Lesson 4: Vector Data Analysis – Overlay Techniques

## INTRODUCTION

In this lesson you will learn about raster data analysis and density surfaces. Map density is defined and explained with examples to further your understanding on this topic. You will learn about converting between vector and raster data and view examples of how this process occurs.

## LESSON OBJECTIVES

By the end of this lesson, you will be able to:

1. Identify vector data analysis overlay techniques.

2. Convert coverage data format to a modern GIS data format.

3. Explain how environmental settings are used to enhance data organization.

## LEARNING SEQUENCE

|  |  |
| --- | --- |
|  | |
| Required Reading | Read the following:  Vector Data Analysis - Overlay Techniques   * Coverage Data Format * Vector Overlays |
| Assignments | Complete the following:   * Lab: Vector Data Analysis - Overlay Techniques * Quiz: Vector Data Analysis – Overlay Techniques |

## INSTRUCTION

**Vector Data Analysis – Overlay Techniques**

## What are Environment Settings?

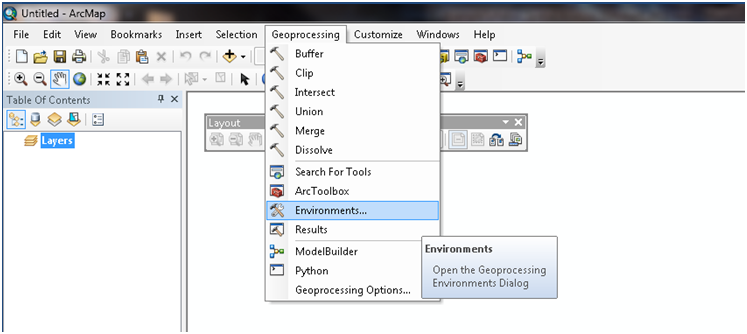
Environment settings are a set of parameters that a user will set prior to executing a tool or operation. Environment settings allow the user to:

* Determine settings for the input, processing, and output of the tool or operation.
* Set defaults for parameters such as output location, analysis extents, cell sizes, and so on.
* Remain organized as well as make tasks quicker and easier as environment settings can be shared between tools and operations.

## Location

In ArcMap, and ArcCatalog, you can access the environment settings dialog box by clicking Geoprocessing from the menu bar followed by environments.

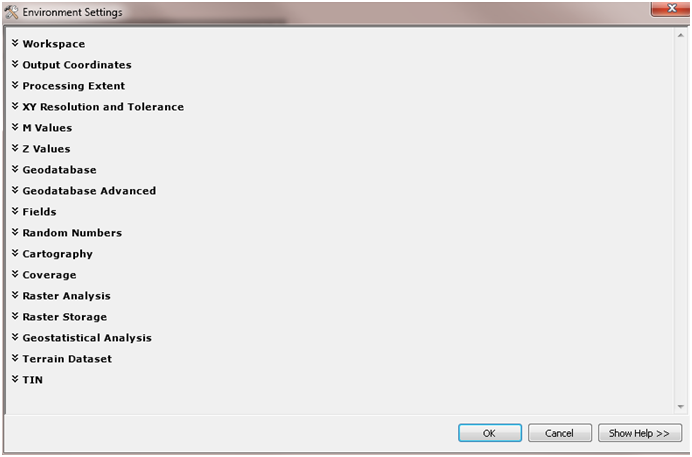
Geoprocessing -> Environments



## Parameters

This is the environment settings dialog box as opened in either ArcMap or ArcCatalog. You can click the double down chevrons to expand each environment setting and change its values. Let’s review a few environment settings.

* The workspace environment setting sets the current and scratch workspace. The workspace is where the tool will look for input values by default, and where the tool will place intermediate and output files by default.
* The output coordinates environment setting sets a geographic transformation for all output coordinates when the tool or operation is ran. You can find more information about other environment settings in ArcGIS desktop help.



