Chapter 18 Creating a New Suit from Old Cloth: Manipulating Raster Mode Cartographic Data

This chapter will be dèjá vu all over again except that in this case you will be performing the operations with raster rather than vector layers. The motivation is quite the same as in the previous chapter. You will gain far greater utility from vector and raster spatial data layers if you have a set of tools that enable you to manipulate those data layers so that they better suit your needs.

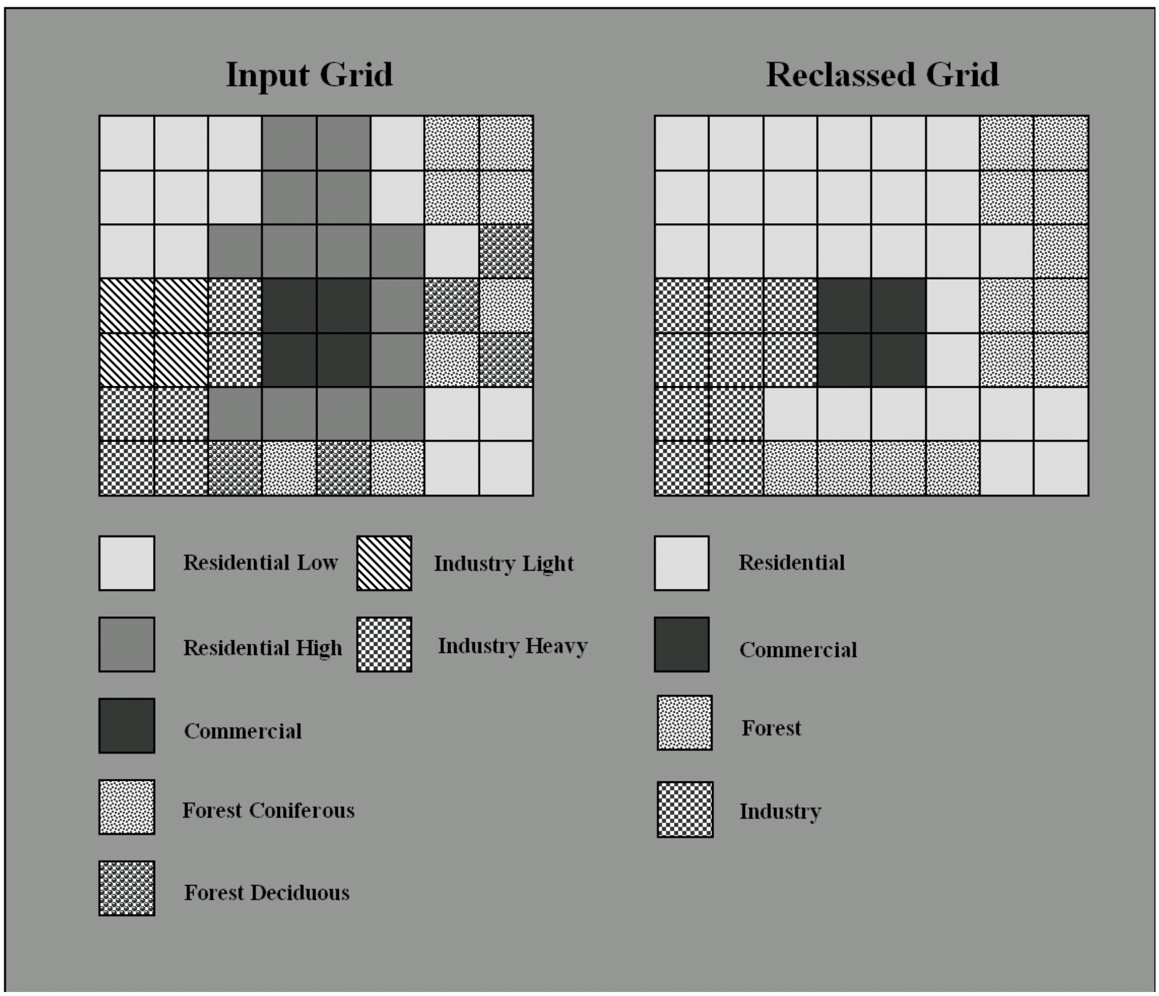
Here we will be working with the Spatial Analyst's Raster Calculator to manipulate grid files. At the out set let me emphasize that the material we review here is represents but a small portion of the database manipulation capabilities available in the Raster Calculator. The goal is, first of all, to provide you with a set of tools for manipulation of grid data that work in parallel to those available for vector data that you learned to use in the previous chapter in the Arc Toolbox. A second goal is to introduce you to the style and conventions of the Raster Calculator's data manipulation tools so that you can easily move beyond the material introduced here.

To help you get oriented, in the following table I have listed the Arc Toolbox operations, with which you are already familiar, along with their equivalent Raster Calculator operations.

Vector Mode Operation: in Arc Toolbox	**Raster Mode Operation: Raster Calculator**	**Comment**
Dissolve	Reclassify	In a grid there are no polygon boundaries, but there are class boundaries. You can use the Reclassify operation to combine classes [e.g., combine different types of forest into a single forest class].
Append	Merge, Mosaic	Exactly the same idea [almost] as append in the Toolbox.
Clip	Selectmask	Exactly the same idea as clip in the Toolbox.
Intersect	Logical AND	You already know how to do this from the introductory course, so we will not repeat it here.
Union	Logical OR	You already know how to do this from the introductory course, so we will not repeat it here.

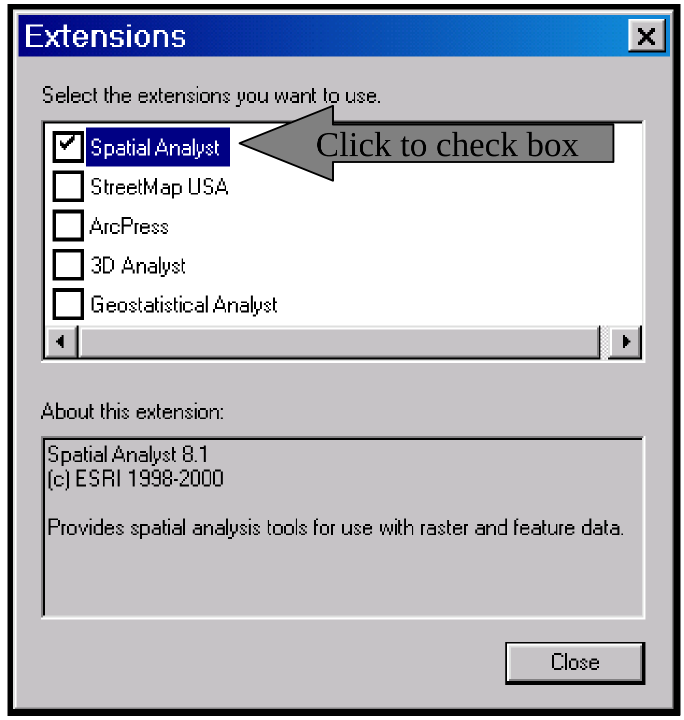
Reclassify. You are already familiar with this operation in its role as a very commonly used part of analytical procedures. You can also think of it as a database management operation because it enables you to create a new grid layer from an input grid layer. The reclassify operation enables you to alter individual values or groups of values on an input layer and save the altered [reclassed] data values on an output layer. For instance, you can use reclassify to combine groups of existing data values in the input layer and then save each group as a single data value in the output layer.

The following illustration provides and example. In the Input Grid layer there are seven land use categories. Notice that for the residential, forest and industrial uses there are sub-types, for example industrial land uses are divided into heavy and light. An analyst might require a more generalized classification that combines the detailed categories, for instance heavy and light industry become simply industrial. The Reclassed Grid layer shows the result of performing the reclassify operation. The analyst has collapsed each of the more detailed land use categories into a single more general category. From the standpoint of database management, this is the raster mode equivalent of the vector mode dissolve in which you remove boundaries between polygons that have the same values for a specified attribute. In this case, the analyst is simply reclassifying multiple values on the input layer into single values on the output layer, which is conceptually very much like removing a polygon boundary.

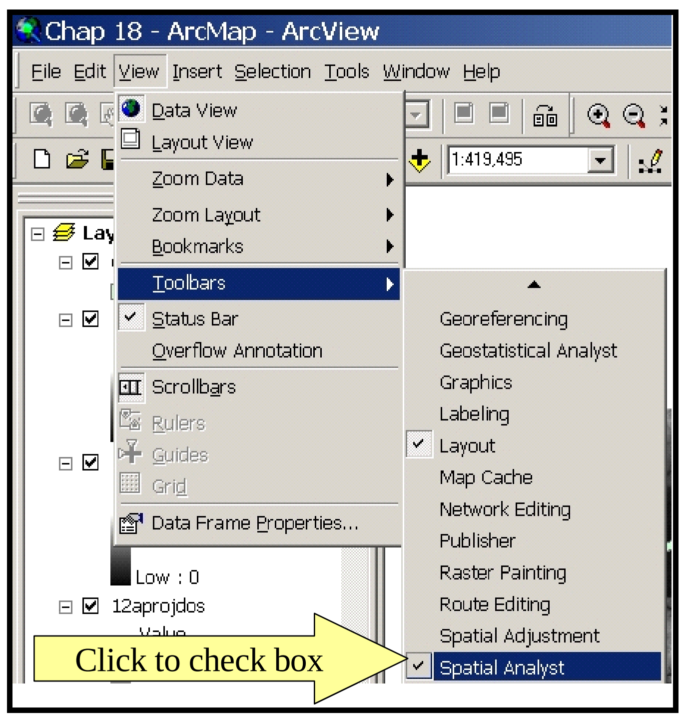


A Few Preliminary Operations. Before you can begin performing the reclassify operation in the Arc software, you need to perform a few operations required to load the Spatial Analyst extension, which is where nearly all raster mode operations take place. Here you will learn how to load the Spatial Analyst and how to display its toolbar. Additionally, you will learn how to specify the directory in which you wish to store the new grid layers you create, the default extent of the grids, and their cell size.

