Job aides to assist in the development and analyses of raster data for Soil Mapping

Contributors in alphabetical order: Jon Bathgate, Jon Bonjean, Joe Brennan, Dwain Daniels, Tom D'Avello, Adolfo Diaz, Chad Ferguson, Suzann Kienast-Brown, Bob Long, Lynn Loomis, Jessica Philippe , Xun Shi, Fred Young

Tom D'Avello (Editor), National Soil Survey Center, Geospatial Research Unit, 2016

Best Practices for Processing Raster Data in Soil Survey Applications

By NRCS Soil Scientists and GIS Specialists Tom D'Avello, Dwain Daniels, Adolfo Diaz, and Suzann Kienast-Brown.

Background

There is a wealth of raster data currently available for the development of covariates that can be used with countless soil survey-related activities. Because of the large spatial extents of many soil survey projects, planning is of the utmost importance in data development. No one wants to have to recreate a covariate that took 5 days of CPU time to complete because a problem was not discovered until the data was ready for analysis. Following a few guidelines during the preliminary stages of data processing can help eliminate or minimize problems during the application stages.

Processing Checklist

A quick workflow for processing raster data includes:

1) Verifying the data source
2) Verifying the projection parameters
3) Verifying the horizontal **and** vertical units
4) Verifying the resolution
5) Verifying the extent in terms of rows and columns
6) Verify raster statistics
7) Setting a snap raster
8) Buffering the project area
9) Using a watershed for hydrologically based derivatives
10) Processing data
11) Storing output in a common folder

Data source

Do you have complete coverage of raster data from the same source for the extent of your project area? A project area commonly will have raster data coverage that is delivered with the appearance of being a uniform, consistent product but actually has more than one source, and different sources utilize different data capture sensors/techniques, quality parameters, processing techniques, etc. Figure 1 shows an area where elevation data can be extracted as a single 3m resolution raster. The join line between two different sources of LiDAR data extends from the upper left to the lower right corner, and the difference in bare earth surface characterization and significant artifact occurrence is very clear. Just as users of "seamless" SSURGO data with a project area that extends across original soil survey area boundaries need to be aware of potential differences in the soil properties/interpretations and their impact on analyses, users of raster datasets for developing covariates need to be aware of differences and consider them in planning.

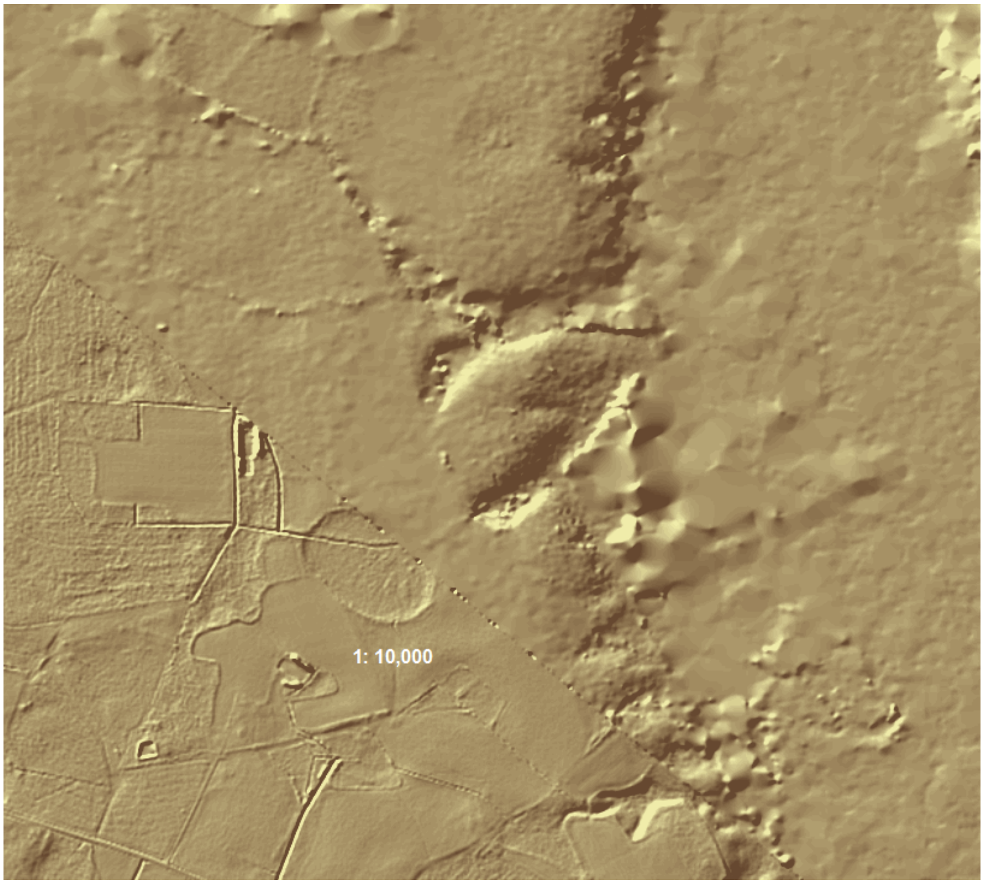


Figure 1. —Hillshade derivative example of a 3m resolution DEM comprised of LiDAR data from two different sources.

Projection parameters

ArcGIS provides dynamic projection capabilities; however, the best practice is to use one common projection for all of the raster data used in the GIS analyses. For DEMs that require re-projection, refer to the Job Aid that provides details related to that operation. If the raster data will also be utilized in applications such as R or ArcSIE, all layers must share common projections, resolutions, and extents. Save data in a GDAL compatible file format for use with R (see list at: http://www.gdal.org/formats_list.html) and in a ESRI GRID format for use with ArcSIE.

Units

Matching the horizontal and vertical units results in assumption-free terrain derivatives. Data provided in a geographic coordinate system, such as decimal degrees, must be converted to a projected coordinate system. Many users have been frustrated trying to interpret a slope-gradient layer generated from mismatched horizontal and vertical units. There are cases, primarily in engineering, in which users maintain DEMs with vertical units that differ from horizontal units, but there is no great need for this in soil survey applications. The primary reason for matching the horizontal and vertical units is to reduce file size. For example, a file with vertical units in meters may be stored as a floating-point, 64 or 32 bit file with a file size of 400 MB. Converting the vertical units to integer feet or centimeters would result in a file size of 200 MB. It is important to note that adoption of this space-saving option will limit the choices in developing slope gradient and curvature layers in GIS packages that accommodate the z scaling parameter (ArcGIS, QGIS). In these packages, you must remember to use the proper z scaling factor. In addition, the 3 x 3 neighborhood utilized by ArcGIS imposes additional limitations (see the May 2016 NCSS Newsletter, page 16). Finally, this space-saving option will help avoid problems when developing compound terrain derivatives, such as wetness index or stream power index, that assume common horizontal and vertical units.

Resolution

Given the wide range of available data and sources, any given set of assembled data will have an array of resolutions. Developing an inventory that lists the data layers and their resolutions helps in keeping track of everything and documenting the development of metadata. A common cell resolution should be used for all analyses. There are no absolute rules for determining a working resolution. The best practice is to:

1) Determine the phenomena(on) of interest for the project
2) Determine the largest resolution that is required to appropriately map the phenomena(on)
3) Select the resolution from your data inventory that satisfies check #2
4) Resample all data to be used in analyses to the resolution selected in check #3

Two common methods available in ArcGIS for resampling are Resample and Aggregate. Resample is more versatile, allowing for increasing or decreasing to any resolution, while Aggregate only provides for decreasing resolution by whole factors. A Job Aid is available that discusses the use of the Resample tool. It is important that you select the bilinear or cubic convolution resampling technique when resampling continuous data types with the Resample tool. The only time you should select the nearest neighbor technique is when resampling categorical data types.

Common reference

Your objective is to develop all of your raster data with common extents, common number of columns and rows, and common cell alignment. There should be a 1:1 relationship between all cells among the data layers to be used in analyses. Figure 2 shows how multiple resolutions and misalignment cause problems.

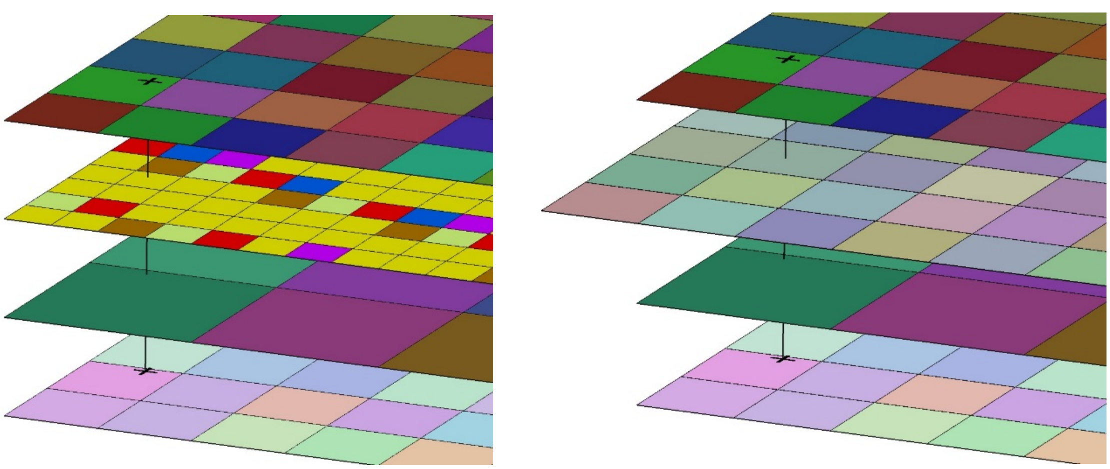


Figure 2.—These images illustrate how the value in one cell of a layer does not apply to the corresponding cells in other layers with different resolutions.

The requirement is that one cell matches all other cells, as shown in Figure 3.

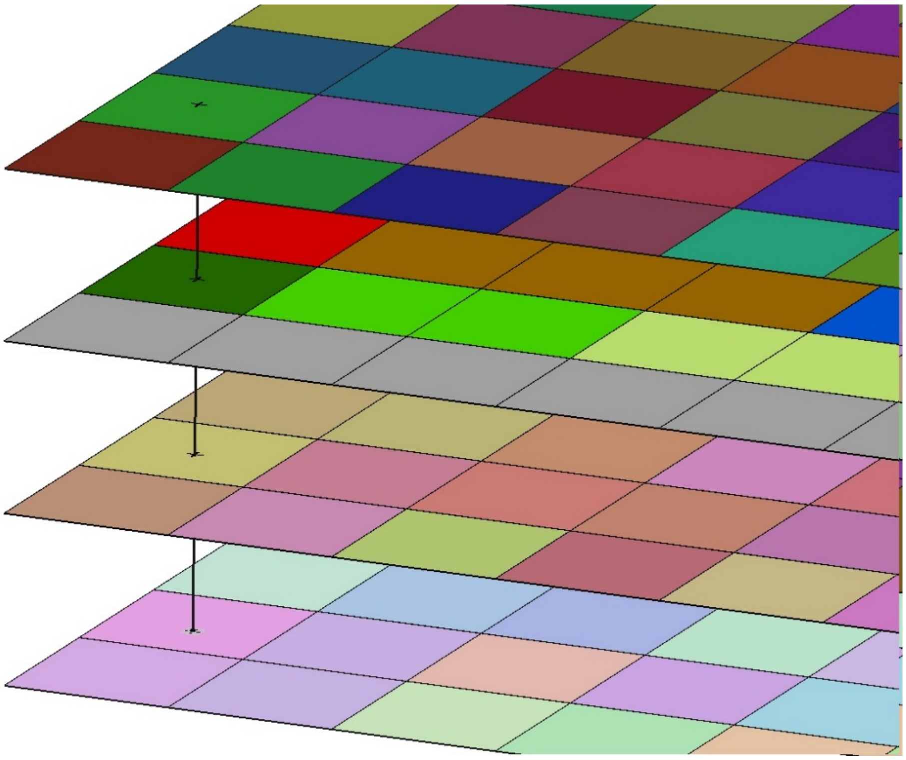
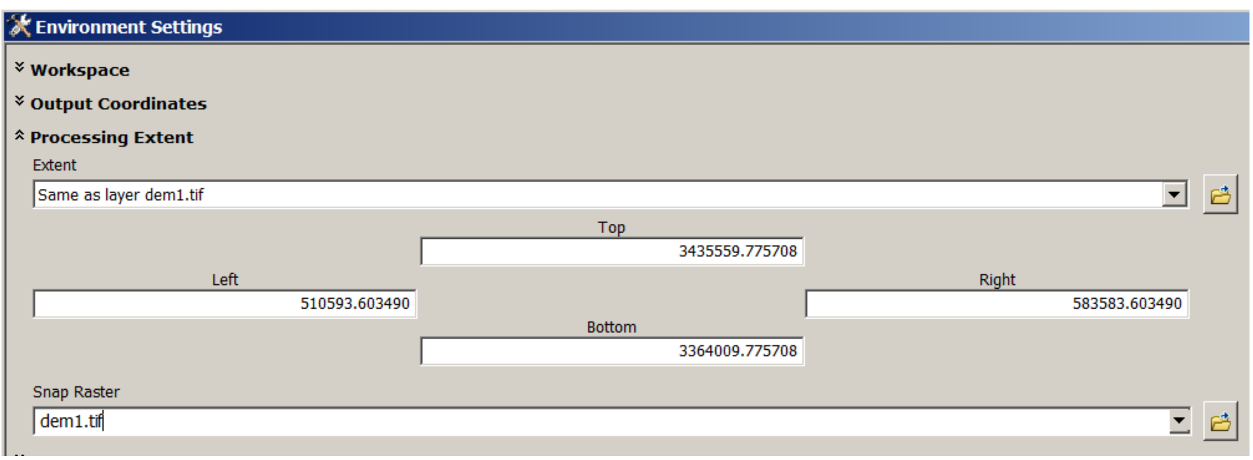
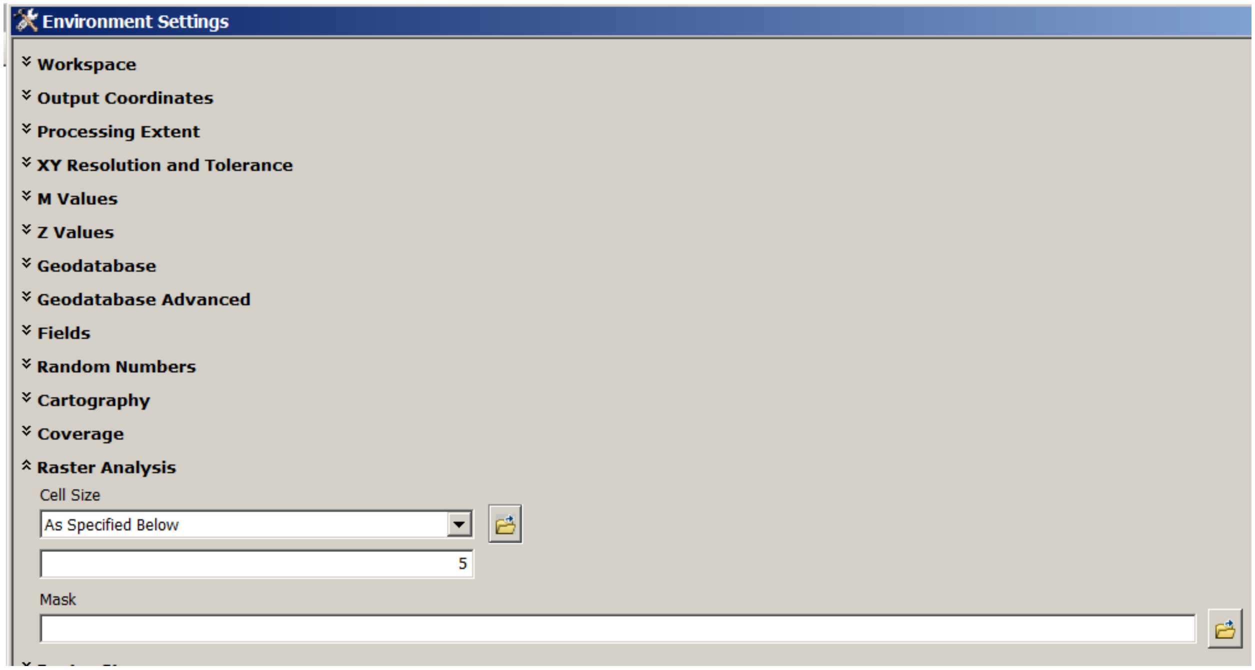


Figure 3. —This image illustrates raster layers that all share a common extent and resolution.