Environment Settings Dialog Box

**Coverage Data Format**

## Coverage

The coverage data format stores vector data such as points, lines, and polygons. Coverage files contain the spatial and attribute data for the geographic features that it represents. Coverage is similar to a shapefile, however, the file format is quite different, and the data is stored in separate feature classes.

|  |
| --- |
| **Key Facts**  **Coverage**   * A data model that stores vector data. * The data contains the spatial and attribute data for geographic features. * Data stored in separate feature classes. |

## Storing Coverages

To store a coverage, there are a few restrictions to keep in mind. A coverage is stored as a directory containing multiple files. The name of a coverage cannot be longer than 13 characters, it cannot contain spaces, it cannot start with a number, it does not have a filename extension, and it must be named using all lowercase letters.

The reason for these restrictions is that a coverage is considered a legacy format, and was created during a time when computers and operating systems did not support longer and more complex filenames.

|  |
| --- |
| **Key Facts**  **Storing Coverages**   * Stored as a directory * Naming:   + Cannot be longer than 13 characters   + Cannot contain spaces   + Cannot start with a number   + Does not have an extension   + Must be in all lowercase letters |

## Coverage Composition

A single coverage is composed of multiple files, similar in the way that a shapefile is composed of multiple files. Each file contains information concerning one part of the entire feature class. The contents of a coverage will appear different in Windows Explorer versus ArcCatalog, and it is not recommended that you ever manually move or change any of the files inside of the coverage folder. It is recommended that you manage the coverages using ArcCatalog, as it simplifies the coverage files into a single entry in the ArcCatalog file viewer.

|  |
| --- |
| **Key Facts**  **Coverage Composition**   * Composed of a set of files * Each file contains information concerning the feature classes * Contents appear differently in Windows Explorer v. ArcCatalog |

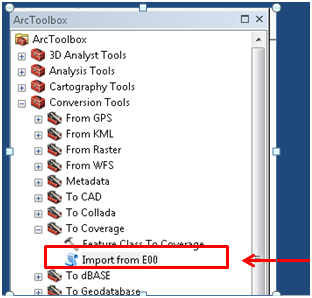
## Converting Coverages

The reason coverages still need to be learned is that quite a bit of data is still provided in this legacy format. Therefore, while it is not recommended that you save newly created data into the coverage format, you must know how to convert the coverages into a more contemporary GIS data format.

To make things slightly more complicated, coverage files are often delivered in an interchange format, known as the E00 format. In ArcGIS, a script tool is provided to convert coverages in the interchange format to a feature data set. This tool is located arc toolbox under the conversion tools toolbox, to coverage toolset, and is named import from E00.

Located in ArcToolbox:

Conversion Tools -> To Coverage -> Import from E00 (E00 being the interchange format)



**Vector Overlays**

## Overlay Operations

Overlay operations involve combining spatial and attribute data from two or more spatial data layers. In other words, we are stacking the data, and looking for where the layers overlap. While the overlay operations may seem simple when used alone, they can actually be quite powerful operations when combined in series.

One thing to note about overlay operations is that overlays require that the data be in the same coordinate system. This is so that the software can perform the highest accuracy of overlay. Some software packages automatically re-project on-the-fly if the data is not in the same coordinate system, but it is preferable if you convert all data sets to the same coordinate system before performing the overlay operation.

An example of an overlay operation would be the clip operation. For example, say I have a data set that contains all the streets in the United States of America. However, I only want to see the streets in Arkansas. I can use the clip operation to cut out all the streets that are only in Arkansas by providing the state outline of Arkansas as the clipping feature.

## Vector Overlay

A vector overlay involves combining point, line, or polygon geometry and their associated attributes. All overlay operations create new geometry and a new output geospatial data set. You should be cautioned that with certain overlay operations, very large attribute tables may result if the overlay operations combine many layers that each possess.

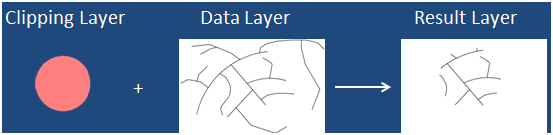
Additionally, it might be possible that the combined attribute tables would cause duplicate attribute fields to exist. In these cases, you should consider reducing the number of transferred attributes to the minimum required, and renaming duplicate fields so that there is no ambiguity.