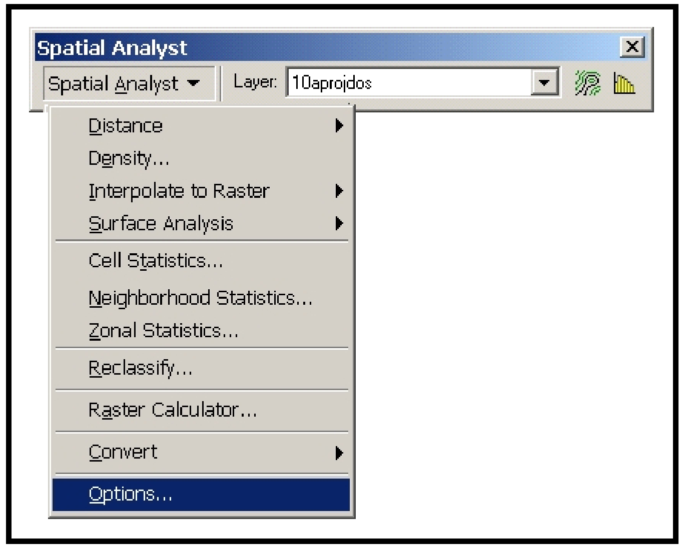
***Loading the Spatial Analyst and Opening the Toolbar*:** In order to work with grid layers, you will need to use the ArcMap Spatial Analyst extension. If the extension is not loaded, then you will need to add it. To load the extension, click on Tools in the Main Menu to open the drop down menu. Select Extensions from the drop down to open the Extensions dialog. In the dialog click to place a check in the box to the left of Spatial Analyst and then click the Close button. You now have access to the analytical features of the extension.



To use the Spatial Analyst, you need to open its toolbar. Click on View in the Main Menu and point to Toolbars on the drop down menu. The program will open a slide out menu listing toolbars. To open the Spatial Analyst's toolbar, click to place a check in its listing in the slide out menu. If you see that the Spatial Analyst already has a check next to its name, do not click as this will close an already open toolbar.

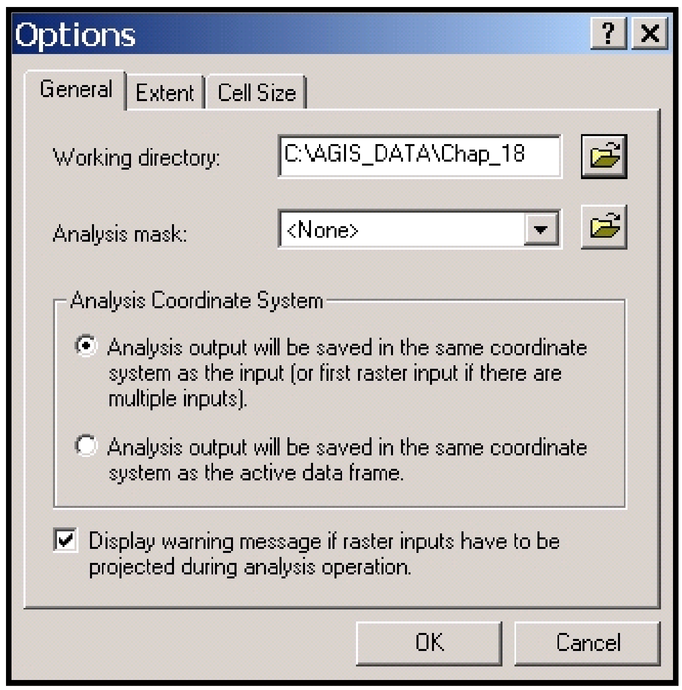


Setting the Working directory and the Analysis Environment: In working with grid data in the Spatial Analyst, you will create many temporary data layers and some new permanent layers. By default Arc Map will store these layers in C:\Temp. This is most definitely NOT where you want to store these files for the simple reason that if you were to continue the analysis on another computer, the files would not be accessible to you. You also need to tell the Spatial Analyst some things about the spatial extent of the to the analysis you will perform and the cell sizes in raster layers you create.

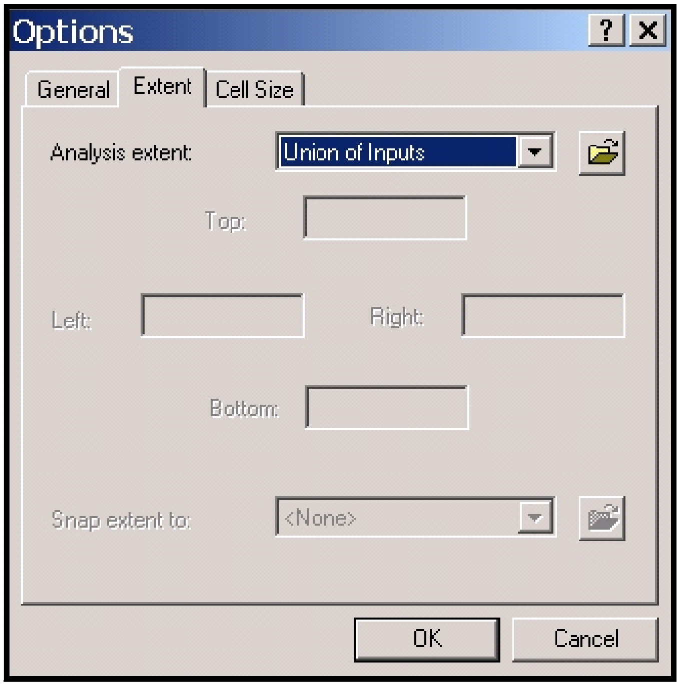


To specify the working directory, click the arrow located on the Spatial Analyst toolbar to open a menu. Click Options in the menu to open the Options dialog, which is the place where you will perform these preliminary operations.

To specify the working directory, click the General tab in the Options dialog and either type the drive and folder name you want for the working directory or click the browse button located to the right side of the Working directory field and navigate to the directory. If the directory does not yet exist you can create it within this dialog by typing the name you want and clicking the OK button. The program will open a dialog that asks if you want to create the directory. Tell it yes. In the image, I have specified the working directory as C:\AGIS_DATA\Chap_18. In your case, you might want to specify H:\Temp18 or something like that.

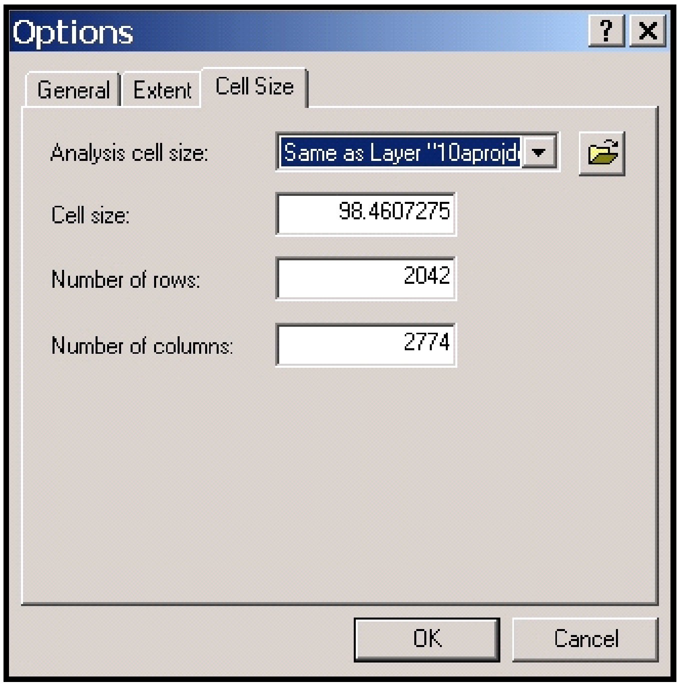


To specify the analysis extent, while still in the Options dialog, click the Extent tab. In this dialog you tell Arc Map how to determine the spatial extent of the analysis you will perform. By default the analysis extent is set at the intersection of the layers you have loaded. This means that analytical functions will affect cell values only in positions occupied simultaneously by all layers in the Data Frame. There are additional options:



1. Union of Inputs. Cells occupied by any of the input layers will be included in the analysis.
2. Same as Display. All cells that appear in the Data View window are included in the analysis.
3. As Specified Below. You type in the coordinate values describing the top, left, right and bottom of the analysis window.
4. Same as <layername>. Use the extent of one of the layers in the data set as a model.

For the operations we will be performing here, you should set the analysis extent to Union of Inputs.



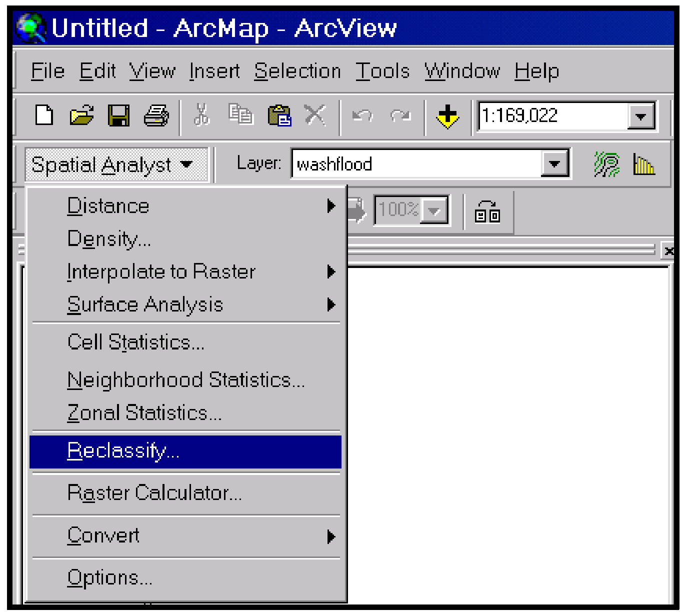
To specify a cell size, while still in the Options dialog, click on the Cell Size tab. Here you can type a cell size you want to apply to new grids, or you can specify the number of rows and number columns and let the program determine the cell size. You can access additional options by clicking the Analysis cell size field:

1. Maximum of Inputs. This option sets the cell size of a new grid to the cell size possessed by the input grid with the largest cell size.
2. Minimum of inputs. Assigns a cell size to calculated grids equal to the smallest size of input grids.
3. Same as <layername>. Assigns a cell size equal to that present in another raster layer either currently in the Table of Contents or located in another database. You use the browse button to specify the name of a layer in another database.

When you have specified the location of the working directory, the analysis extent, and the cell size, click the OK button to register your selections and close the dialog.