Use the Geoprocessing Environment in ArcGIS to specify the processing extent and snap raster in the Processing Extent drop-down menu. As a general rule, the snap raster should be set to the origin layer from which derivatives are being developed to ensure cell alignment.

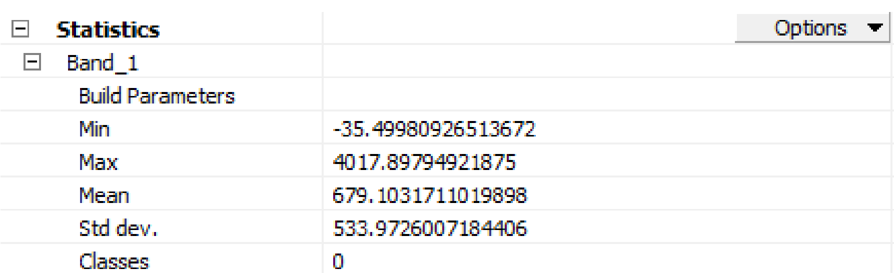


Specify the cell size in the Raster Analysis drop-down menu:



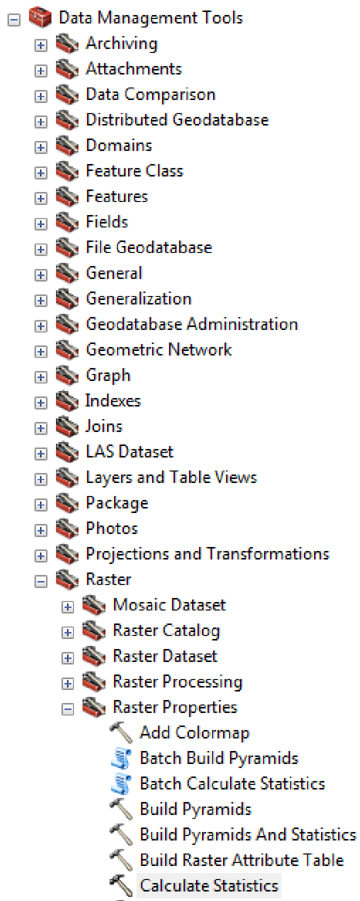
Raster statistics

Calculating raster statistics allows ArcMap to properly stretch and symbolize raster data for display and can also serve as a quick quality assurance review by computing the value range of your data. The statistics that are calculated include the minimum and maximum pixel values, the mean and standard deviation of the calculated pixel values, and, if the dataset is thematic (such as a land cover dataset), the number of classes. If your dataset is continuous, such as a DEM, there will be no classes. Below is an example showing the statistics for a DEM whose minimum value may be questionable.

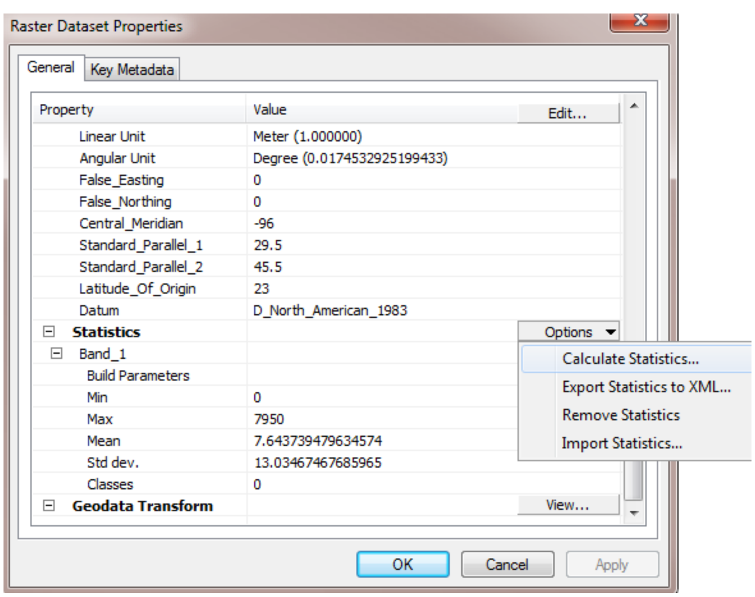


Statistics	Options
Band_1	
Build Parameters	
Min	-35.49980926513672
Max	4017.89794921875
Mean	679.1031711019898
Std dev.	533.9726007184406
Classes	0

There are multiple ways to calculate statistics for a raster dataset. The Calculate Statistics tool in the ArcGIS Data Management toolbox is the most direct way.



Another convenient way is through the Raster Dataset Properties interface, which is accessed by right-clicking on the raster dataset within ArcCatalog. Although you can access the raster properties in ArcMap, you cannot calculate them.



Buffering the project area

Raster processes use a neighborhood to determine the value of the middle cell. The edge of a raster layer will have many No Data (null) values, producing an “edge effect” along the border (see Figure 4).

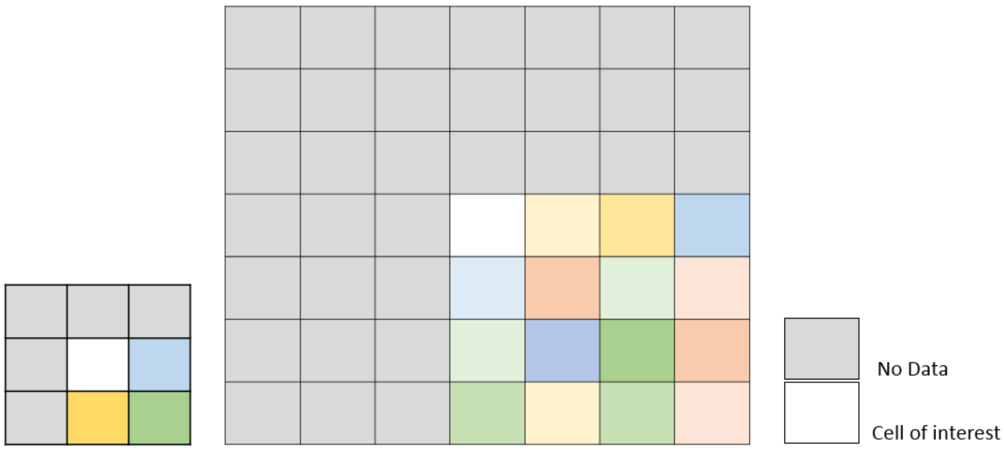
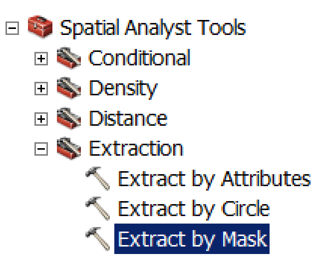


Figure 4. —The typical 3 x 3 neighborhood is on the left, and the larger 7 x 7 neighborhood is on the right.

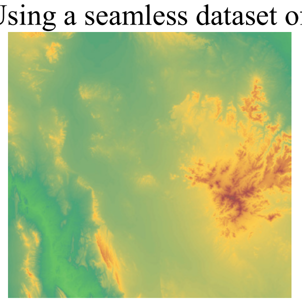
In Figure 4, the value for the center cell will be determined from 3 cells in the 3 x 3 neighborhood, rather than the typical 8, and from 15 cells in the 7 x 7 neighborhood, rather than 48. It is important to note that algorithms based on focal statistics have an optional parameter to “Ignore No Data in calculations.” If that option is chosen, No Data input cells will remain as No Data in the output.

Given the variability that can occur at the edge of a raster dataset, you should make sure that your input raster layers extend well beyond the project area. There are no hard rules, but a buffer of 1,000 to 2,000 meters is reasonable. This buffer generally is created from a polygon file defining the project area, but a raster file may also be used. The Extract by MASK tool in the ArcGIS Spatial Analyst toolbox is the most direct way to subset data from a larger extent:



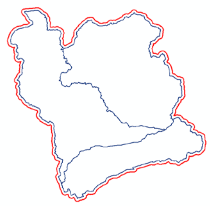
A typical workflow includes:

1) Using a seamless dataset of desired projection, resolution, and units

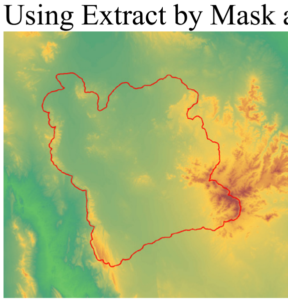


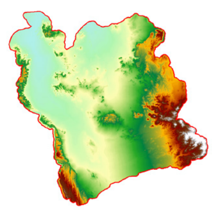
2) Ensuring the dataset defining the project area matches the projection parameters of layer #1

3) Selecting buffer layer #2



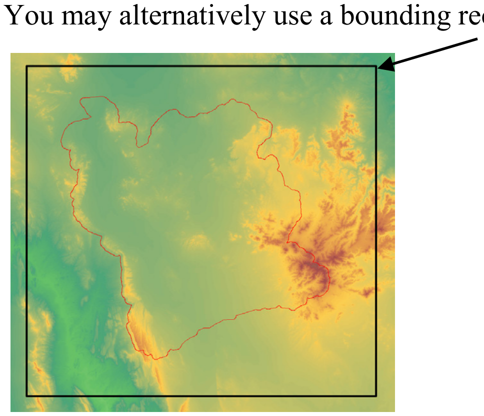
4) Using Extract by Mask and the Buffer layer to subset data





5) Setting the geoprocessing environment, extent, snap raster, and resolution to output layer from step #4
6) Using the buffer layer as a “clip” file for extractions on all other layers required for the project
7) Developing covariates from layers

You may alternatively use a bounding rectangle as a “clip” layer to extract subsets.



If you are processing smaller areas to facilitate computing but your eventual goal is one layer of large extent, you can “clip” the larger buffered layers using a buffer that slightly exceeds the original project area boundary. Adjacent layers can then be merged using Mosaic or Mosaic to New Raster to reform intact project areas with minimal match problems.

Using complete watersheds for hydrologically based covariates

Hydrologically based derivatives that have proven useful for soil survey applications include:

Flow accumulation
Slope length
Stream link
Stream order
Stream power index
Upslope contributing area
Watershed
Wetness Index

Brief descriptions of these covariates are available on the NRCS's Soil Geography website. The cell values for hydrologically based derivatives need to correspond among watersheds for analysis purposes. Using the entire extent of a watershed is the only sure way to satisfy this requirement. The 10- or 12-digit HUC boundaries serve as the most convenient reference. Follow the steps in the previous section and buffer the HUC 1,000 to 2,000 meters or use a bounding rectangle with a border well beyond the watershed edge, then "clip" input data to be processed (using the preceding process).

If the eventual goal is one layer of large extent, this procedure may be used to produce a mosaic dataset with minimal edge-matching problems.

For example, in Figure 5, the 1000 meter buffer file would be used to create a DEM subset. All hydrologically based covariates would be developed using this DEM. This process would be repeated for all adjacent watersheds using corresponding 1000 meter buffer files. The output covariate layers could then be subset using the 30 meter buffer file. The covariates from the 30 meter buffer would be used in the mosaic operation. The assumption is that adjacent layers will share relatively comparable data values within close proximity to their shared watershed divides. As a result, there is a smoother match when multiple layers are mosaicked into larger extents (see Figure 6).

Figure 5. —Example of buffered watersheds from HUC layer.

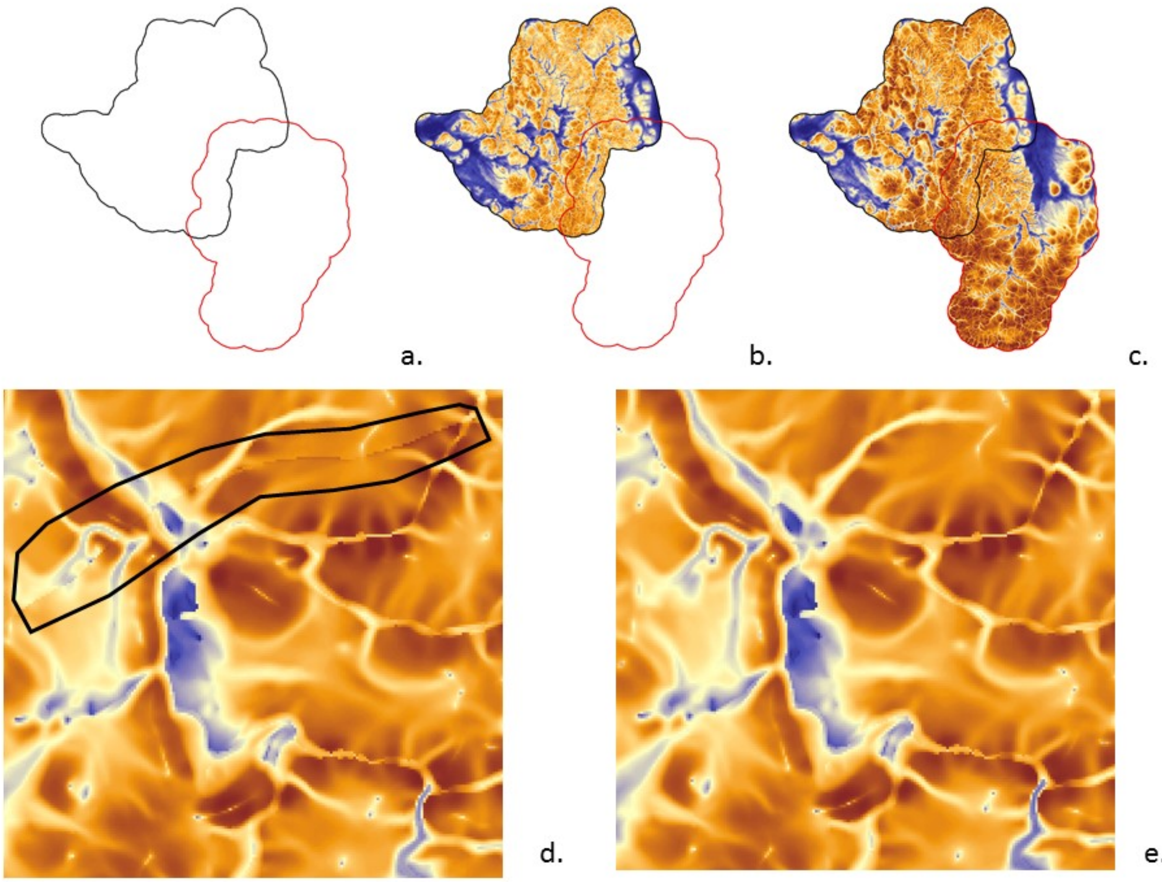


Figure 6.—Example of processing hydrologically based data. Image **a** shows two adjacent watersheds that have been buffered 1000 meters; image **b,** the wetness index layer for one of the watersheds; image **c,** the mosaicked wetness index; image **d**, the visibly poor match within the black polygon resulting from the mosaic of the 1000 meter buffered files; image **e,** the same area from a mosaic developed using the 30 meter buffered extracts of the original 1000 meter buffered wetness index layer with no visible match problem.

Storing output in a common folder

Place all covariates to be used for analyses in a common folder. This helps in data management, such as maintaining back-ups. More importantly, applications using raster data like R require all files to be stored in one common folder.

Develop a naming convention for files for use throughout the office. A Job Aid is available that provides useful guidance for data management. In general, GDAL compatible file formats, such as ERDAS Imagine or GeoTIFF, are preferred. If ArcSIE is being used for modeling purposes, use ESRI GRID format. It is also useful to store these files in a common folder.

Data and file management

Digital soil mapping projects generate many files. Developing a system to organize, name, back up, and document data for other soil scientists and end users of the final product is essential. The USDA Geodata Management Document pertains to data requirements for Service Centers, and is a reasonable system to adopt if you are already familiar with it. However, it is a dated system, centered on shapefiles and does not address the variety of products you will be developing and managing.

Develop a folder and file naming convention that works for your office. There will be interim products developed that will need to be named and managed, which will also require adoption of a naming convention. Developing a system to back data up is essential. External drives are inexpensive, and could serve as a backup device. It would be reasonable to have several drives of the same data for redundancy and worst-case planning.

Ideas for naming folders:

1. Quadrangle
2. Watershed (10 or 12 digit HUC)
3. Project_area
4. Climate
5. Biota
6. Elevation
7. Geology

The first three would be “project area” centric, while the last four are “theme” centric. The resulting file names would likely be “theme” or “project area” centric respectively. It is preferable to have unique file names for all data, independent of the folder they are stored in. For example:

Folder
Alpa_Quad
File
Alph_slp
Alph_asp
.
Folder
Beta_Quad
File
Beta_slp
Beta_asp
.

