
- Where the locations are that are most susceptible to mud slides

Overlay analysis often requires the analysis of many different factors. For instance, choosing the site for a new housing development means assessing such things as land cost, proximity to existing services, slope, and flood frequency. This information exists in different rasters with different value scales: dollars, distances, degrees, and so on. You cannot add a raster of land cost (dollars) to a raster of distance to utilities (meters) and obtain a meaningful result.

Additionally, the factors in your analysis may not be equally important. It may be that the cost of land is more important in choosing a site than the distance to utility lines. How much more important is for you to decide.

The following lists the general steps to perform overlay analysis:

1. Define the problem.
2. Break the problem into submodels.
3. Determine significant layers.
4. Reclassify or transform the data within a layer.
5. Weight the input layers.
6. Add or combine the layers.
7. Analyze.

Steps 1–3 are common steps for nearly all spatial problem solving and are particularly important in overlay analysis.

## 1. Define the problem

Overall objective:

- Where the next Arab riot is most likely to occur?
- Where the locations are that are most susceptible to revolutions in MEAN region?

The problem components include:

Population demographic: the number of people and age group in any country.

Income: the total economic output of the country as measured by GDP

Educational attainment: the number of people with formal education
Unemployment: percent of the people looking for a job without success

## 2. Break the problem into submodels

Population submodel: the best location for a riot is a country with high population. Economic submodel: the best location for a riot is a country with low GDP per capita.

Unemployment: optimal location for a riot is a country with high unemployment build on.
Education submodel: favorable country for a riot is one with high educational attainment.

## 3. Determine significant layers

The most important factors are economic, education and demographic.
For certain factors, the layers may need to be created.

## 4. Reclassification/transformation

Factors above are transformed to a common ratio scale. Common scales can be predetermined, such as a 1 to 9 or a 1 to 10 scale, with the higher value being more favorable, or the scale can be on a 0 to 1 scale, defining the possibility of belonging to a specific set.

## 5. Weight

Certain factors may be more important to the overall goal than others. Therefore, before combining the two layers, the most important criteria (economic) should be multiplied twice as much as education level.

## 6. Add/Combine

In overlay analysis, it is desirable to establish the relationship of all the input factors together to identify the desirable locations that meet the goals of the model. For example, the input layers, once weighted appropriately, can be added together in an additive weighted overlay model. In this combination approach, it is assumed that the more favorable the factors, the more desirable the location will be. Thus, the higher the value on the resulting output raster, the more desirable the location will be.

Other combining approaches can be applied. For example, in a fuzzy logic overlay analysis, the combination approaches explore the possibility of membership of a location to multiple sets.

## 7. Analyze

The final step in the modeling process is for you to analyze the results. Do the potential ideal locations sensibly meet the criteria? It may be beneficial not only to explore the best locations identified by the model but to also investigate the second and third most favorable sites.

The identified locations should be visited. You need to validate what you think is there is actually there. Things could have changed since the data for the model was created. For example, views may be one of the input criteria to the model; the better the view, the more preferred the location will be. From the input elevation data, the model identified the locations with the best views; however, when one of the favorable sites is visited, it is discovered that a building has been constructed in front of the location, obstructing the view.

Taking the input from all of the steps above, a location is selected.