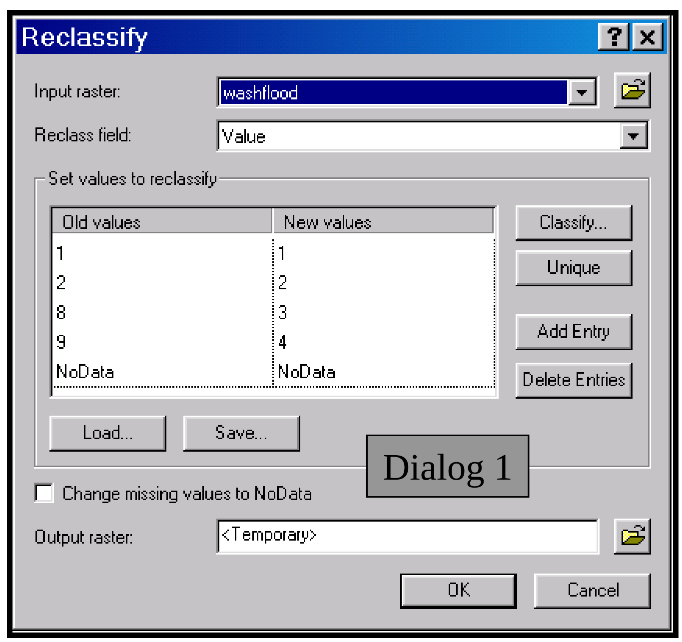
Reclassify Operation. The raster mode equivalent of dissolve is reclassify. You will remember from the previous chapter that the Arc Toolbox dissolve operation enabled you to remove common boundaries from polygons that share an attribute value in a field specified by you. For example, if for a polygon layer you created a field in which all data values are identical, then in the dissolve operation you could use the values in that field to remove all interior boundaries. You used this type of dissolve to create a county layer by removing all interior boundaries from a layer containing census tract boundaries.

You can also use dissolve to selectively eliminate boundaries. For example, suppose that you are carrying out an analysis in which you wish to differentiate areas that are flood prone and areas that are unlikely to flood. In the flood prone data layer you have available, there are four attribute values: 1 = USGS floodprone areas, 2 = other flood prone areas, 8 = areas of water, and 9 = areas that are not flood prone. If you created a field in which all flood prone and water polygons have one value and all non flood prone polygons have another value, for instance zero for flood prone areas and one for non flood prone areas, and then perform the dissolve operation using the field containing the zero and one values, the result would be a new layer in which there are only two polygon types: not flood prone and some type of flood prone or water. Combining classes in this way is exactly what the Reclassify operation does for raster data.

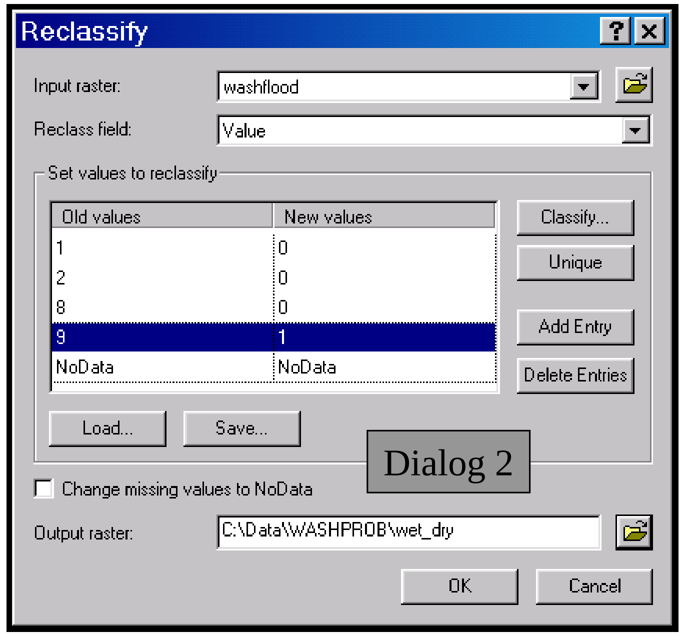


To begin the Reclassify operation, click the arrow in the Spatial Analyst toolbar to open the menu and select Reclassify. The program will open the Reclassify dialog.

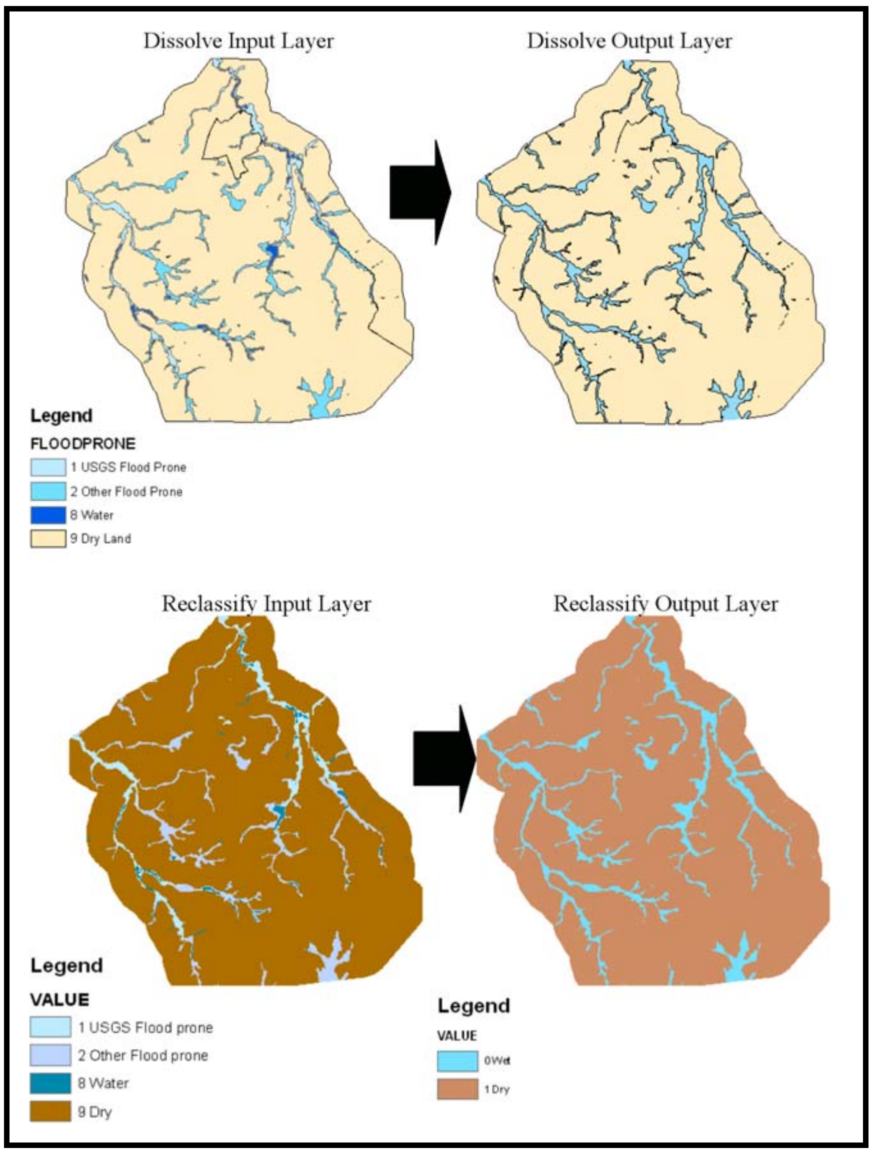
Dialog 1. In the first dialog you will see the Old Values [i.e., existing values] listed in the left portion of the Set values to reclassify window along with default New values to be assigned in the new layer. If more than one old value appears in one of the lines of the Old values list, click the Unique button. This will place a single [unique] old value in each line of the Old values window. Given a small number of classes, the easiest way to proceed here is to click in each entry in the **New values** field and type the appropriate value to replace the value listed in the **Old values** field. You want to specify a new value of zero for ranges of old values that are flood prone or water [values 1, 2, and 8] and a value of one for areas that are not flood prone. Use the Output raster field to specify a name and location for the new layer you are creating.



Dialog 2. In this dialog you can see the New values and the Output raster fields as they will appear after you have edited the dialog. To complete the operation, click the OK button. The program will enter the new grid layer in the Table of Contents.



The upper portion of the illustration depicts the application of dissolve to the vector mode flood prone data. The lower portion depicts the same result using the reclassify operation on a raster mode version of the same data set.



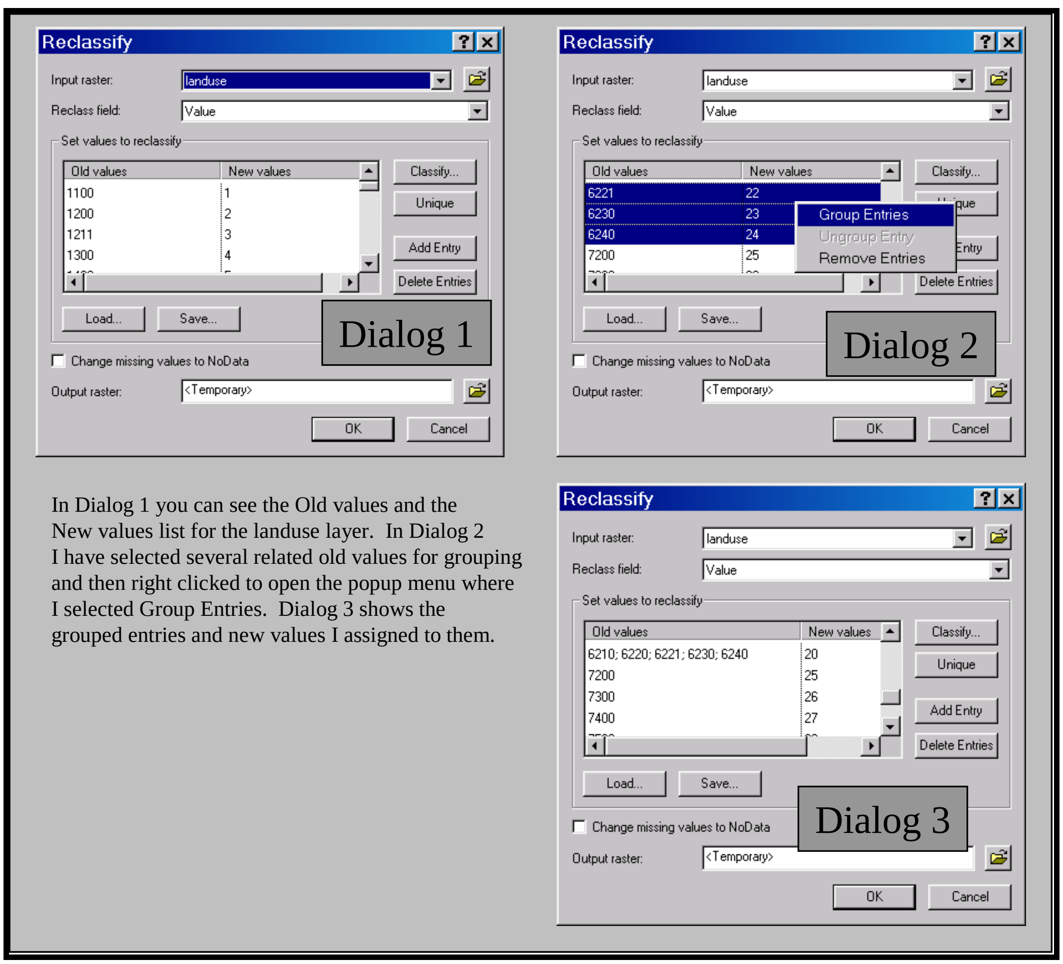
The reclassify technique just outlined works fine for categorical data sets in which there are relatively few unique values. In cases in which there are a large number of values, alternative approaches will enable you to work more efficiently. One thing you can do is group entries of old values into a single new value by selecting groups of old values to include in a single new value. To begin this process, open the Reclassify dialog and then carry out the following steps:

1. Dialog 1 in the illustration. Click on the Unique button to list a single value on each line of the Old values field.

2. Dialog 2. Click the first old value that you want to be a part of the grouped set. Next, if all of the old values you want to group are listed contiguously, hold the Shift key down and start clicking the list members you wish to include in the group and then right click in the New values field. The program will open a popup menu. In the menu click Group Entries.