Would be preferable to:

Folder
Alpa_Quad
File
Slope
Aspect
Folder
Beta_Quad
File
Slope
Aspect

Why bother with unique names?

It is better than trying to sort out 30 files all named "slope", or "aspect" when you want to share or trouble-shoot data or processes.

Metadata

Creating and maintaining a minimum level of documentation for metadata is important. As time passes, personnel move on and the metadata should at least serve as a "recipe" of steps and parameters for the derivation of each data layer.

File Structure developed by SSO 12-5, St. Johnsbury, VT

Each quarter-quad has its own file structure, consisting of at least the following file folders:

DbfFiles, Final Map, Geodatabases, GRIDlayers, GRIDresults, Hillshades, Shapefiles

Each set of quarter-quads that is put together for a field season of work has a similar file structure, and this document details the steps to be taken to create all the files that will go into the Final Map, GRIDlayers, and GRIDresults folder. The following is a list of all the files and how they should be named; this example is from the averill_sw_nine area. If there are any files in these folders with any other names, they will be deleted.
More details on how these files are created come later in this document.

Folder	File Name	What is this file?
GRIDlayers	avsw	1m DEM of 9 qquad work area, clipped to 500m outside of 9qquad boundary, after each qquad has been shaved
	avsw99	avsw filtered using a 9 by 9 rectangular neighborhood
	avsw995	The 5m DEM that we use for inference (Filtered and resampled)
	Multiwetsm	Smoothed wetness index
	Slope30	Slope layer generated using 30m neighborhood
	Slope30re6	30m slope layer reclassed for ablation till
	Slope30re7	30m slope layer reclassed for outwash (includes A slopes)
	Slope60	Slope layer generated using 60m neighborhood

	Slope60re	60m slope layer reclassed to our 5 slope classes
GRIDresults	Cabot	Cabot inference results
	Colonel	Colonel inference results
	Dixfield	Dixfield inference results
	Harden	Hardened map of Cabot, Colonel, and Dixfield
	Hardenent	Uncertainty map
	Hardenexg	Uncertainty map
	Hardph	Hardened map with slope phase
	Hardphmu	Hardened slope phase map combined into map units
	Pdat	Poorly drained ablation till from model
	Allat	Ablation till inference results
Final Map	Allpm	Raster map created from certified parent material layer
GRIDS		
	Btmu	Basal till map units
	Btmu84	Btmu with slivers removed, 8- then 4-connected
	Final	Final raster map (all parent materials put back together)
	Rkymu	Rocky map units
	Rkymu84	Rockymu with slivers removed,8- then 4-connected
	Vrky	Very rocky map units
	Vrky84	Vrky with slivers removed, 8- then 4-connected
	Atmu	Ablation till mapunits
	Atmu84	Atmu with slivers removed, 8- then 4-connected
Final Map Shapefiles	Allpmpolys	Certified parent material exported to shapefile
	Averill_border	Outer border of mapping area
	Polys	Polys before labeling
	FinalPolys	Labeled polys
	SimpPoly	Labeled polys that have been simplified to remove vertices

NAMING CONVENTION FOR ELEVATION TERRAIN AND GENERALIZATION DERIVATIVES PROPOSED by Dwain Daniels, GIS Specialist CNTSC

The following naming convention is proposed to provide consistency in naming elevation derivatives and generalization products. The components of the file name are:

Area identification
type of surface (bare earth or first return if applicable)
cell size
pyramid used (if applicable)
derivative identification.

.Examples.

ms052_be_5m_5pyr_fel
This is the file name for a 5meter resolution filled elevation dataset created from a bare earth terrain dataset with a 0.5 meter vertical resolution pyramid applied for Leflore County, Mississippi.

When another dataset is created the extension would be added, as in the example the flow accumulation would have the extension _acc added. **ms052_be_5m_5pyr_fel_acc**

ms052_be_5m_5pyr_cir3_slp
This example would be the name for a slope gradient raster dataset that has had a focal mean calculation performed on the elevation raster in a 3 cell radius circle shape.

All raster elevation data is stored in the file geodatabase raster dataset format. This naming convention will **NOT** work with ESRI GRIDs that have a 13 character limit on dataset name. *If you are using ArcSIE for modeling efforts, follow the naming limits for ERSI GRIDs. ArcSIE will implement other raster formats in the future, but is currently based on the ESRI GRID.*
The extension used in the name of the primary and secondary elevation derivatives and selected generalization products that are most commonly used created are:

Table 1. Raster derivative type file extensions.

Extension	Derivative	Value Type	Description
_elev	Elevation	Floating Point 32Bit	Units above Mean Sea Level, default is meters, if units are not meters the unit is identified in the extension, e.g. elevft.
_hsd	Hillshade	Unsigned Integer 8Bit	Sunlight reflection off ground surface.
_slp	Slope Gradient	Floating Point 32Bit	Cells carry a value calculated from the 8 surrounding cells of the maximum rate of change in the ratio of vertical and horizontal distance.
_reggrp	Unique Regions	Unsigned Integer 16 Bit	Cells carry a sequential number of individual regions starting in the upper left corner of the matrix.
_rcls	Reclassification	Unsigned Integer 8Bit	Individual or groups of values in the original raster are converted to designated integer values in the output raster.
_nib#	Nibble	Unsigned Integer 8Bit	Regions of cells determined by a nibble mask are replaced by the cell values surrounding them. A number can be added to indicate what values were replaced.
_cir#	Focal Statistic	Floating Point 32Bit	Cell values are a statistic calculated from values of surrounding cells based on a geometric shape such as a circle, square, rectangle, etc. The number indicates the extent based on the number of cells in the geometric shape. Unless designated, the statistic is assumed to be the mean.
_fel	Filled Elevation	Floating Point 32Bit	Depressions in the surface are filled for continuous flow.
_fdr	Flow Direction	Unsigned Integer 8Bit	Each cell carries a value indicating one direction of flow. D8 flow direction.

_acc	Flow Accumulation	Signed Integer 32Bit	Each cell carries the value indicating how many cells exclusively flow into it.
_pfel#	Partial Filled Elevation	Floating Point 32Bit	Depressions in the surface are filled to a designated threshold depth. A number can be added to indicate the threshold depth value.

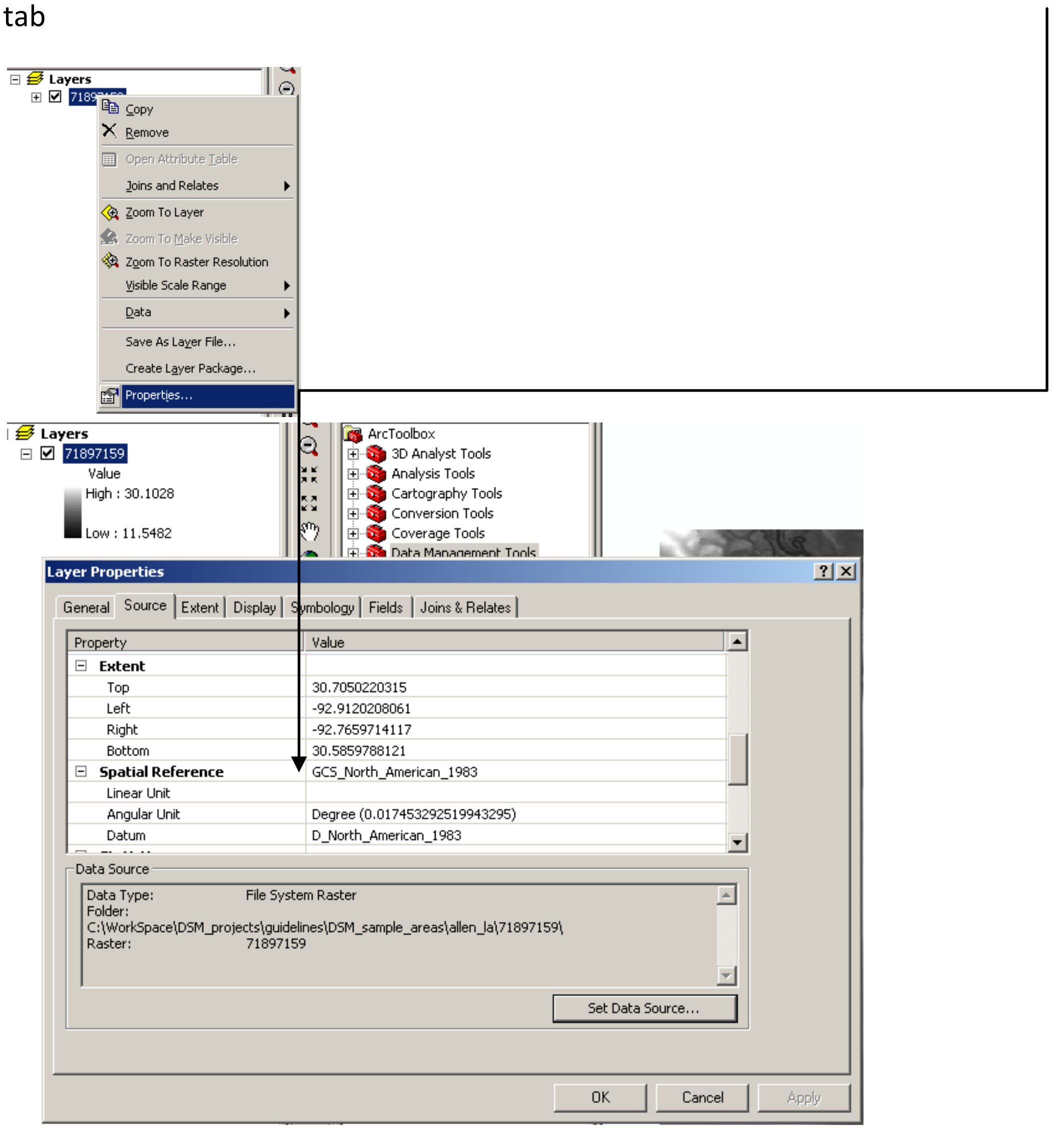
Projecting DEMs

Background

Elevation data from the USGS download sites and the elevation data contained in the CORE datasets by MO are provided in geographic coordinates. This data needs to be projected to a Cartesian system like UTM, State Plane or other accepted coordinate system for useful output of derivative layers like slope and curvature. Elevation data downloaded from the NRCS Data Gateway is in UTM. It is assumed that vertical units will be the same as the horizontal units before development of terrain derivatives.

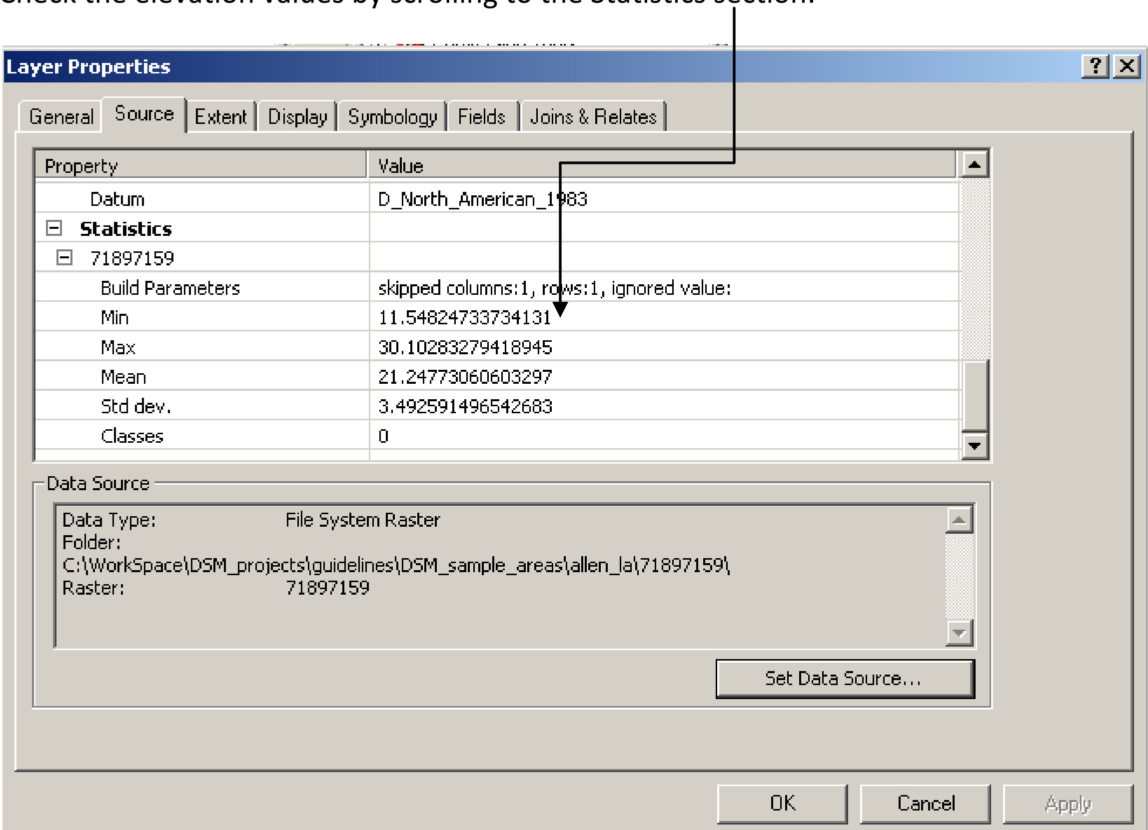
Check your data

Right-click the DEM data layer, select Properties and look for **Spatial Reference** on the **Source** tab



This data has a Spatial Reference of GCS_North_American_1983, and a Datum of North_American_1983.

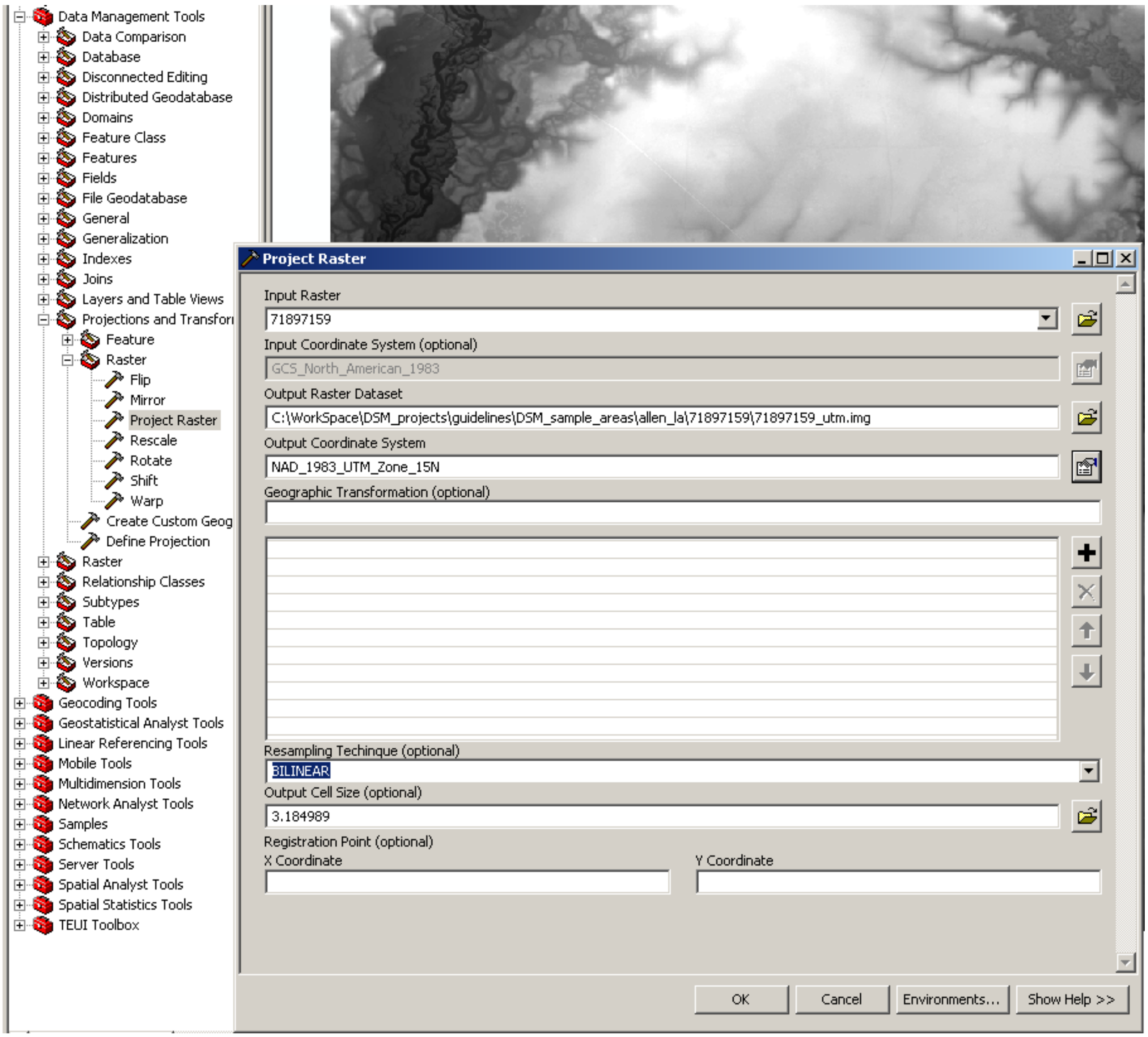
Check the elevation values by scrolling to the Statistics section:



The elevation units are in meters.