## Basic Cases of Overlay - Clip

The first basic case of overlay is the clip function. The clip function defines the area for which features will be output based on a “clipping” polygon. I like to think of it in terms of cookie dough and a cookie-cutter. The cookie dough is the data layer that we want to reduce in size based on our cookie-cutter.

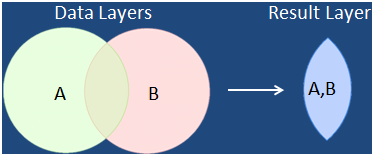
The clipping layer is the cookie-cutter that will determine how we cut out the part of the cookie dough we are interested in. It is important to note that in the clip operation only the geometry and attributes of the data layer are transferred to the result layer.

Neither the geometry nor attribute of the clipping layer are included in the output, just as you would not make your cookies in the oven with the cookie-cutter still surrounding the cookie dough.



## Basic Cases of Overlay - Intersection

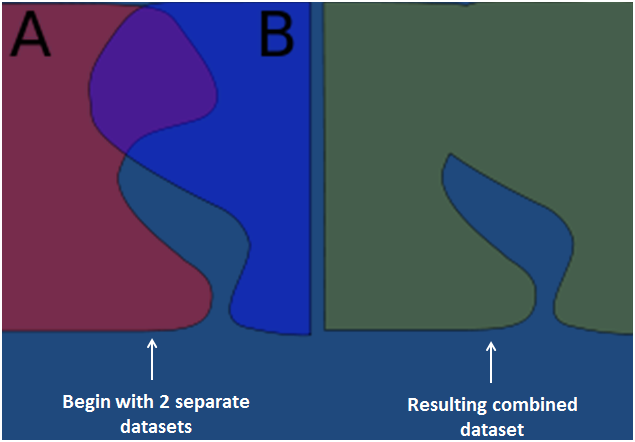
The next basic case of overlay is the intersection operation. The intersection operation combines data from two or more input layers, but the combination only exists for the regions where all input layers contain data. This is similar to the clip function; however the attributes and geometry of all input data layers are transferred to the result layer.



## Basic Cases of Overlay - Union

The third basic case of overlay is the union operation. The union operation is an overlay that includes all data from all of the input data layers. Since the union operation combines all attributes, and all geometry, no geographic data are discarded in the union operation, and all corresponding attribute data are saved for all regions.

Here is an illustration of a union overlay where we initially had two separate data sets, A and B, that, when union is applied, result in a single combined data set.



## Vector Overlay Problems

When performing vector overlays, there are a few problems that you should be aware of. In the case where the input features represent common features that are represented in both layers, they may have slightly different geometry. This creates sliver polygons when the overlay operations are performed.