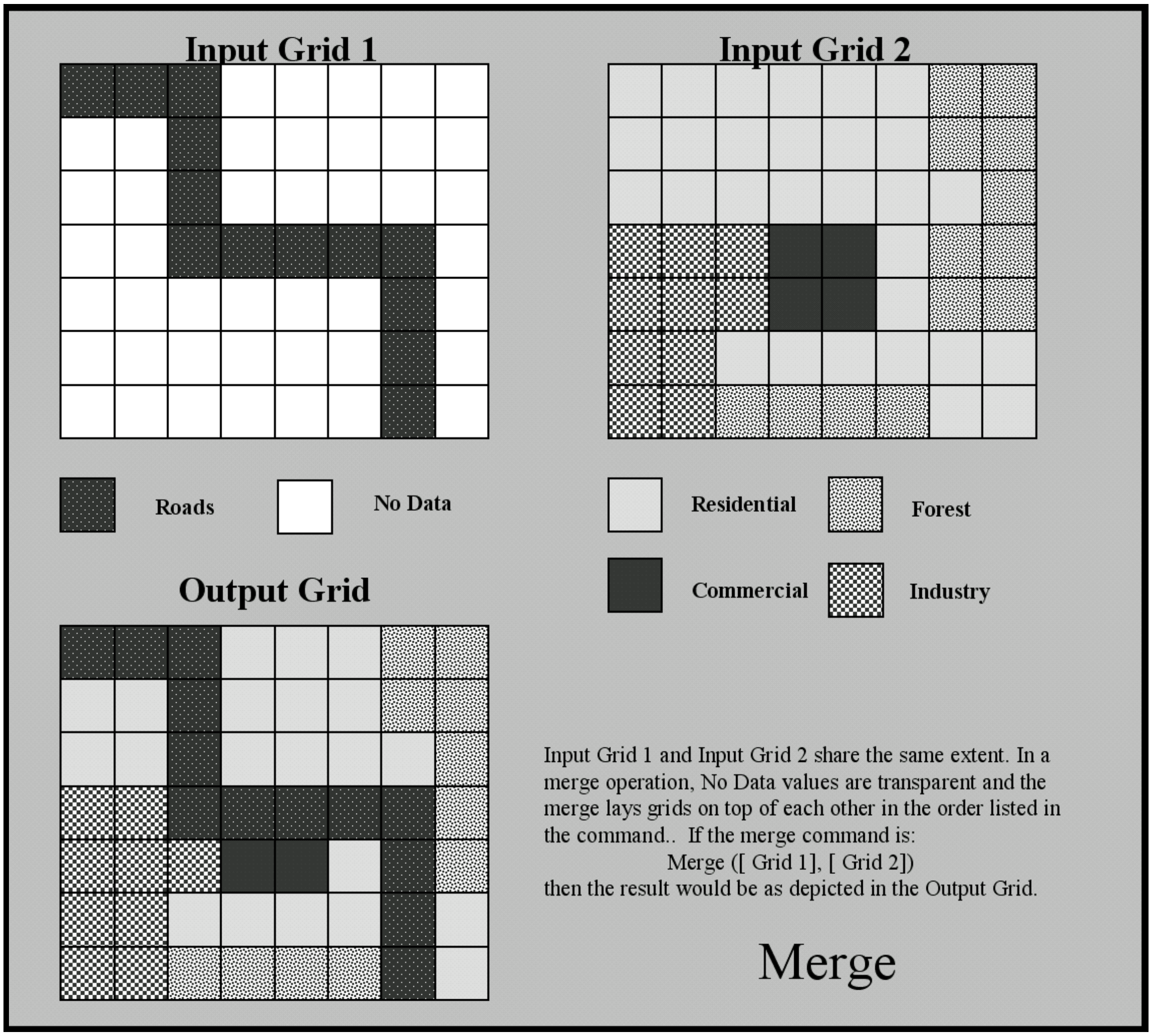
If the list members that you wish to group are not contiguous in the listing of old values, then begin by clicking the first item on the list that you want to include in the group and then hold the Control key and click the remaining members of the group. After you have selected all items that you wish to include in the group, right click in the New values field to open the popup menu and click on Group Entries.

3. Dialog 3. Depicts the result of the grouping of the entries I selected in Dialog 2. You can continue in this fashion until you have completed all grouping operations. Finish the reclassification by clicking the OK button

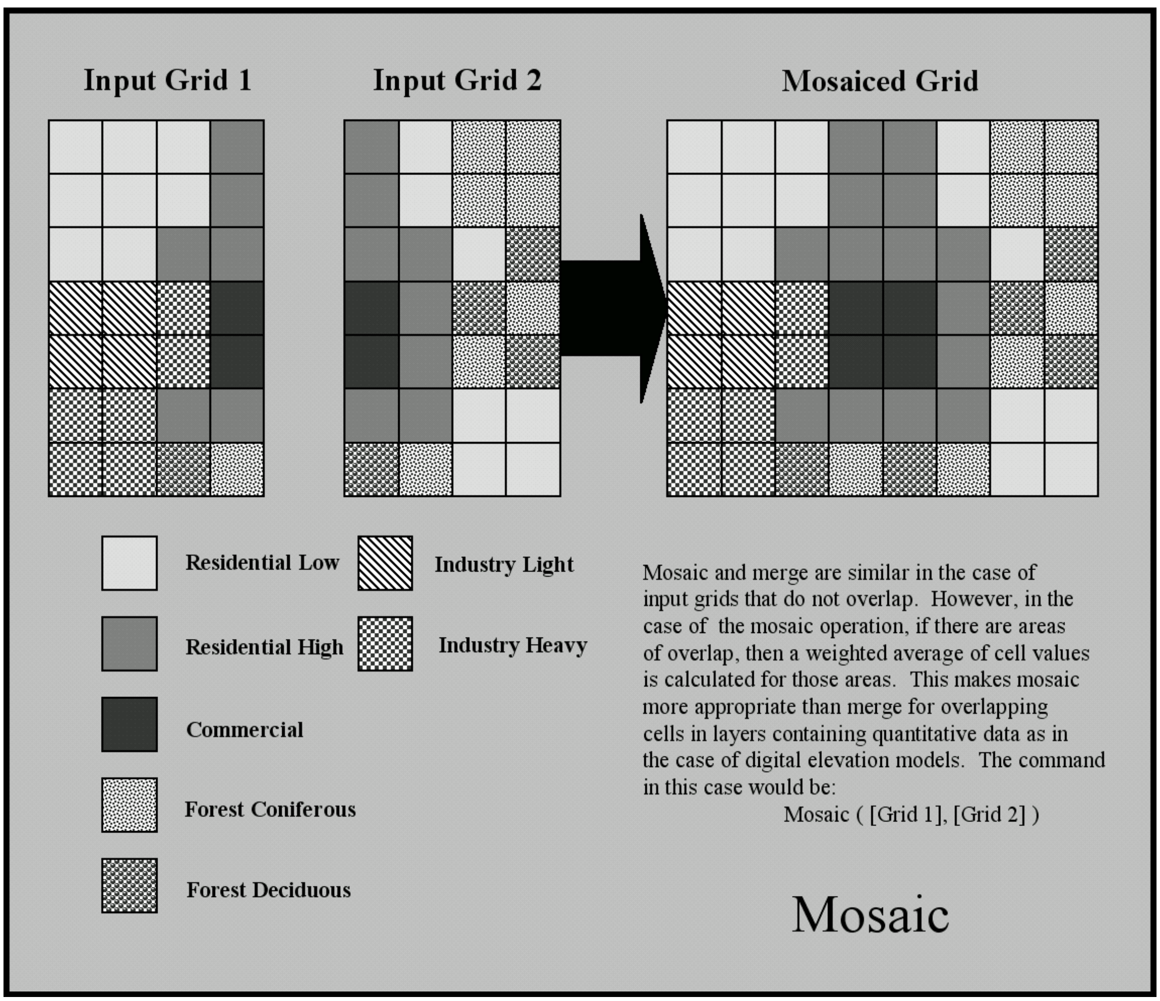


In Dialog 1 you can see the Old values and the New values list for the landuse layer. In Dialog 2 I have selected several related old values for grouping and then right clicked to open the popup menu where I selected Group Entries. Dialog 3 shows the grouped entries and new values I assigned to them.

Merge and Mosaic. You will use the merge and mosaic operations to combine grid layers. In some application situations, there will be little or no difference in the results whether you use one or the other. Merge joins grids, all of which must be in the same coordinate system, that overlap totally, partially, or that are adjacent. With Merge all No Data cell values are transparent. In the case of grids that overlap, the order of overlaying is determined by the order in which the analyst enters the grid layers into the Merge statement: grids listed earlier in the list are "on top" of grids listed later. In the output layer, the data value of each cell will equal the first non No Data value encountered. In the case of overlapping cells, this means that the first layer with a non No Data value will determine the data value of the corresponding cell of the output layer.