Projection example 1 – Geographic Coordinate system to UTM, elevations units are in meters

Open Data Management Tools
 Projections and Transformations
  Raster
   Project Raster
    Fill out dialog



Make sure you select **Bilinear** or **Cubic** for Resampling Technique. Nearest Neighbor or Majority should **never** be selected when using continuous data. There are mathematical differences between the Bilinear and Cubic methods, with Cubic being more complex. As a result, processing time can be longer for the Cubic method. In addition, the Cubic method has a greater potential to extrapolate elevation values compared to the Bilinear method. The Cubic method yields a smoother surface than Bilinear. The differences between the two methods are probably not detectable for soil survey purposes.

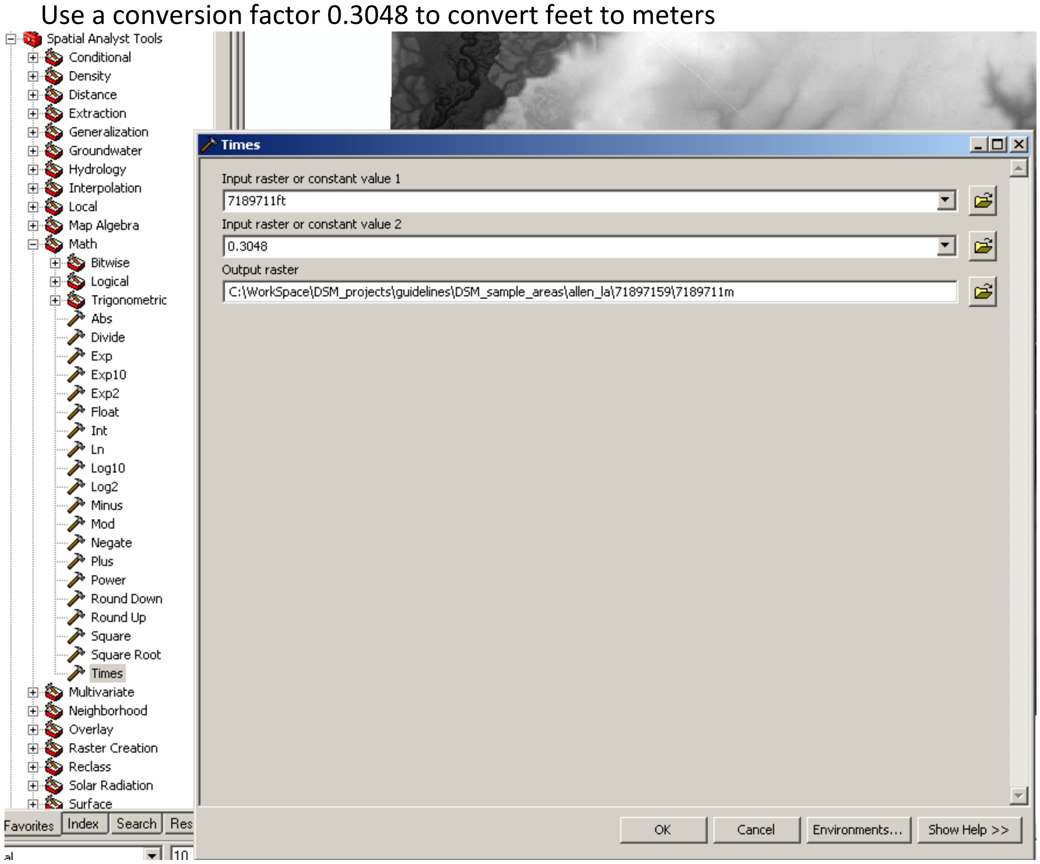
If you have determined an optimal cell size, set your output cell size at this point.

In this example, the ouput file was saved as an Imagine file. The most robust extension available for developing terrain derivatives, ArcSIE, requires raster data in GRID format. To save as a GRID, do not put a file extension in the name, and restrict the length of file name to less than 13 characters, with NO spaces.

Projection example 2 – Geographic Coordinate system to UTM, elevations units are in feet

Follow all steps as in example 1 and produce an interim layer with x,y units in meters, and z units in feet.

Open Spatial Analyst Tools
Math
Times
Use a conversion factor 0.3048 to convert feet to meters

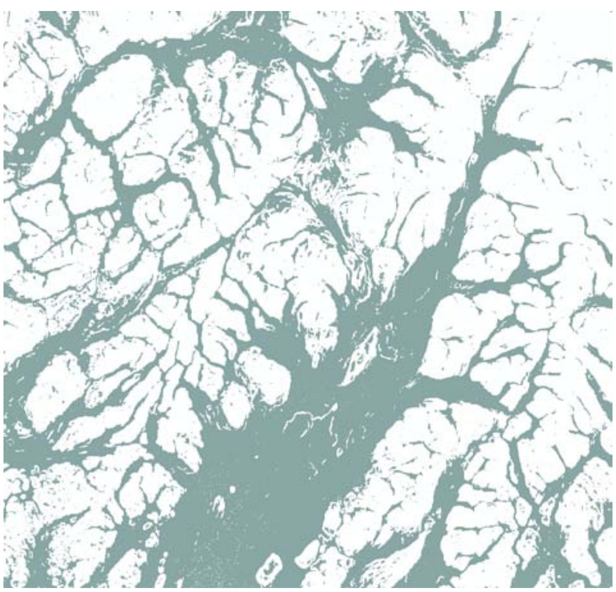


The resulting file has all coordinates in meters and is ready for further analyses.

A sample of methods to create raster subsets

The problem:

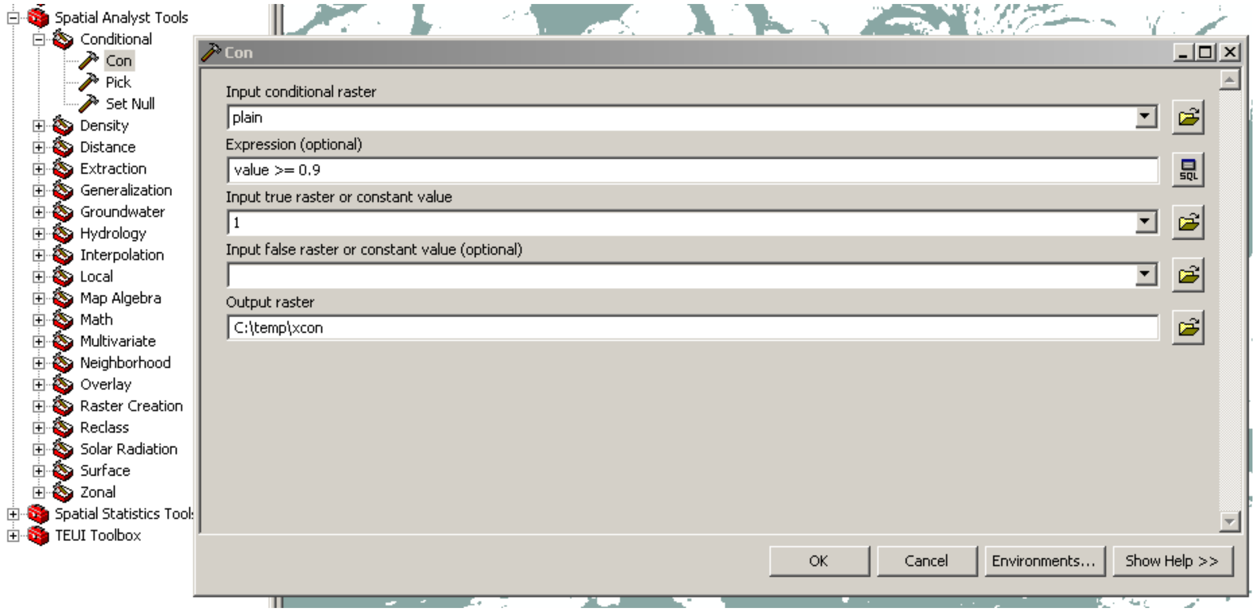
You have the following DEM and wish to exclude the valleys



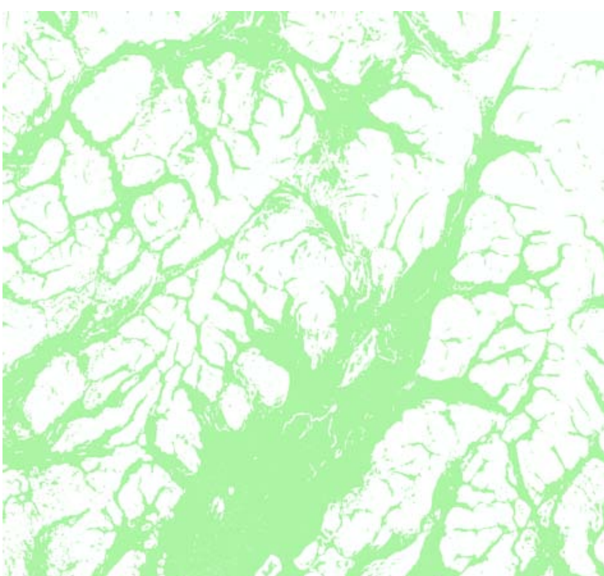
You have a layer depicting “valleys”, with values ranging from 0 to 1, with increasing values being more like a “valley”. You decide values from 0.9 to 1 represent valleys, as depicted above.

Method 1 - Extract by Mask

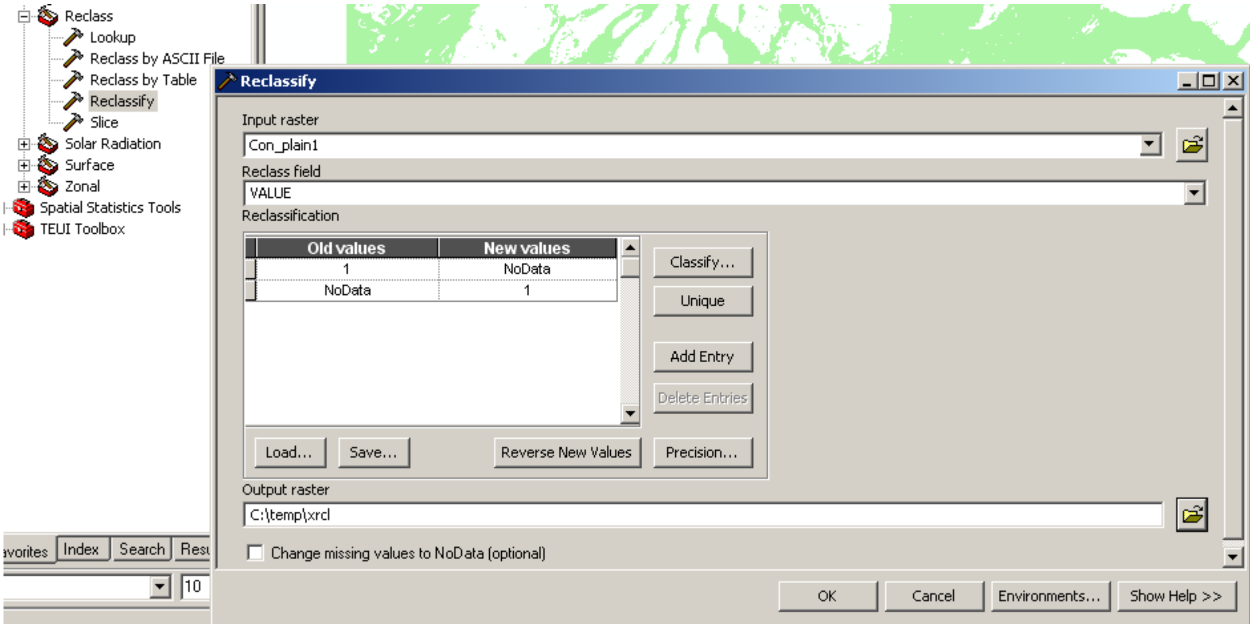
Open the **Con** tool under **Conditional** Toolbox



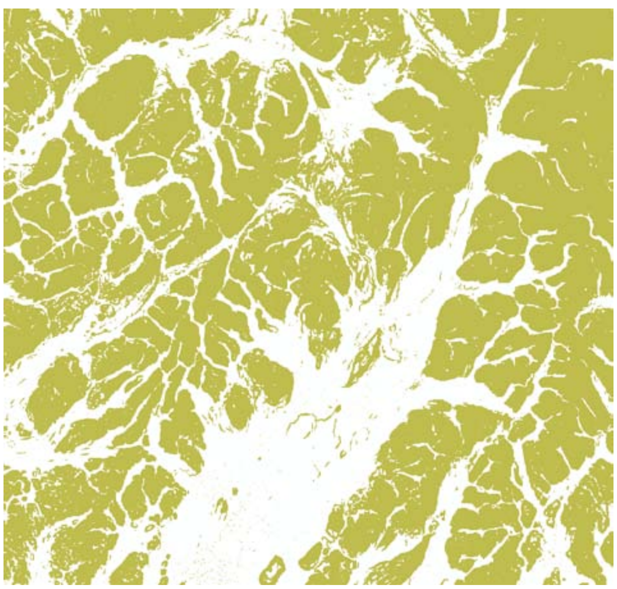
The condition will be values greater than or equal to 0.9. When that condition is met, a value of 1 will be assigned to the new raster “xcon”. Everything else in “xcon” will receive NoData.