**Example 1**

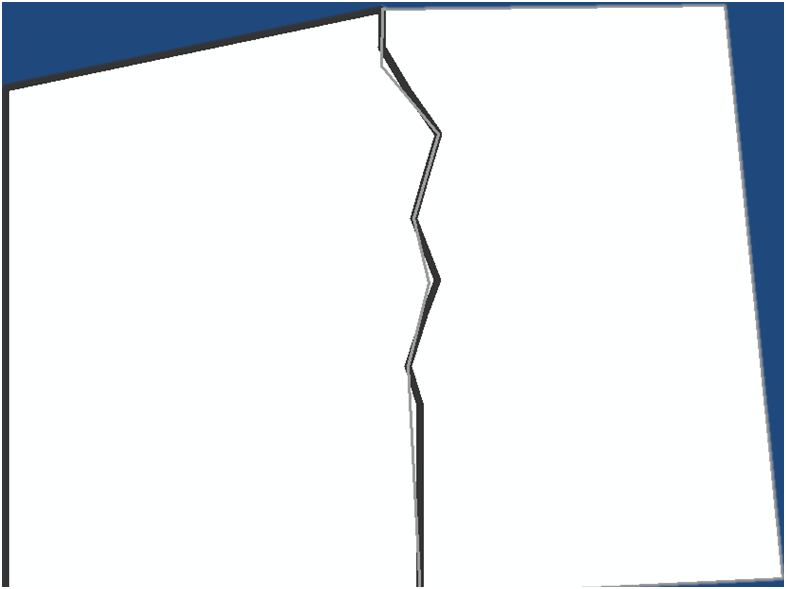
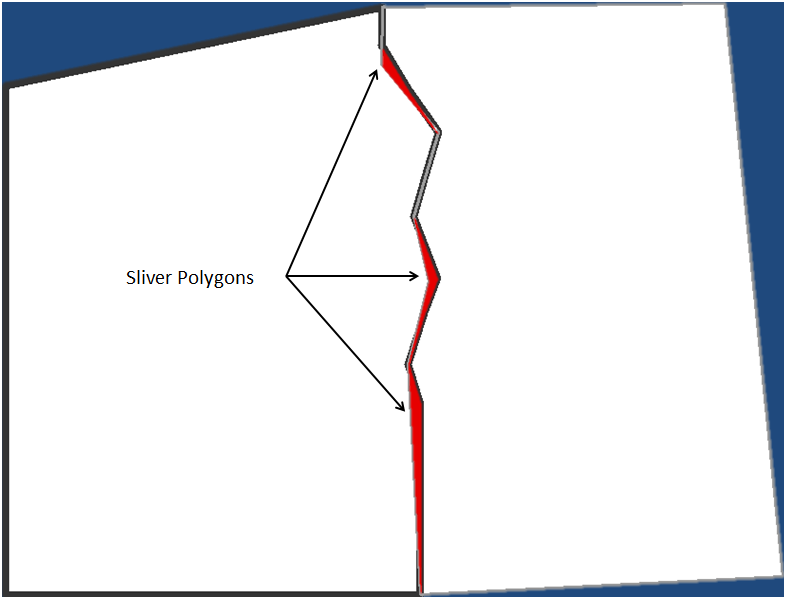
Imagine we have one input data set that represents the national boundary of Canada created at a 1 to 5,000,000 scale, and that we have a second input data set of the national boundary of the United States of America created at a 12 to 1,000,000 scale. If we stack these two layers on top of each other, it is probable that the boundary lines will not overlay exactly, as they were created at different scales, and therefore the Canadian boundary line has probably been simplified more than the United States boundary line. The areas where the two input data sets do not match will result in very small and typically insignificant sliver polygons that need to be dealt with.

There are several methods that exist to reduce the occurrence of sliver polygons, such as manually snapping the boundaries together, or introducing a fuzziness to the edges of the overlay operation, or simply manually deleting the sliver polygons.

**Example 2**

To illustrate the idea of a sliver polygon, here are two different data sets, one outlined in a bold black line, and the other outlined in a thinner great line. Where the two data sets overlay, they do not match exactly for whatever reason. If we perform an intersect, clip, or union operation on these two input data sets, we are going to create three sliver polygons.

The sliver polygons created between these two input data sets are highlighted in red. The sliver polygons are probably insignificant, and are the results of data sets being produced by different agencies, at different times, and probably at different scales. It is up to you to determine how you would like to solve for the sliver polygons.

## SUMMARY

In this lesson you learned about raster data analysis and density surfaces. You learned why and when density maps would be created with examples to help you understand. Point density, line density, and kernel density were illustrated to show you the difference of each and how they are used. This lesson also explained converting between vector and raster data with images to show the processes. The conversion from polygon to raster was also covered including the three methods which are maximum area, maximum combined area, and cell center. The raster to vector conversion process was also explained with examples to further your knowledge on how this process occurs.

## ASSIGNMENTS

1. Lab: Vector Data Analysis - Overlay Techniques

2. Quiz: Vector Data Analysis – Overlay Techniques