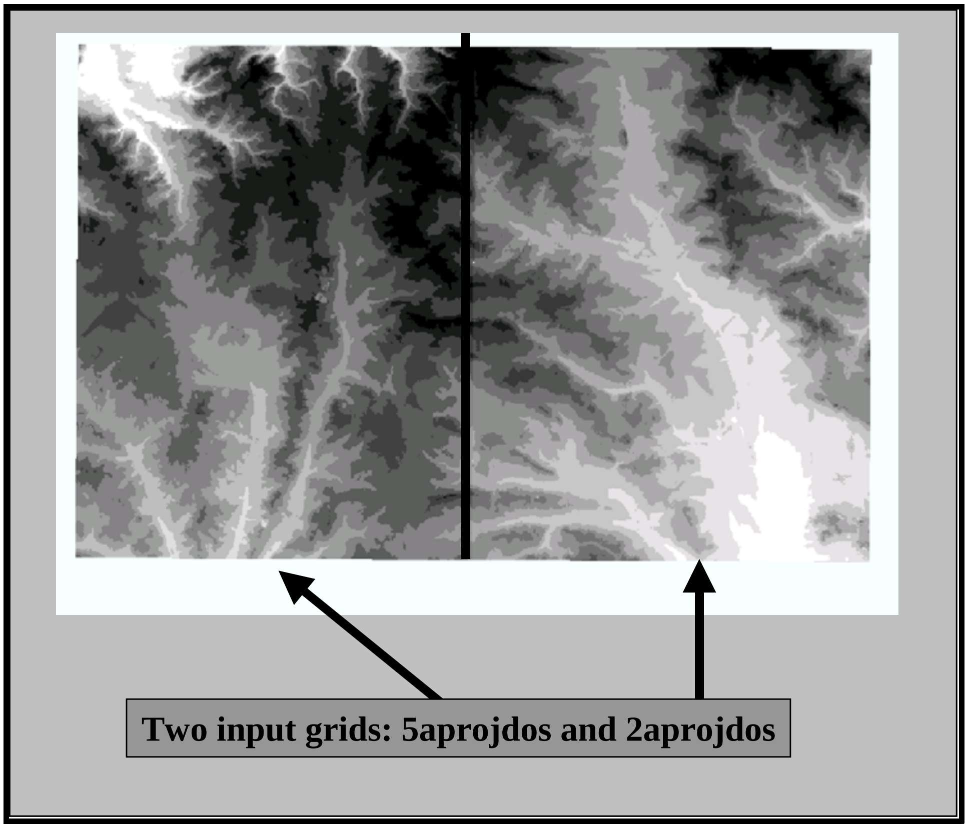


In many cases, especially if the input grids have no overlapping areas, these two operations will produce nearly the same results. However, in the case of grids that do overlap, the mosaic operation produces a result that is different from that of the merge operation. In the case of the mosaic operation the program calculates the output data value for overlapping cells as a distance weighted average of the input cell values in the overlapping area. In this case, the order in which the analyst enters the grid layers into the statement has no effect of the result of the calculation. Clearly, you will not want to use mosaic in the case of overlapping grid layers that contain categorical data. Calculating a distance weighted average for land use codes will produce totally meaningless results. However, in the case of quantitative data set calculating such an average is reasonable.

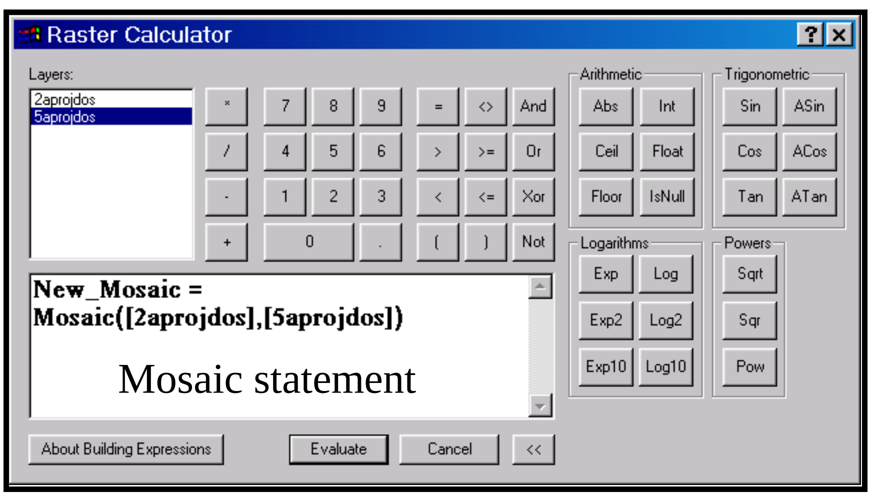


Mosaic Operation. To merge two or more adjacent layers, a set of digital elevation models for instance, begin by setting the options in the Spatial Analyst. Specifically, set the output folder to the place where you wish to keep the new layer you will calculate. At this point you should load the grid layers you want to mosaic. In this case I have added two elevation layers: 2aprojdos and 5aprojdos. Note that if you have set the default Spatial Analyst directory in the Option dialog, then you do not need to load the grids into the table of contents. You can simply type their names into the expression box of the Raster Calculator. As I can rarely remember file names, I find it easier to add the layers. In the case of very large layers, this approach can take longer as you have to wait for the layers to load.

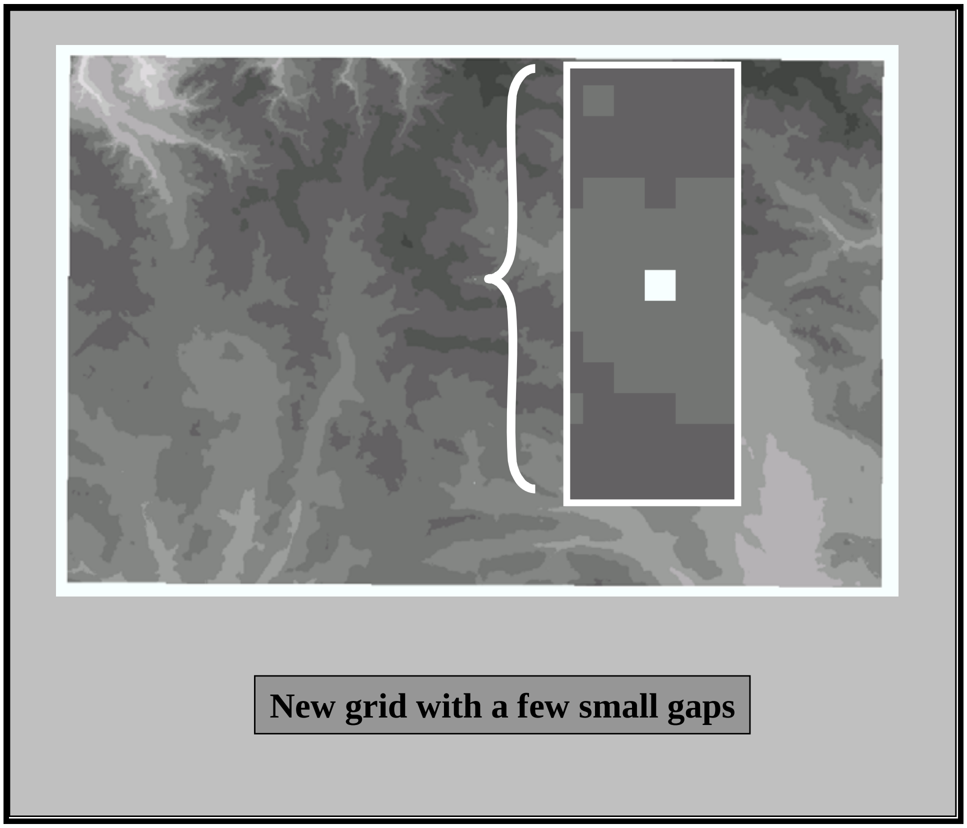
Next, open the Raster Calculator by clicking the arrow in the Spatial Analyst tool bar and then selecting the Raster Calculator. In the example presented here I will simply mosaic two elevation layers. To stitch together more layers you would simple follow the same format, adding the additional layers to the map algebra statement. There is limit of forty-nine grids that you can mosaic in one statement. The illustration depicts the two input layers.



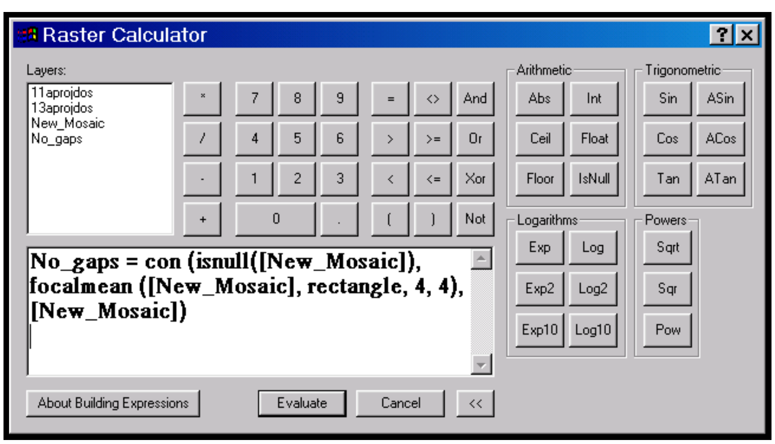
In the expression box of the Raster Calculator type the name you want to assign to the new layer. Here I have called the layer: New_Mosaic. Next, type the equals sign and the word Mosaic followed by a left parentheses. In the Layers window you can see the names of the grids that are currently in the Table of Contents. To insert a grid into the expression, double click its



name in the layers window. Follow the first grid name by a comma and then double click to insert the next grid. Continue in this manner until you have added all of the grids that you wish to mosaic. While entering information into the expression box, do no press the Enter key to move to the next line. Simply allow the software to wrap the lines. After the last grid, insert the right parentheses and then click the Evaluate button. The program will enter the new grid [New_Mosaic] into the Table of Contents and display it.



In some cases there might be small gaps between some of the input grids. Cells in these gaps have the value, No Data. In the case of elevation data and other continuous spatial series you can interpolate values for the blank data cells from surrounding cells for which you know the valid values. This is pretty safe, assuming the gaps are not too large. In the layer, New_mosaic, there are a few small gaps. In the illustration I have enlarged a small portion of the mosaiced layer to show one of the gap cells.