Results of the Con statement, Green = 1, everything else = NoData

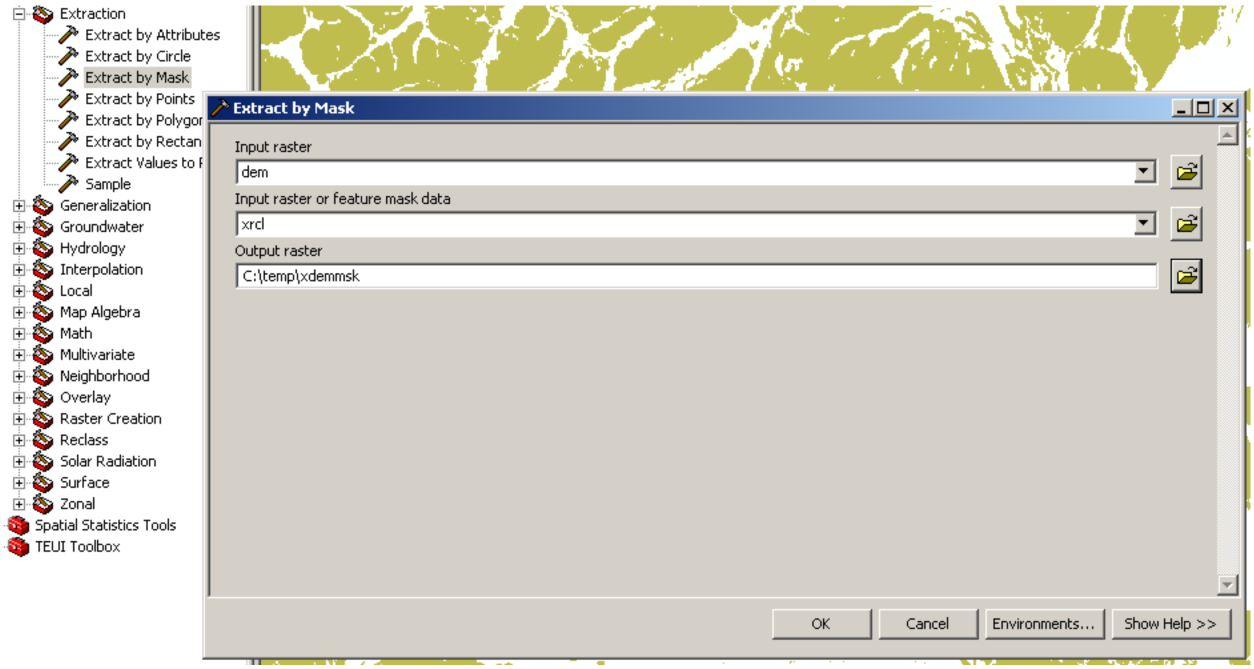
You want the “NoData” areas, so use Reclass to convert NoData to 1



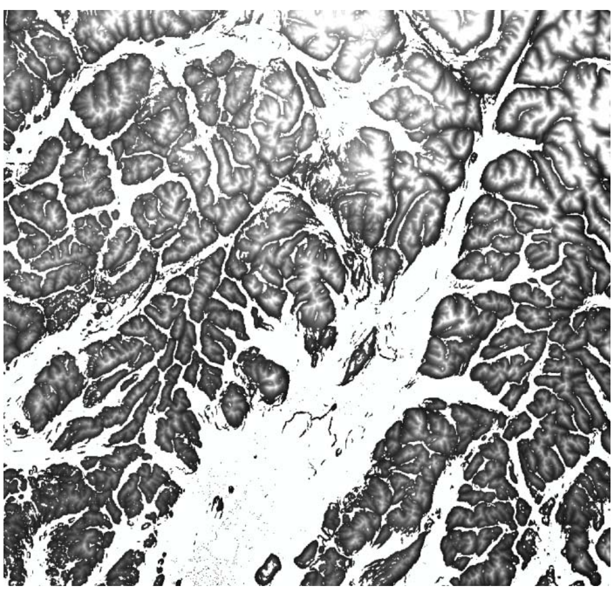
Giving you what you want



Use **Extract by Mask** to get DEM of “non-valley” areas



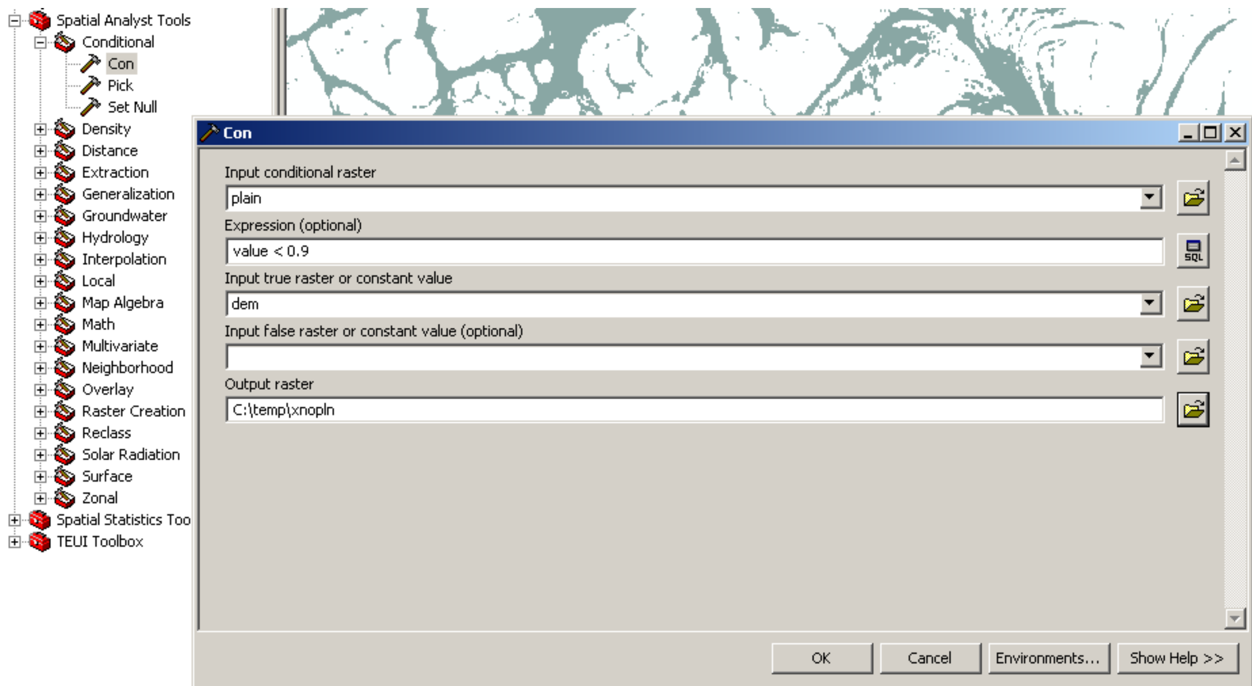
DEM without “valleys”



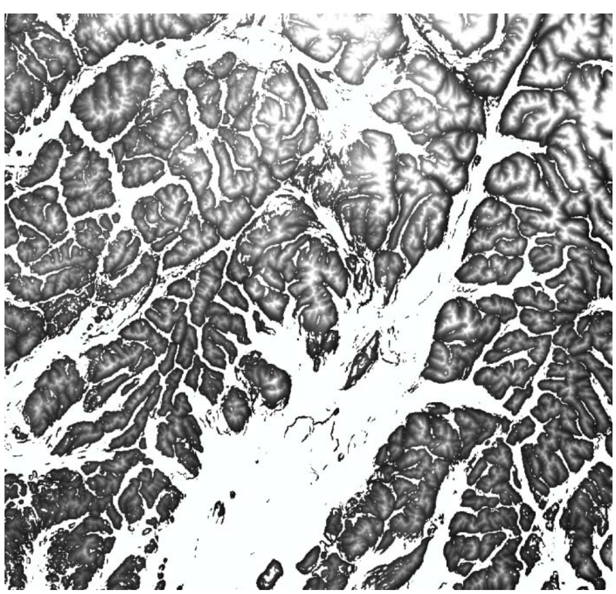
Isn’t there a more direct way?

Method 2 – Con statement

If Values greater than 0.9 are “valleys”, we could select everything less than 0.9, set the true raster to **dem**, and get what we want in one step. Open the **Con** tool in the **Conditional Toolbox**, and change the expression to **Values < 0.9** and the **true raster** to the DEM.



DEM without “valleys” in one step

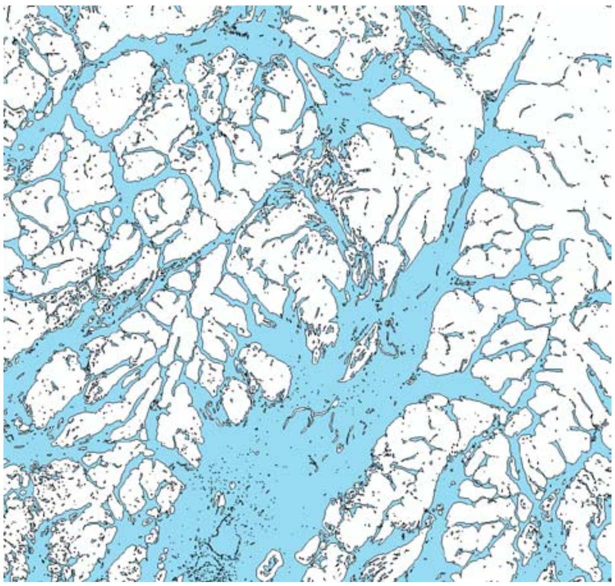


You could change the expression to **value >= 0.9** to get DEM of “valley” only

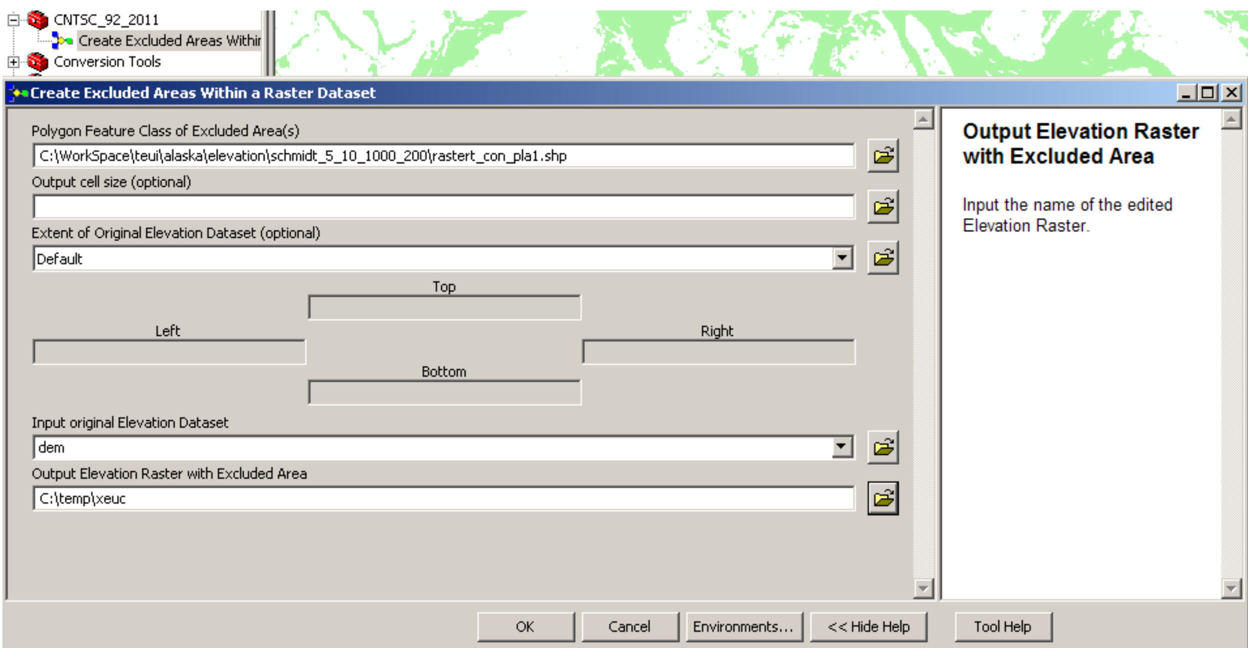
Method 3 – Exclude areas with CNTSC tool, which uses Set Null and Euclidean Distance

http://www.nrcs.usda.gov/wps/PA_NRCSConsumption/download?cid=stelprdb1258045&ext=tbx

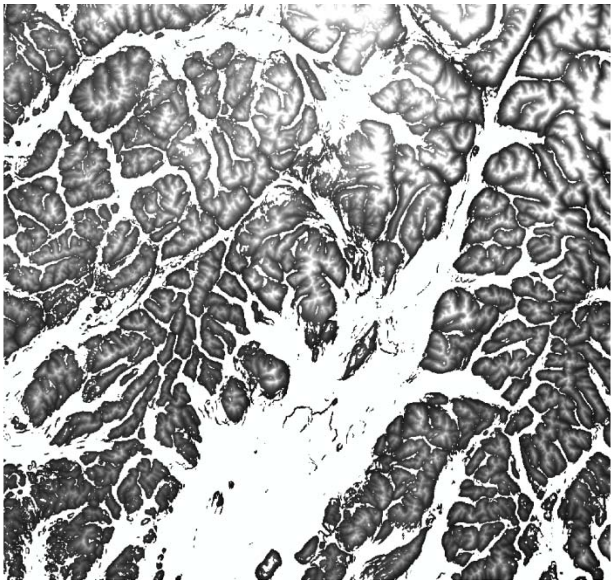
You have a polygon file with "valleys" identified, which you want to exclude



Input the polygon file, input DEM, and output file in GUI



Results are elevations without “valleys”



Resolution and Neighborhoods

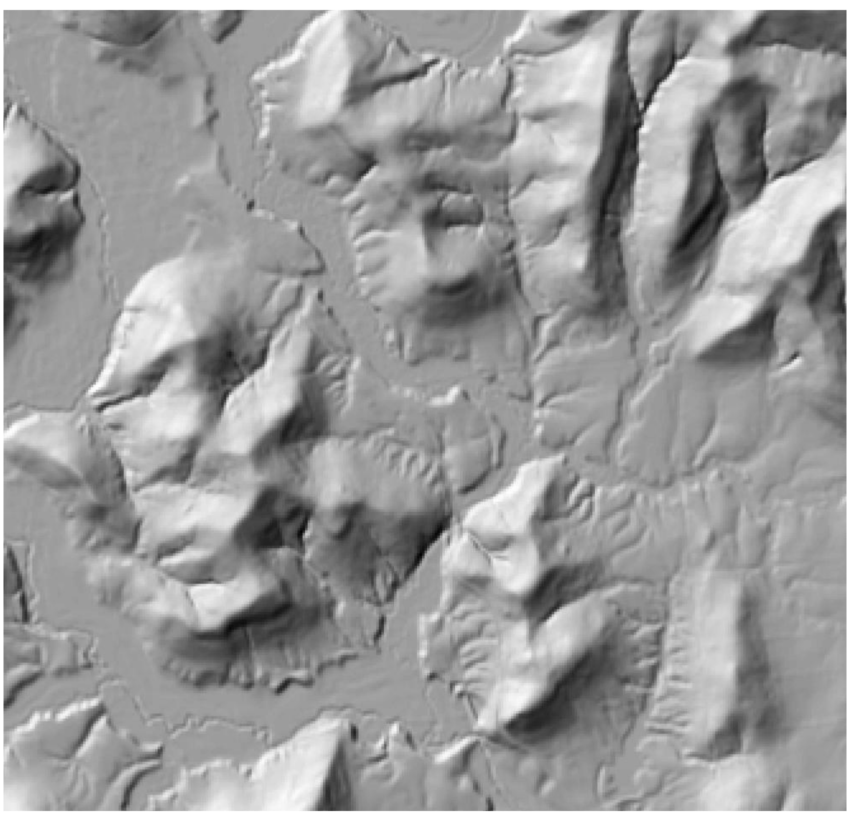
GIS software uses two data models to represent reality, vector and raster. The vector model is most familiar to people, and uses point, lines and polygons to represent features. The raster model uses a regular grid of specified dimension to represent features. Modeling efforts are typically performed in the raster environment.

Definitions

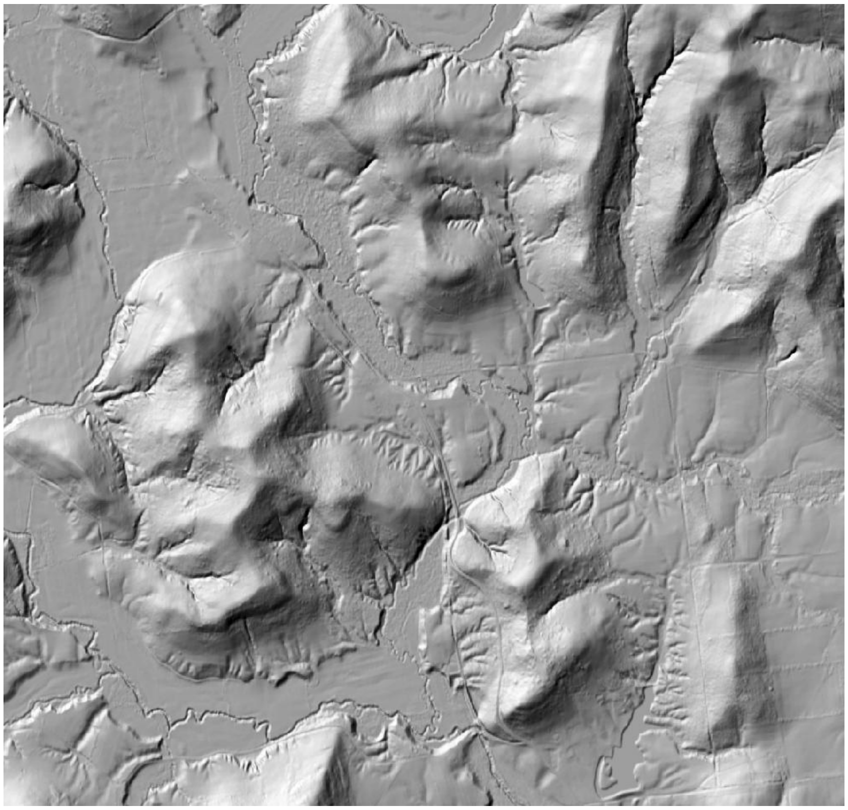
Resolution – the dimension of a grid cell. High resolution refers to small grid cell dimensions, more grid cells and larger file size. Higher resolution generally corresponds with the higher definition of features. Some common raster datasets and their typical resolutions include the 30, 10 and 3 meter USGS National Elevation Data, 30 meter TM satellite imagery, 2 and 1 meter NAIP imagery and 5 meter IFSAR data. DEMs produced from LiDAR data are typically delivered at resolutions of 3 meters and smaller.

Neighborhood – the analysis window used for many raster operations such as calculation of slope, curvature and aspect. Most GIS packages use a neighborhood of 3 x 3 cells for these operations, which can not be altered by the user. ArcSIE is an extension for ArcGIS that supports floating neighborhoods, which means a user may select a neighborhood size independent of the cell resolution.

Resolution example

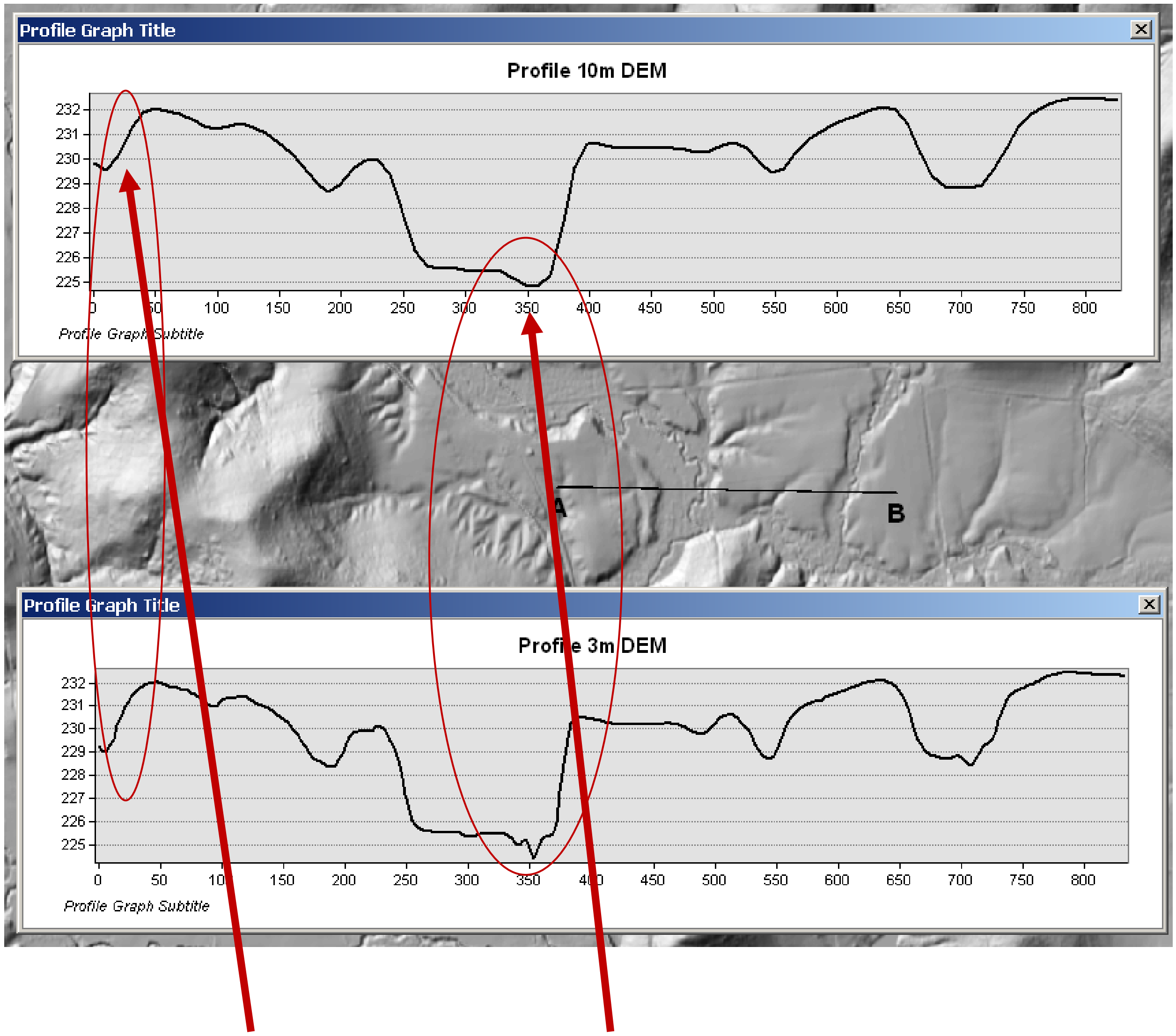


Hillshade produced with 10 meter DEM



Hillshade produced with 3 meter DEM

The hillshade from the 10 meter DEM appears smoother than the 3 meter DEM. That “smoothing” may be a negative if ones goal is to capture the finer features depicted in the 3 meter data, as shown in the cross-sections below, or if slope classes derived from the 10 meter DEM don’t correspond to the actual terrain.



In addition, gradient is typically less on steeper slopes, and short-term variation is muted in the 10 meter DEM when compared to the 3 meter DEM.

Optimal resolution

Selection of resolution is often determined by the data one has available. However, if LiDAR is available, or point data is being developed into a surface, the user has some flexibility in selecting the resolution for a project. The resolution choice will be a balance of computational efficiency and the representation of desired features for the goal of the mapping project. If 10 meter data satisfies the required mapping objective, higher resolution data will only increase processing time and storage requirements.

When high resolution data is available, the choice of using a coarser resolution becomes an option. The selected resolution should represent the features required for the objectives of the mapping project. There are no absolute rules for determination of an optimal resolution. In practice, soil scientists often utilize multi-resolution data, with the original, high resolution data used for visualization, and resampled, coarser resolution data for modeling and mapping.

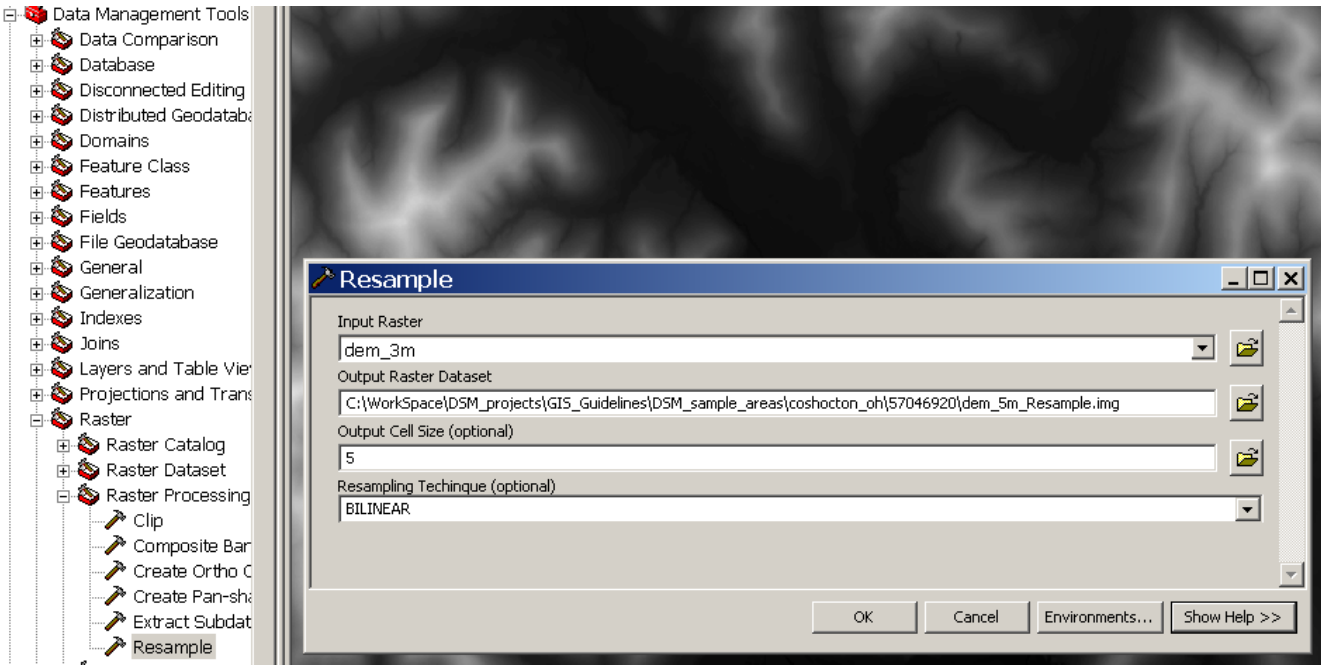
A useful guide for selecting grid resolution was proposed by Hengl (2006):

Resolution = Scale Factor x 0.0005.

A 1:12,000 scale mapping project would have a suggested resolution of 6 meters. That would be a good starting point to begin the evaluation of appropriate resolution.

Resampling

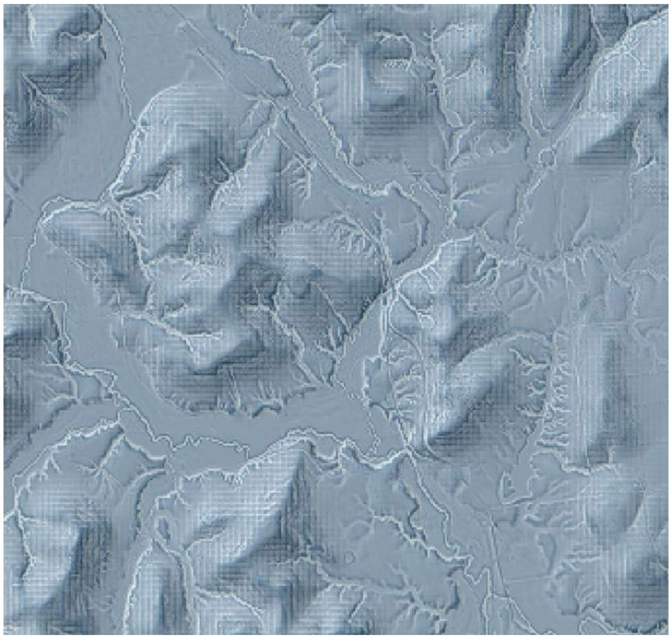
The process of changing cell size is called resampling. If 1 or 3 meter LiDAR data is too detailed for your needs, the resampling tool in ArcGIS can be used to convert to 5, 7 or 10 meter resolution. The Resample tool is located under the **Data Management Tools/Raster/Raster processing Toolbox**:



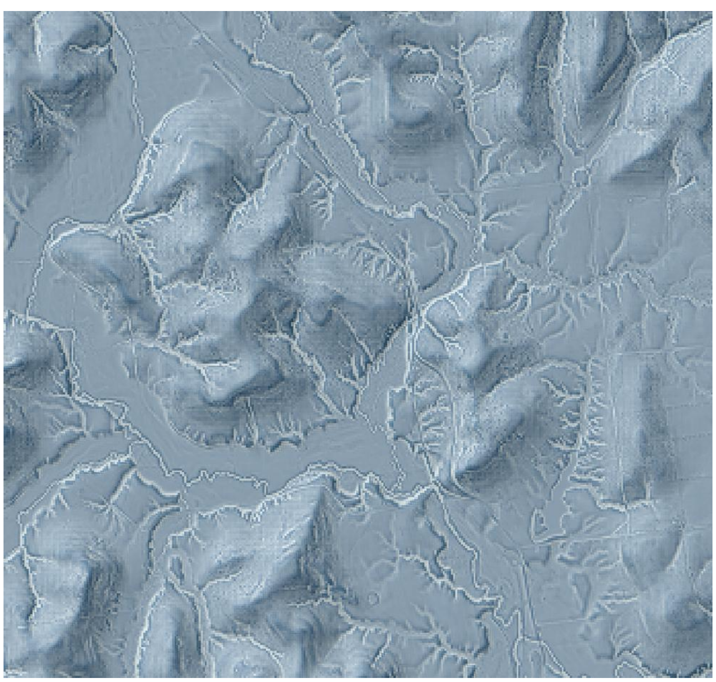
Make sure you select **Bilinear** or **Cubic** for Resampling Technique. Nearest Neighbor or Majority should **never** be selected when using continuous data. There are mathematical differences between the Bilinear and Cubic methods, with Cubic being more complex. As a result, processing time may be longer for the Cubic method. In addition, the Cubic method has a greater potential to extrapolate elevation values compared to the Bilinear method. The Cubic method yields a smoother surface than Bilinear. The differences between the two methods are probably not detectable for soil survey purposes.

(*The nearest neighbor method does not alter the value of the input cells and **should** be used when resampling categoric data, such as surficial geology, imagery or land cover*.)

An example of the resampling output is shown below:



The image above is a curvature map and was produced from a DEM that was resampled from 3 meter to 10 meter resolution using the Nearest Neighbor technique. The blocky pattern introduced by the Nearest Neighbor resampling technique is readily apparent.



The image above is a curvature map and was produced from a DEM that was resampled from 3 meter to 10 meter resolution using the Bilinear technique, without the blocky artifacts.

Neighborhood size

ArcGIS uses a moving window, with a dimension of 3x3 cells for calculation of terrain attributes like slope, curvature and aspect**.**

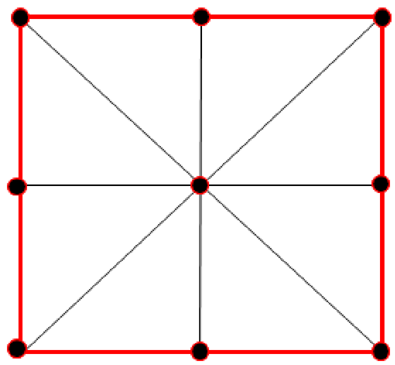
<table>
<tr><td> </td><td> </td><td> </td></tr>
<tr><td> </td><td>**x**</td><td> </td></tr>
<tr><td> </td><td> </td><td> </td></tr>
</table>

This process uses the values from the outside rows and columns to determine the value of the center cell “x”. The window then moves one cell over and the calculations are repeated until all cells are processed. (*Cells along the edge of a map do not get the full benefit of all surrounding neighbors, resulting in the “edge effect” commonly seen in derivative products. Using buffered project areas with plenty of overlap is the best way to minimize the “edge effect” and produce maps that will have reasonable joins.)*

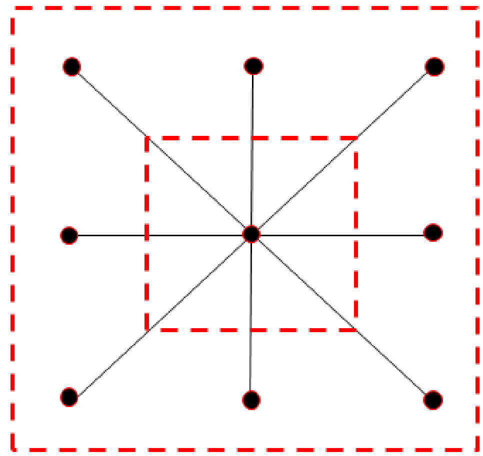
This has been a standard analysis window since the inception of raster based processing. However, it restricts the user to the resolution of their data, introduces directional bias from the diagonal cells and provides only one alternative to changing neighborhood size, resampling to a coarser resolution. What are the alternatives?

Neighborhood size with ArcSIE

ArcSIE supports floating neighborhoods, which are independent of cell size



Cellsize-determined Neighborhood



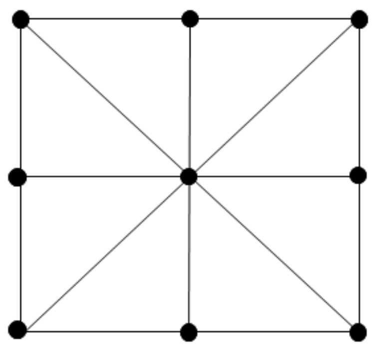
floating neighborhood

- User-specified, not restricted or limited by the resolution of the DEM.
- Do not need to modify the original DEM.
- Can be used to implement different shapes of neighborhood.

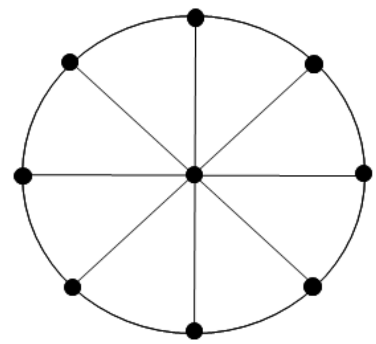
Neighborhood shape

ArcSIE supports circular and square neighborhood shapes, providing more flexibility to the user. The circular neighborhood offers many advantages over the square neighborhood.