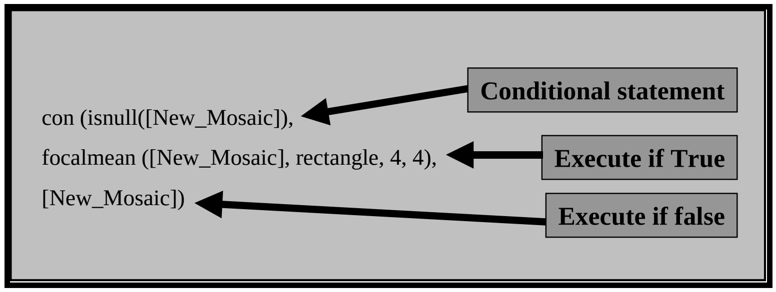


The map algebra statement that enables you to "fill in the gaps" is the conditional function in conjunction with the focal mean function. The conditional statement has the following syntax:

con(test condition, True, False)

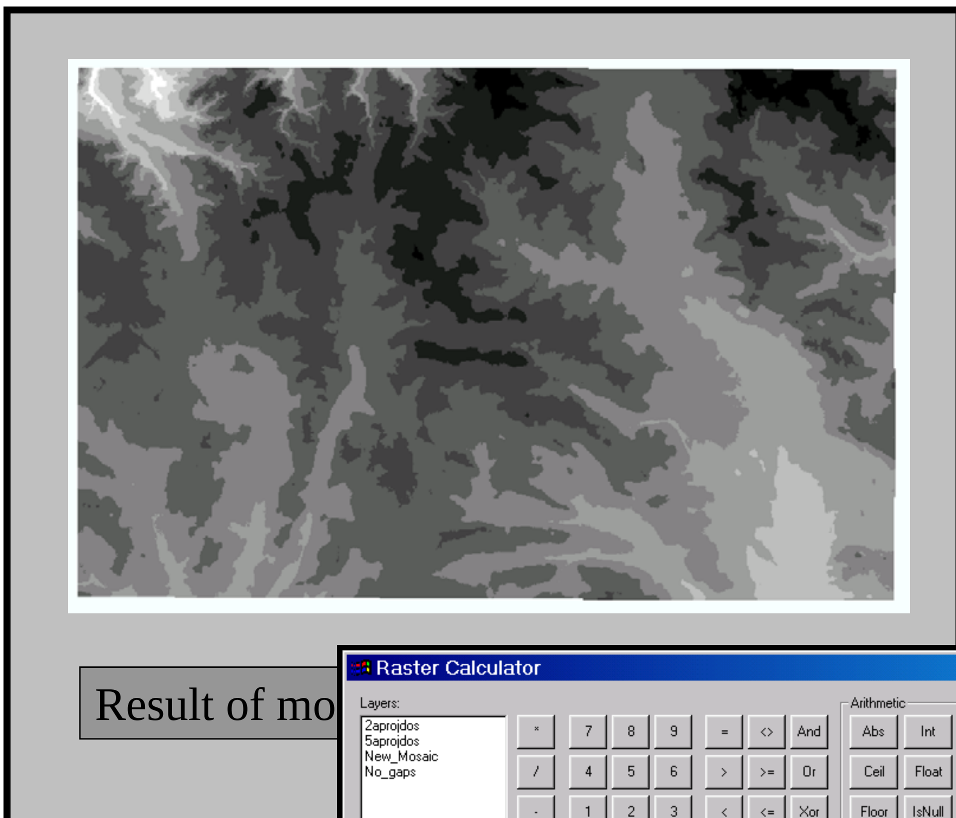
The test condition is a conditional statement that the program will evaluate as either true or false. If the condition is true, then the program executes the True portion of the statement. If the conditional statement is false, then the program executes the False portion of the statement. In the illustration, I have set the new grid, No_Gaps, equal to the conditional statement. In the conditional statement the condition is: isnull([New_Mosaic]. This statement checks the data value of every cell in the grid, [New_Mosaic]. If the cell value is equal to No Data, then the condition is true and the program executes the True portion of the statement:

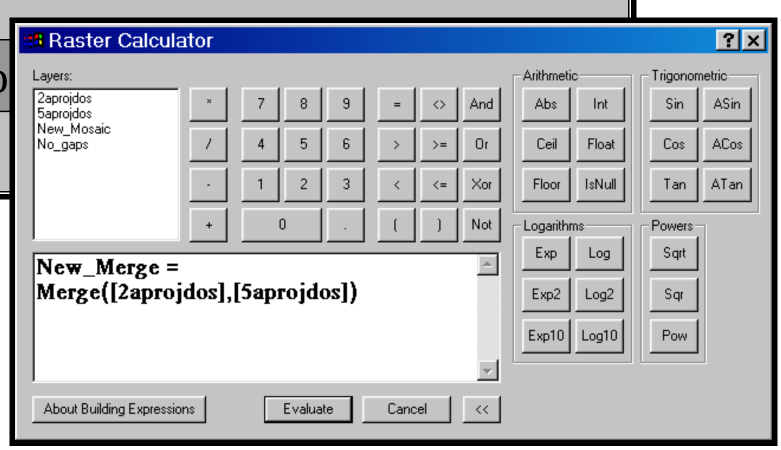
focalmean (New_Mosaic], rectangle, 4,4)



This function takes the average of the cells in a 4 by 4 neighborhood and assigns the result to the No Data cell. If the cell value the program is evaluating in equal to anything other then No Data, then the condition is false. In this case the function assigns the existing data value, that is the value in the [New_Mosaic] grid, to the cell.

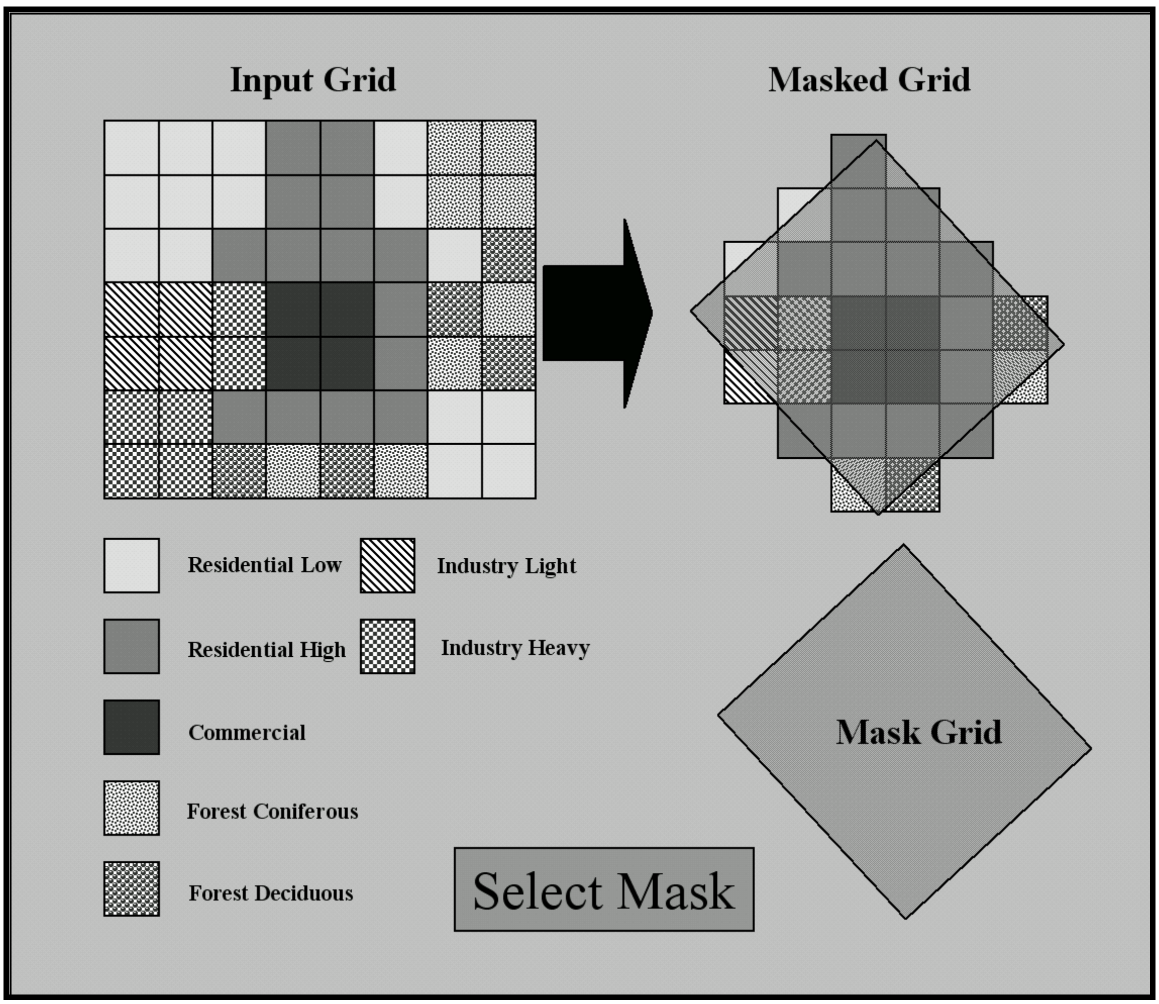
The illustration summarizes the use of the con statement along with the isnull and focalmean statements to provide interpolated data values to No Data cells in a grid layer. Remember that you would not want to do this in the case of categorical data sets, because taking the mean of such values is, shall I say, meaningless. Sorry I just couldn't help it. The resulting layer from the mosaic and the focal mean operation is displayed here.



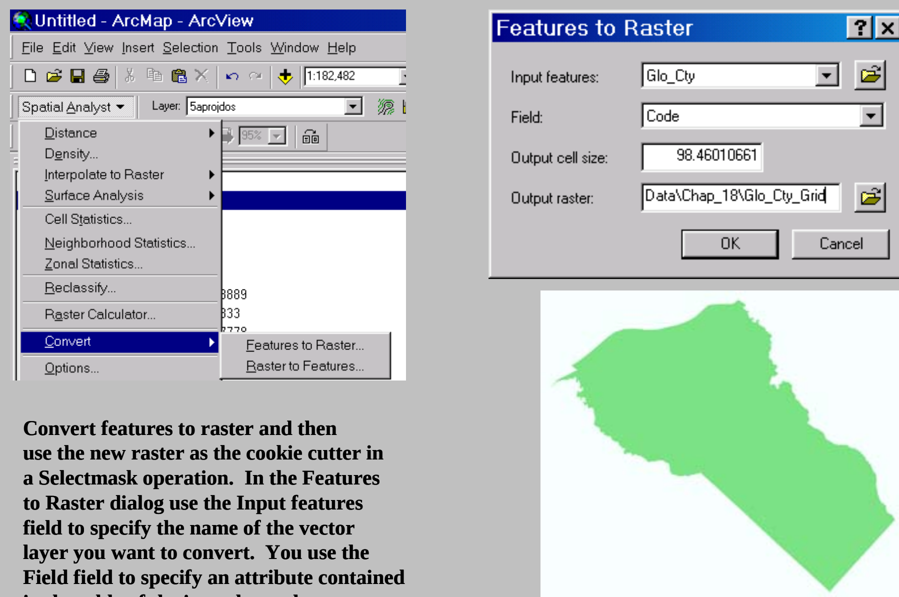


Merge Operation. When the input layers are adjacent, but not overlapping, the main difference between mosaic and merge is syntax. The result produced by the statement in the illustration will be exactly the same as if you substituted the word mosaic for merge. Remember, however, that the two operations deal with overlapping layers in very different ways: mosaic calculates the average of input cells; whereas merge assigns the first valid grid cell value to the area of overlap. In the example in the dialog box, cells in 2aprojdos have precedence over cells in 5aprojdos. This is the appropriate operation in the case of categorical data values such as land use.