Shape of the Neighborhood: Square vs. Circular

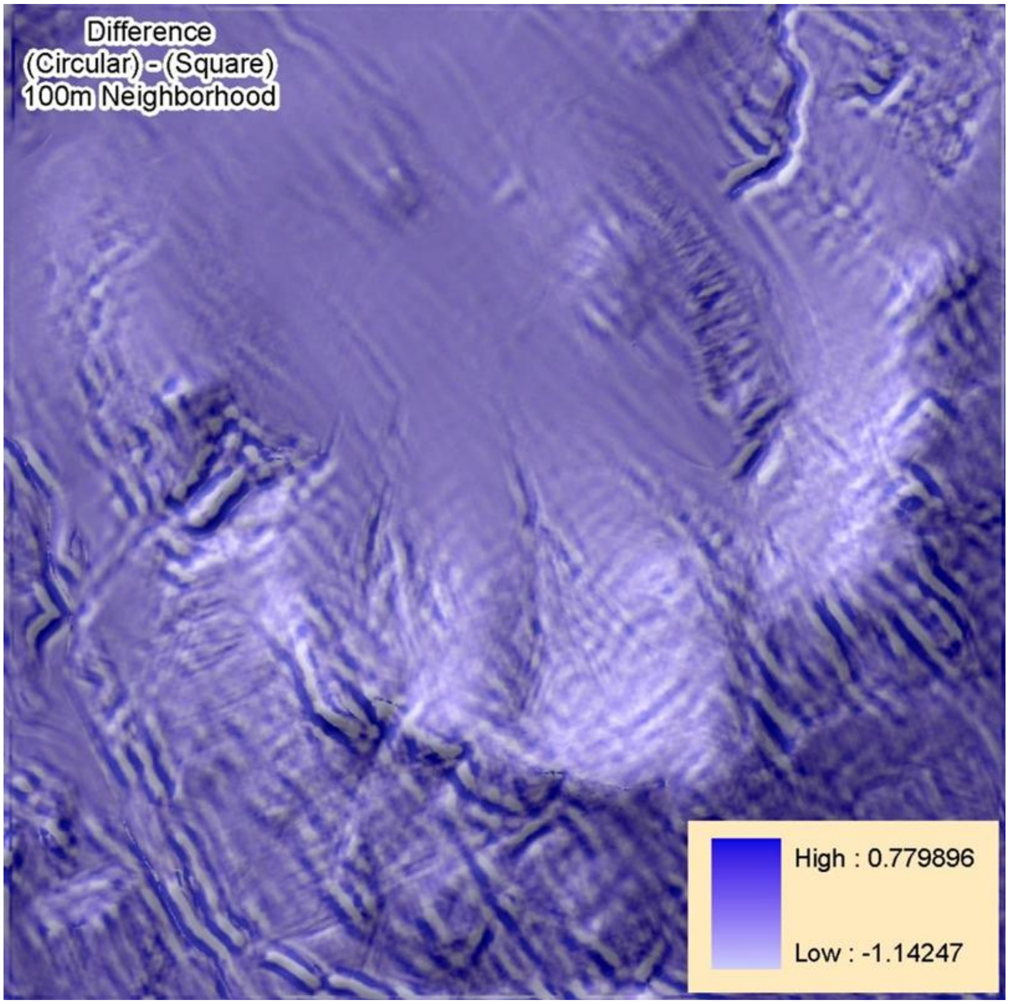


square neighborhood



circular neighborhood

- Reduces directional bias
- More accurate results (when compared to both analytical results for simulated surface and field measurements on real-world terrain)
- The difference between the square and circular neighborhoods becomes more significant when the ratio between neighborhood size and cell size increases. This has important implications on high-resolution DEMs.



An example comparing curvature produced with square vs. circular neighborhoods.

The image above was produced by developing two profile curvature maps; one using a 100 meter circular neighborhood and one using a 100 meter square neighborhood. The square neighborhood map was subtracted from the circular map using the **Minus** tool in the **Math toolbox**. If the curvature maps produced identical results, the "difference" map would be 0. If there are differences and the error is random, we should not see a pattern. The image above shows differences and a pattern that may indicate bias along the diagonals of the square neighborhood (Zhou, 2004).

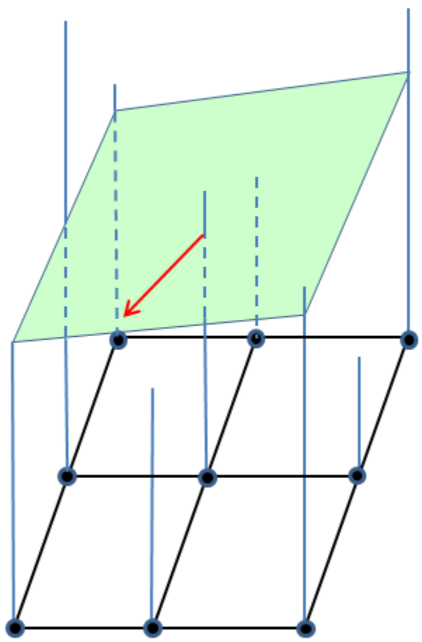
The square neighborhood introduces bias from the unequal distances between corner cells (c) and adjacent (a) cells to the analysis cell.

c	**a**	**c**
a		**a**
c	**a**	**c**

A circular neighborhood would have a uniform distance around the analysis cell.

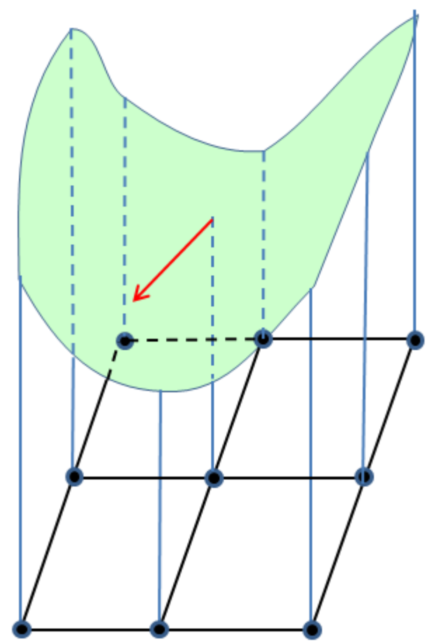
Neighborhood calculation methods

ArcSIE offers several algorithms for calculating terrain attributes, providing more choices for the user. Other packages use predetermined algorithms.



Quadratic fitting

Idea: Use a plane to fit the nine elevation values in the neighborhood; Then calculate derivatives for the central location.
Implementation: the Evans-Young Method and the Horn Method (adopted by ArcInfo).
Pros: less sensitive to artifacts in DEM



Lagrange fitting

Idea: Use a twisted surface to fit the nine elevation values in the neighborhood; Then calculate derivatives for the central location.
Implementation: the Zevenbergen-Thorne Method and the Shi Method (recommended by ArcSIE).
Pros: more accurate for smooth surface

References

Burrough, P.A., 1986. Principles of Geographical Information Systems for Land Resource Assessment, Clarendon, Oxford.

Hengl, T., 2006. Finding the right pixel size. Computers and Geosciences, 32: 12833-1298.

http://www.geom.unimelb.edu.au/gisweb/NOModule/NOtheorylesson.htm#Majority

Star, J.L. and J.E. Estes, 1990. Geographic Information Systems: An Introduction, Prentice Hall, Englewood Cliffs, NJ. An introduction to GIS with a strong raster orientation.

Shi, X., Zhu, A., Burt, J., Choi, W., Wang, R., Pei, T., Li, B., Qin, C. 2007. An experiment using a circular neighborhood to calculate slope gradient from a DEM. Photogram.Engin. And Rem, Sens. 73:143-154.

Thompson, J., Bell, J., Butler, C. 2001. Digital elevation model resolution: effects on Terrain attribute calculation and quantitative soil-landscape modeling. Geoderma, 100:67-89.

Zhou, Q., and X. Liu, 2004. Analysis of errors of derived slope and aspect related to DEM data properties. Computers & Geosciences, 30:369–378.

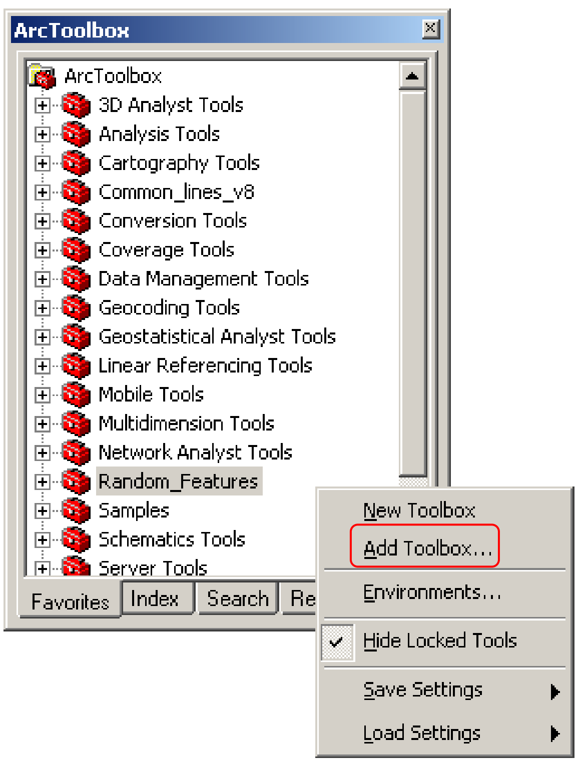
MAKE RANDOM SELCTION OF POLYGONS FOR SOIL SURVEY UPDATES AND EVALUATIONS INSTALLATION:

Download the file:

hhttp://www.nrcs.usda.gov/wps/PA_NRCSConsumption/download?cid=stelprdb1258044&ext=zip

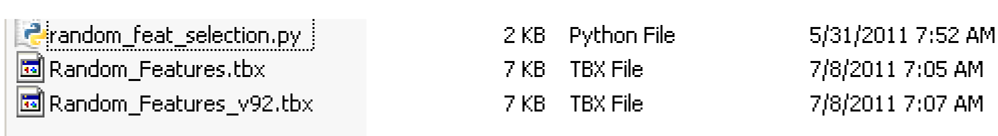
Unzip the file to a directory of your choice.

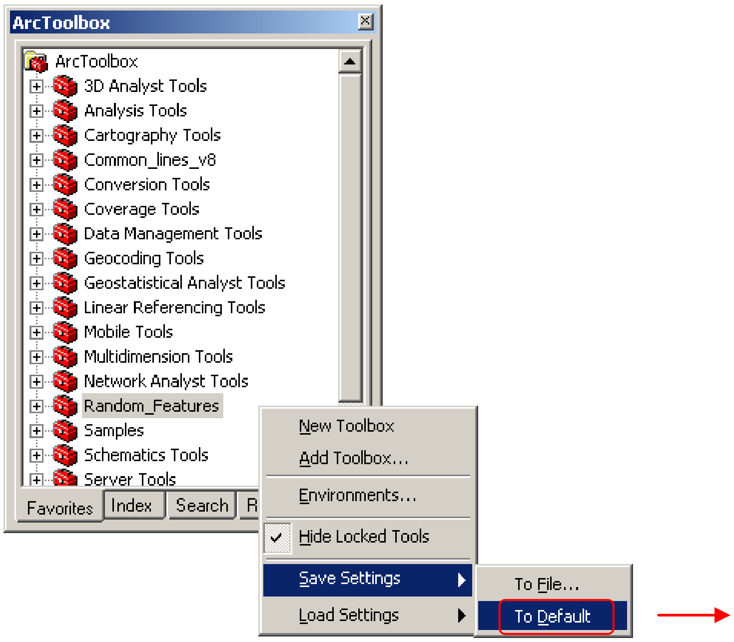
Open ArcMap and open ArcToolbox .



3. Right-click anywhere in the whitespace of ArcToolbox and select **Add Toolbox**

Navigate to where you unzipped the Random Features file and select a toolbox that corresponds with your version of ArcMap and select Open.   Random_Features is for ArcMap version 9.3 and Random_Features_v92 is for ArcMap version 9.2. (To determine which version of ArcMap you are running go to Help > About ArcMap.)



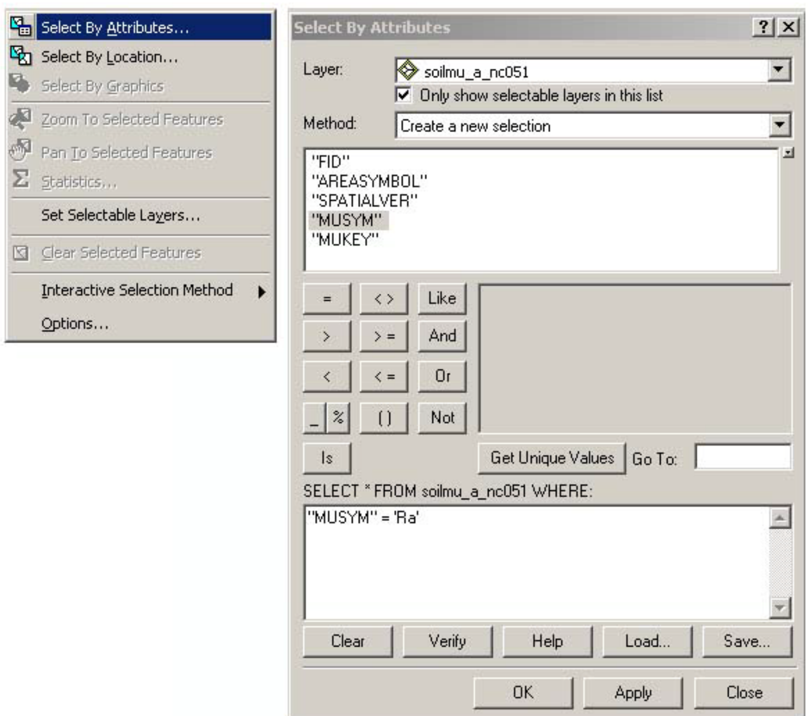


5. In order to save the Random_Features Toolbox permanently into ArcMap, right-click anywhere in the whitespace of ArcToolbox and select **Save Settings > To Default**. The tool should now load every time you open ArcMap. (If ArcCatalog is also open, it is good practice to repeat this process in ArcToolbox there as well.)

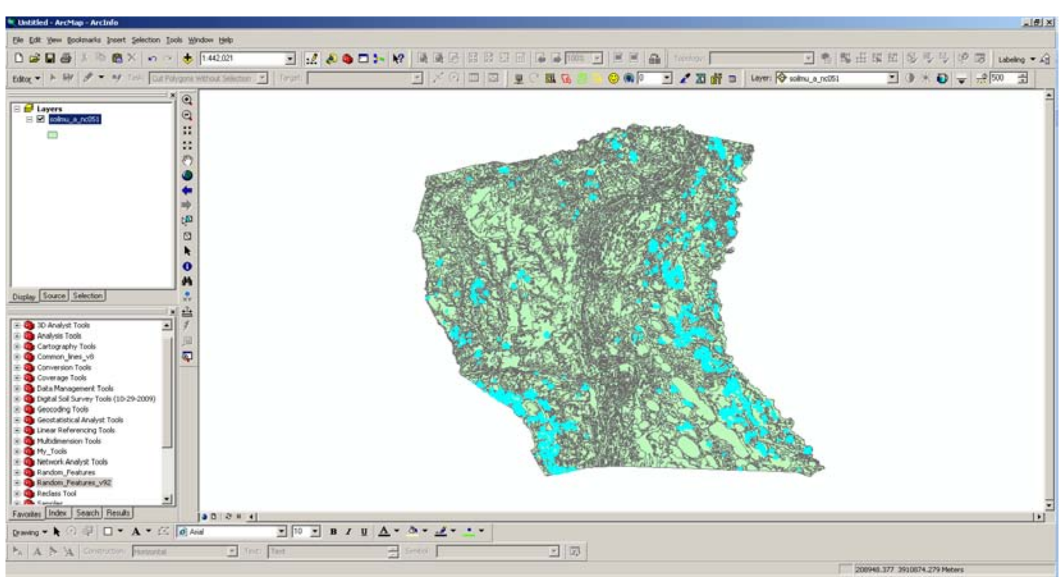
USING THE TOOL:

The Random Features tool will run on a Feature Layer (Geodatabase or Shapefile) that may or may not currently have selected features. Normally, a soil scientist would want to target an individual mapunit and randomly select features for evaluation or documentation purposes. As in the following example.