Mask. The raster mode equivalent of clip is mask. In this case you will use one grid layer as a cookie cutter and another as the cookie dough. In the illustration there are two input grids. The first, which serves as the cookie dough layer, is simply identified as Input Grid; whereas the second, the cookie cutter, is the Mask Grid. For the Mask Grid assume that all cell values within the darker gray border have valid data values and all cells [not shown] outside the border are equal to No Data. The Select Mask function takes the Input Grid and overlays the Mask Grid on top of it. The program selects all cells in the Input Grid that have their centers within the Mask Grid's set of valid data values. In the output layer [i.e., the cookie], the program assigns these selected cells the same data value as the corresponding cell of the input [cookie dough] layer. In the output or masked grid, the program assigns all input grid cells that have centers outside the mask the value No Data. In the illustration only the cells that you can see through the mask layer in the upper right portion of the diagram will have valid data values in the output layer.



SelectMask. After you have mosaiced or merged a set of grids, you might wish to cookie cut or mask the resulting large grid layer so that map readers can see the grid in the familiar shape of a county, state or other mapping area. The SelectMask operation requires two input layers. One of these layers is the grid that you are going to cookie cut. The other is the cookie cutter or mask layer. Both of these input layers must be in grid format. In some instances the area you wish to use as the cookie cutter will be a vector layer. This, of course, is not a problem as you can very easily convert a vector polygon to a raster layer and then use the converted layer as the cookie cutter. For instance, in the exercise you completed in the previous chapter, you created a shapefile that depicts the outline of Gloucester County, New Jersey. You can convert this file to a grid very easily. First add the shapefile to the Table of Contents. Next, click the arrow in the Spatial Analyst toolbar and point to Convert in the drop down menu. The program slides out a menu from which you select, Features to Raster. The program opens the dialog.



Convert features to raster and then use the new raster as the cookie cutter in a Selectmask operation. In the Features to Raster dialog use the Input features field to specify the name of the vector layer you want to convert. You use the Field field to specify an attribute contained in the table of the input layer that you want to assign to cell values in the new raster layer. This is irrelevant here. Cell size is the size you specified in the Cell Size tab of the Spatial Analyst Options dialog. Use the Output raster field to provide a location and name for the new raster layer.

In the Features to Raster dialog use the Input features field to specify the name of the vector layer you wish to convert. In the Field window click the down arrow and select an attribute field containing data values you wish to assign to cells in the new raster. In this case which field you select is of no importance, as all the masking operation is going to do is look for valid data and No Data value cells. The program sets the Output cell size field to whatever value you specified for new grids in the Options dialog. Finally, specify a location and name for the new raster and then click the OK button. The program will create the new grid and store it in the location you specified. See the illustration above.

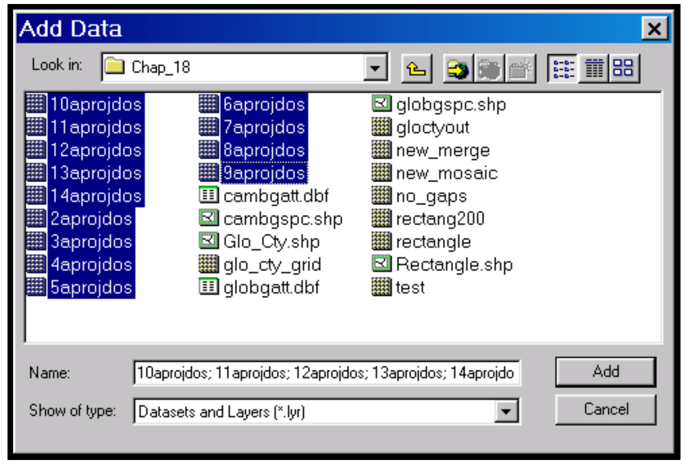
All cells in the cookie cutter layer that have valid data values are "inside" the cookie cutter. Corresponding cells in the cookie dough layer will retain their current data values after being cut into cookies by the cookie cutter. All cells in the cookie cutter layer that are "outside" the cutter have the value, No Data. Corresponding cells in the cookie layer will not retain their cookie dough cell values, but rather will be replaced with the No Data value.

Now that you understand the idea behind masking, the implementation is easy. All you need to do is use the Raster Calculator to execute the Selectmask operation, which has the following syntax:

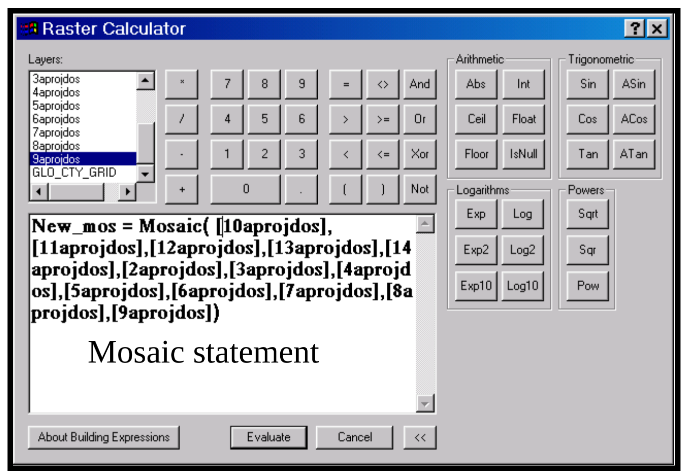
New_Grid = Selectmask([inputgrid], [maskgrid])

Where: The [input grid] is the grid layer you want to cut out into the shape of the valid data values in the [maskgrid].

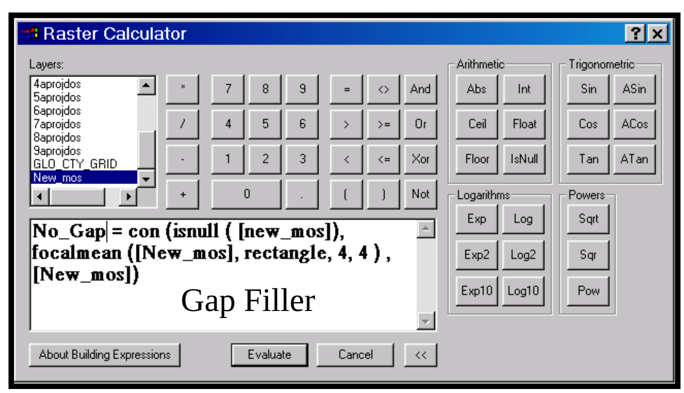
The first step in creating a mosaic of grids it to load the layers you wish to stitch together. In this case I have decided to add all of the elevation grids needed to create a digital elevation model that extends beyond the area covered by Gloucester County, New Jersey.



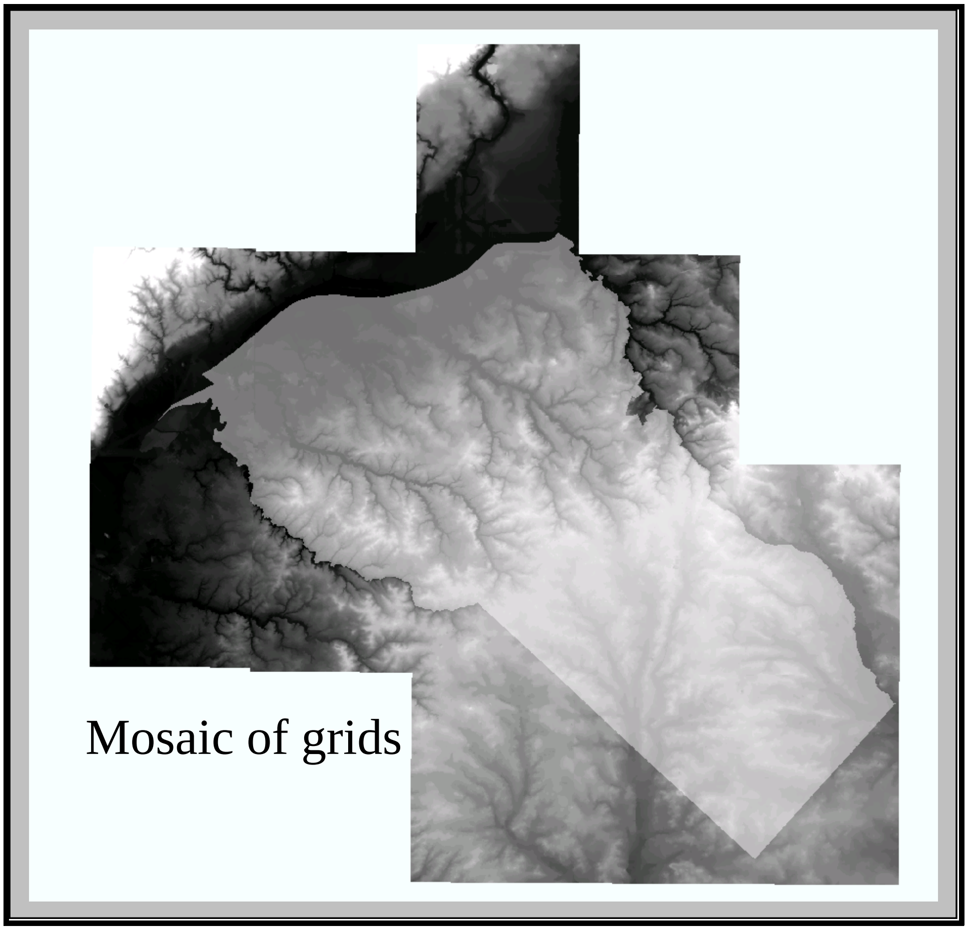
Create the mosaic. The next operation is to create the mosaic by using the Mosaic operation in the Raster calculator of the Spatial Analyst. In the statement I am creating a new grid called, New_mos. Notice that the listing of grids that will be stitched together begins right after the word, Mosaic and is contained with parentheses. Each grid name must be in brackets and each grid name is separated from the next with a comma. As you know the easiest way to ensure the correct spelling and format of grid names is to enter each one into the expression by double clicking its entry in the Layers window. After you have entered all of the grids and inserted the final parentheses, execute the statement by clicking the evaluate button.



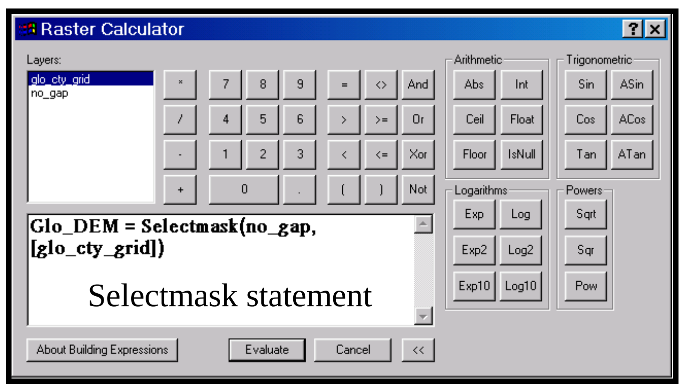
The new grid, new_mos, probably contains a few gaps in which the cell value is No Data. Use the conditional statement in the illustration to get rid of the gaps and fill them with the focalmean calculated from valid values in nearby cells. The resulting grid, No_Gap will be free of unwanted No Data cell values. The statement checks the value of each cell in the grid, [new_mos] to see if the cell is equal to No Data. If it is, then the statement calculates the average value of the cells in a 4 by 4 window surrounding the cell and assigns the result to the No Data cell. If the cell the statement is checking has a valid data value, then the value remains unchanged.



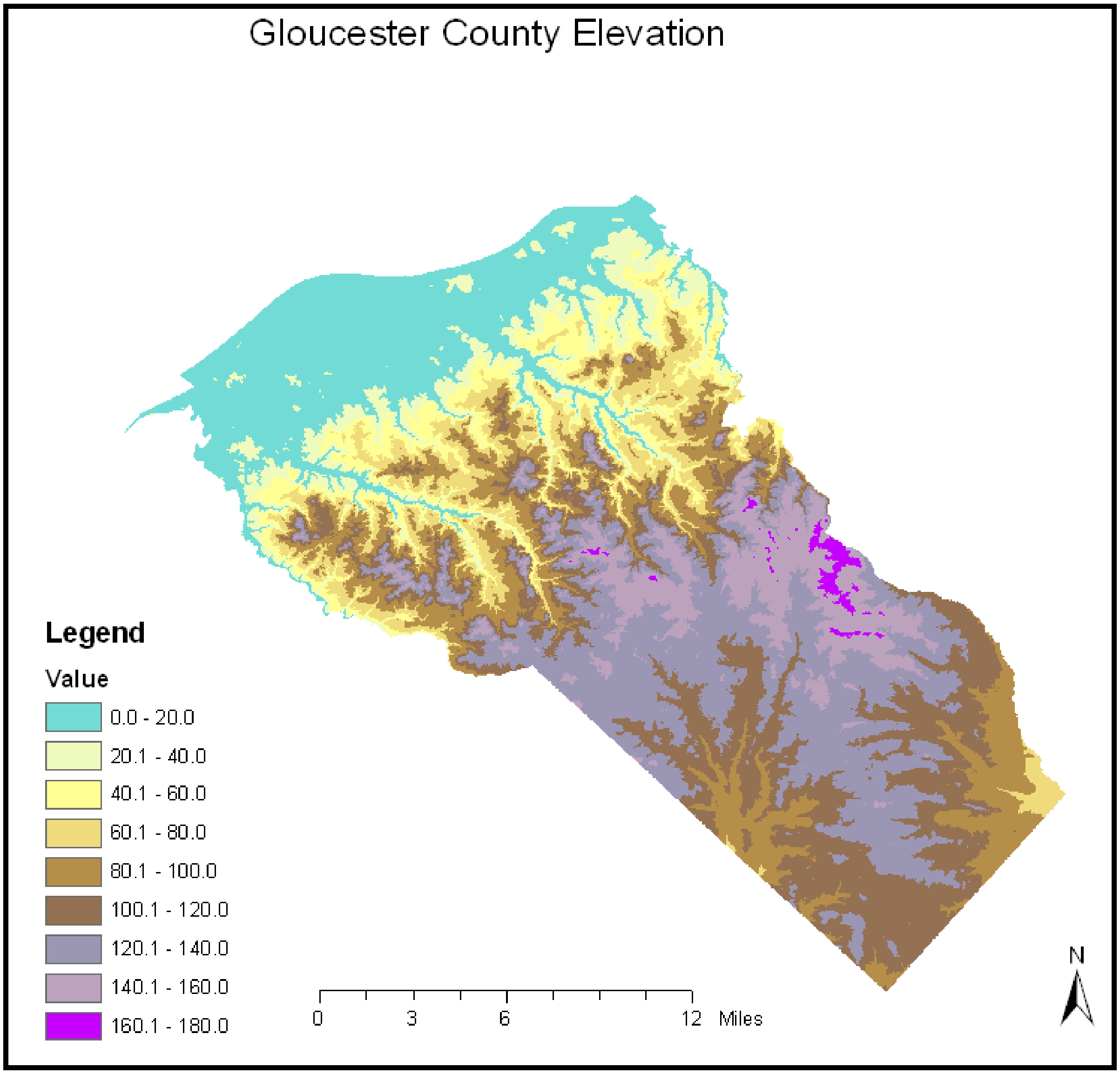
The illustration depicts the results of the original mosaic and subsequent gap filling of the grid. In the illustration I depict the elevation grid with an overly of the Gloucester County grid. So that you can see the elevations below the county outline, I have used a light colored, transparent color for the grid cells in the county layer. The only remaining task is to do the cookie cutting so that the rather blocky looking mosaic is cut into the shape of the county.



You access the SelectMask operation through the Raster Calculator. In the illustration I am creating a new grid layer called Glo_DEM. The input grid layer is no_gap, which I created with mosaic and the gap filling operation. The mask or cookie cutter is called glo_cty_out, which is a grid layer depicting the form of Gloucester County, New Jersey. I created this grid by converting a shapefile to grid. The result of the expression in the Raster Calculator is a digital elevation database of Gloucester County, New Jersey.

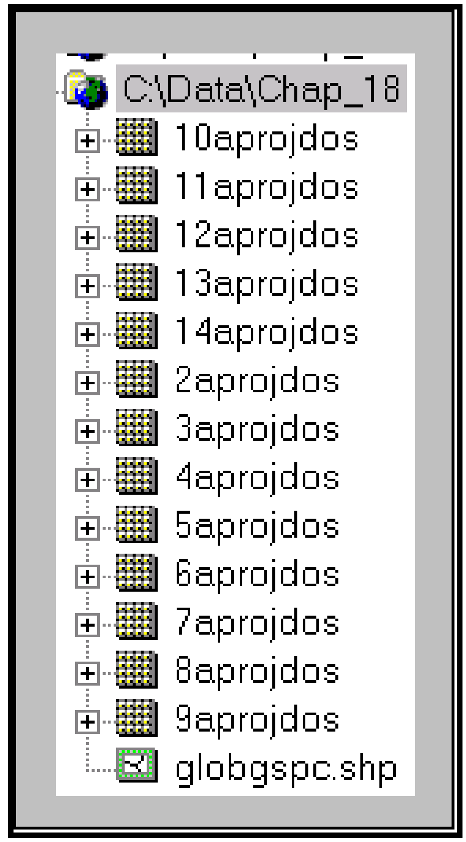


The illustration depicts the final result of the mosaic, gap filling, and cookie cutting operations.



Workshop: In this workshop you will create a digital elevation database for Gloucester County, New Jersey.

Preliminary operations. The very first thing you need to do is copy the workshop database from the open area to your network drive. The folder containing the database is called Chap_18. The contents of the database consist of a set of grid layers that together encompass the area in the vicinity of Gloucester County along with a shapefile that depicts the Gloucester block groups. As is always the case, you will find that processing goes more quickly if, when you start working on the workshop you copy the entire database from your network drive to the local hard drive. Just remember to copy it back when you are finished!



Ths workshop requires that you complete the following operations:

1. Use the Arc Toolbox dissolve operation to perform a dissolve on the shapefile, globgspc.shp. The result will be a shapefile that contains the outline of the county. Call this file Glo_Cty.
2. Load the Spatial Analyst and open its toolbar
3. Load the grid elevation data layers. 2aprojdos, 3aprojdos, etc.
4. Set the Spatial Analyst options for default directory, extent, and cell size. Remember to set the extent to "union of inputs."
5. Convert the county outline layer to a grid layer.
6. Mosaic the grids. Use the mosaic operation in the Raster Calculator to stitch all of the grids together, resulting in a single large grid layer.
7. Use the conditional operation along with the focalmean function to fill in any gaps in the stitched together layer.
8. Use the masking operation to cookie cut the grids into the shape of Gloucester County.

File	Globgspc.shp
Task 1	Add a new field to the attribute table and populate the field with 0. The population part should be automatic.
Task 2	Use the dissolve operation to dissolve all interior boundaries. The dissolve field will be the one you just created. You will find the dissolve operation in Arc Toolbox in the Data Management toolbox and Generalization tool shelf.
Product	Shapefile of outline of Gloucester County. Call it Glo_Cty. You will use this later

File	Glo_Cty.shp and grid layers
Task 1	Load the Spatial Analyst and then add all of the grid files. These files have names that begin with a number and then contain the following letters: aprojdos. For example, one of the files is 5aprojdos. The gird files contain elevation data in feet above sea level.
Task 2	Set the Spatial Analyst options. For the working directory specify the location where you have your data, Chap_18 for instance. For the cell size specify same as one of the grid layers you loaded. For the extent specify union of inputs.
Task 3	Use the Spatial Analyst's Convert function to convert the shapefile, Glo_Cty to grid. Call the output Glo_Cty_Grid
Product	Cookie cutter file, Glo_Cty_Grid. Please map it for me.

File	All of the elevation grid files
Task 1	Use the Mosaic operation to combine all of the grids into a single layer. Call this layer Glo_merge.
Task 2	Use the con and the focalmean operations as explained in the text to clean up any gaps in the merged grid, Glo_merge. Call the result Glo_nogap.
Product	Elevation layer cleaned of gaps. File name Glo_nogap. Please print Glo_nogap.

File	Glo_nogap and Glo_Cty_Grid
Task 1	Use the Selectmask operation with Glo_nogap as the input [cookie dough] layer and Glo_Cty_Grid, the cookie cutter file to create an elevation layer for Gloucester County. Call the layer Elevation
Task 2	Design a cartographically correct layout of county elevations.
Product	Print a layout of the elevation layer you created.

Assignment: please hand in the maps requested in the step by step workshop outline.