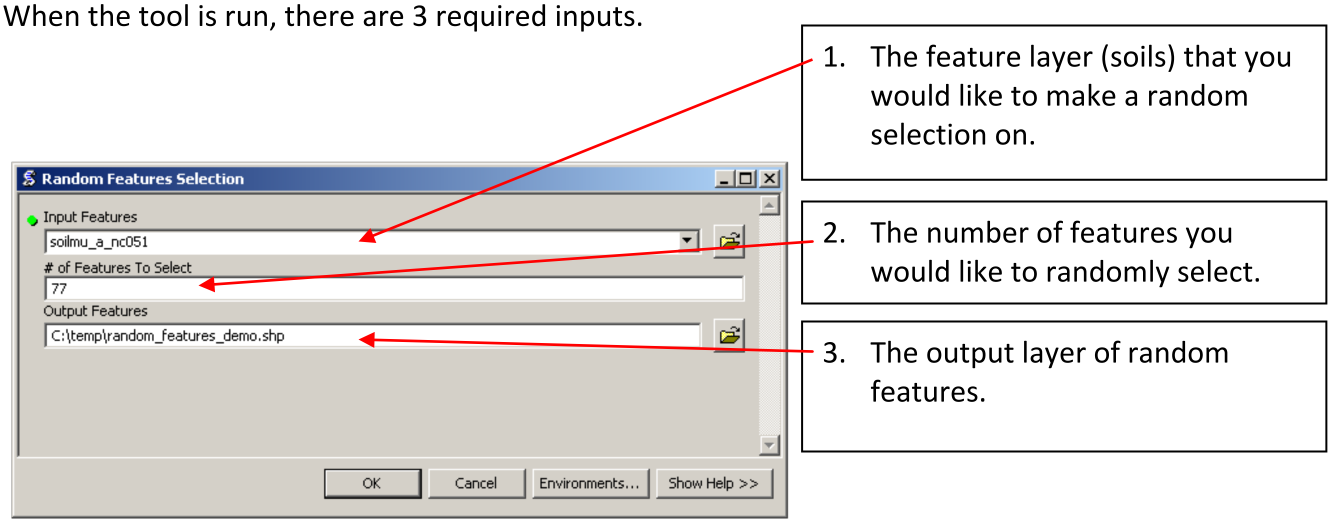
Select the mapunit of interest using Selection > Select by Attributes.



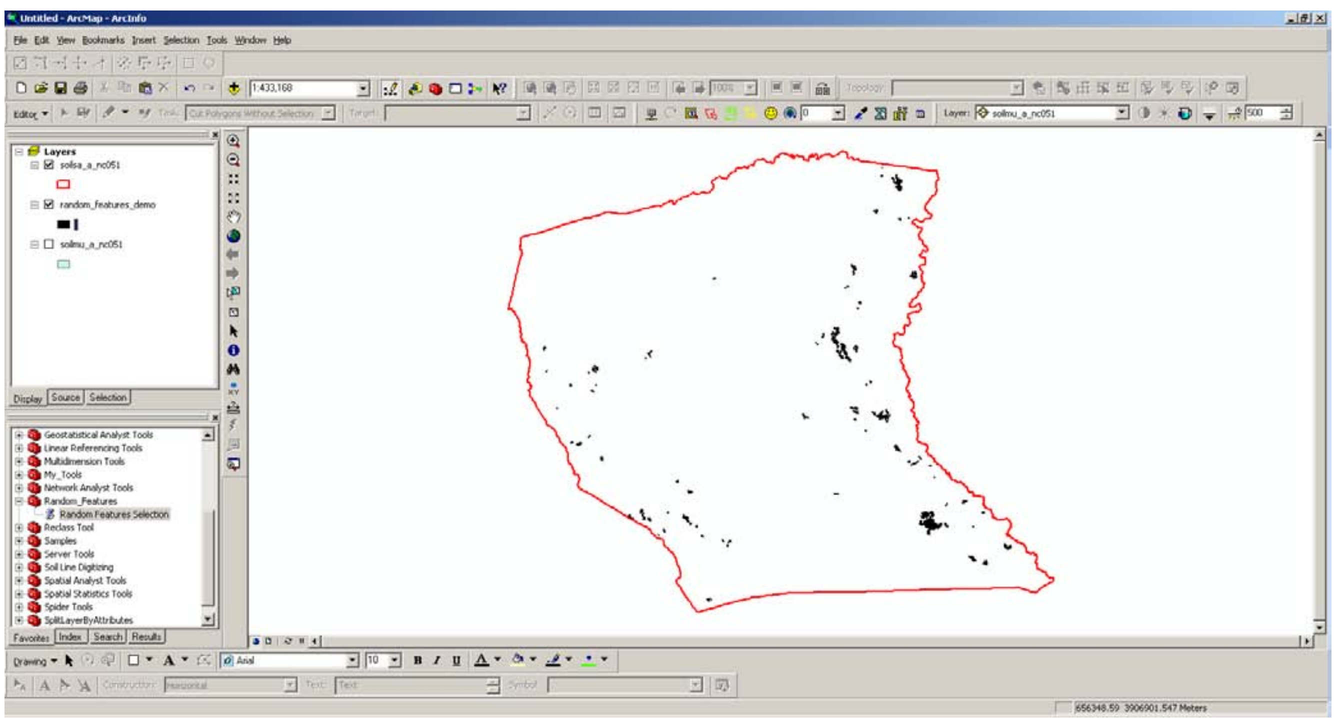
Resulting in 381 selected features.



When the tool is run, there are 3 required inputs.



If the layer does not automatically load, you will need to add it using the Add Data  button or grab it from ArcCatalog.



CONSIDERTIONS:

You will likely want to randomly select a greater number of features than you are actually targeting as some of the selected features will be impractical to get to (distance from road, a shopping center now covers one of the features, etc.)

You are not required to have features selected to use the tool.

QUESTIONS: charles.ferguson@nc.usda.gov

Multiresolution Index of Valley Bottom Flatness (MRVBF)

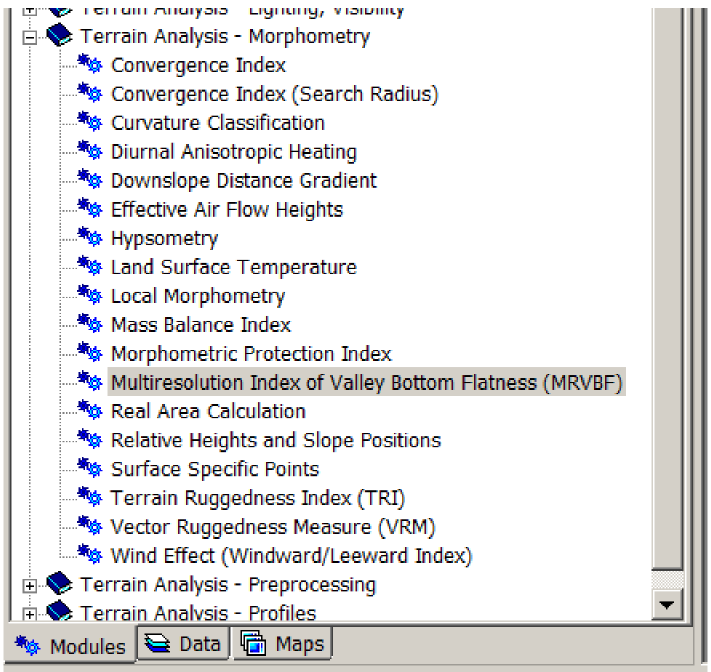
Background

This module identifies valley bottoms from digital elevation models. There is a corresponding routine within the module to identify ridge tops. The full reference for the publication is:

Gallant, J.C., Dowling, T.I. 2003. A multiresolution index of valley bottom flatness for mapping depositional areas. Water Resources Research, 39(12), 1347.

The model uses slope and elevation to classify valley bottoms as flat, low areas. This is accomplished through a series of neighborhood operations at progressively coarser resolutions with the goal of identifying both small and large valleys.

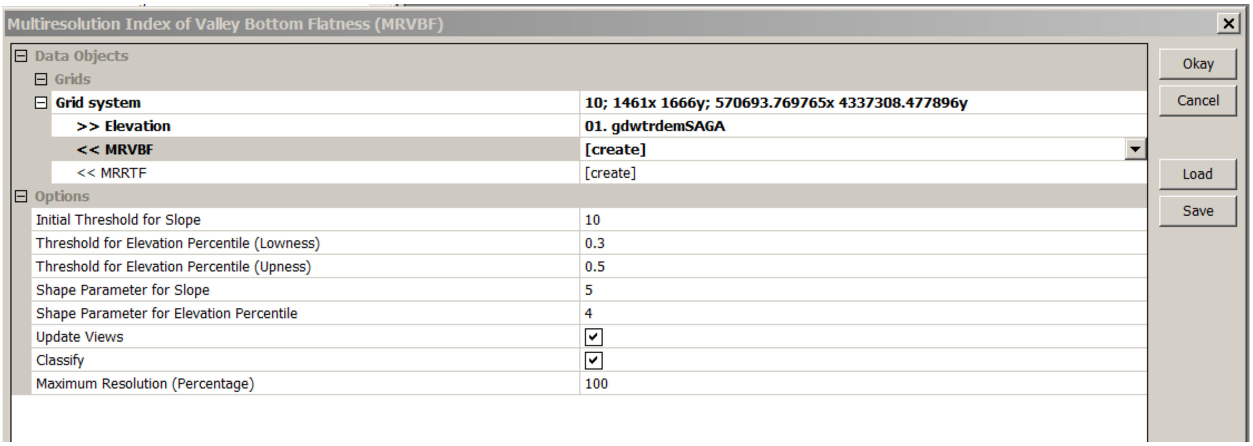
Available under the *Terrain Analysis – Morphometry* Module in SAGA



MRVBF GUI

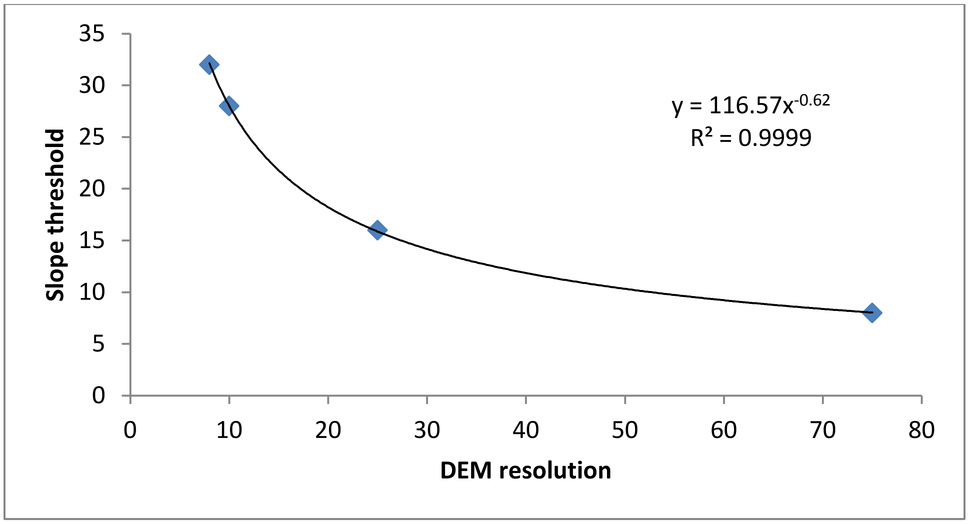
Parameters of the model:

1) **Initial threshold for slope**
2) **Threshold for elevation percentile (Lowness)**
3) **Threshold for elevation percentile (Upness)**
4) **Shape parameter for slope**
5) **Shape parameter for elevation**



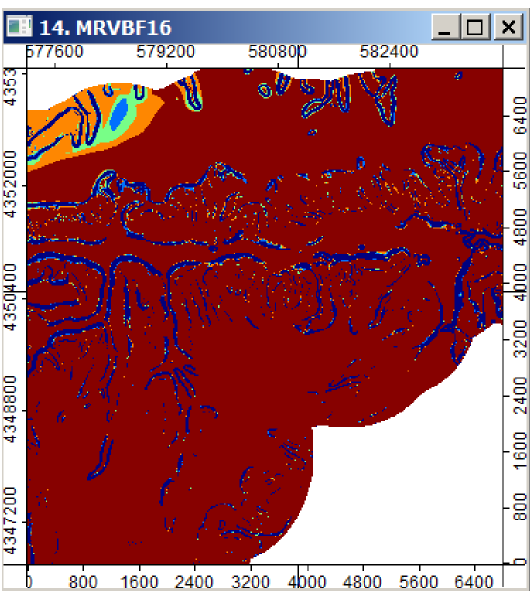
Initial Threshold for slope

“The MRVBF algorithm was developed using 25 m resolution DEMs but can be applied at any resolution provided appropriate adjustments are made. The link between size and flatness of valley bottoms is incorporated into the algorithm by reducing the slope threshold by a factor of 2 at each step, and it is assumed that the relationship between slope threshold, resolution, and MRVBF value does not vary between landscapes or with different resolution DEMs. If the DEM resolution is substantially different from 25 m, the initial slope threshold must be adjusted to retain the relationship between slope and resolution. An initial resolution of 75 m, one resolution step larger than the base 25m resolution, would use a slope threshold of 8% instead of 16% for the first step, while an initial resolution of 8 m, one step smaller than the base, would use a slope threshold of 32% for the first step.” Gallant and Dowling, 2003

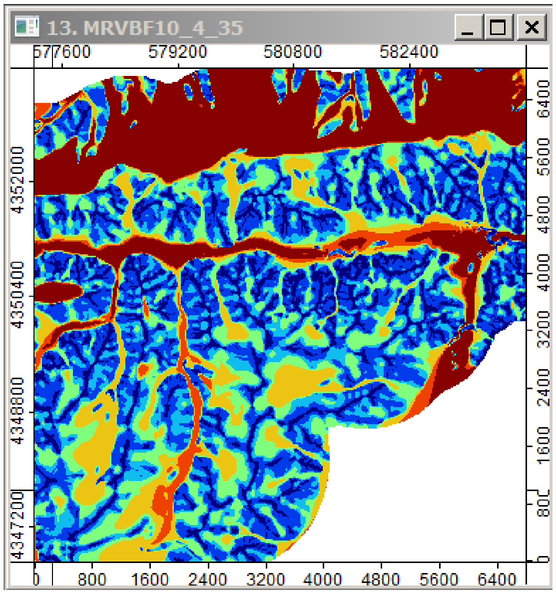


Use the relationship above to select a starting slope threshold based on the resolution of your DEM.

In this example, the default slope threshold of 16 was chosen:



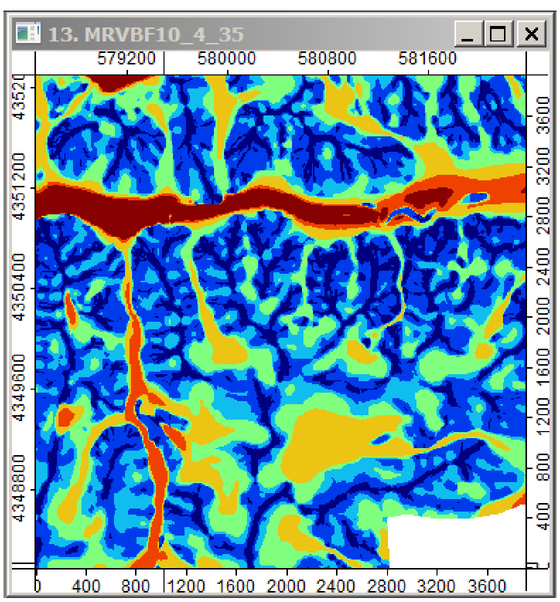
The brown colors are “valley bottoms”, and would be considered grossly over mapped. The sample area is in the Midwest, has 43 meters of local relief, and a maximum slope of about 20 percent.



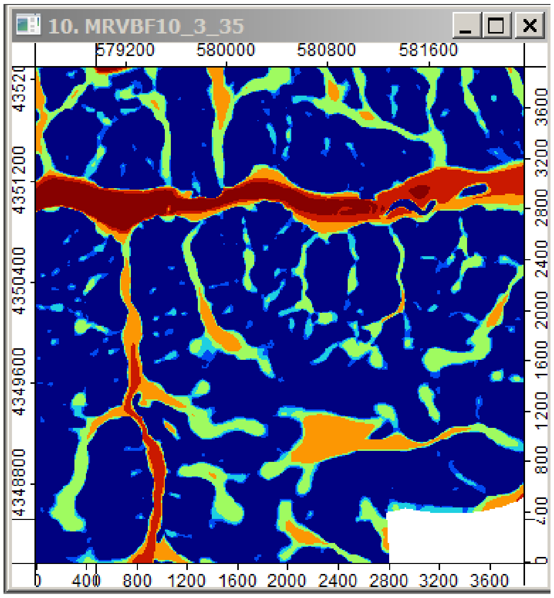
A value of 10 for the slope threshold is more realistic, and also identifies some of the narrow tributaries.

Threshold for Elevation Percentile (Lowness) and (Upness)

The lower the value the more conservative the classification for the respective Valley or Ridge Top. The following examples demonstrate the changes for Valley Bottoms:

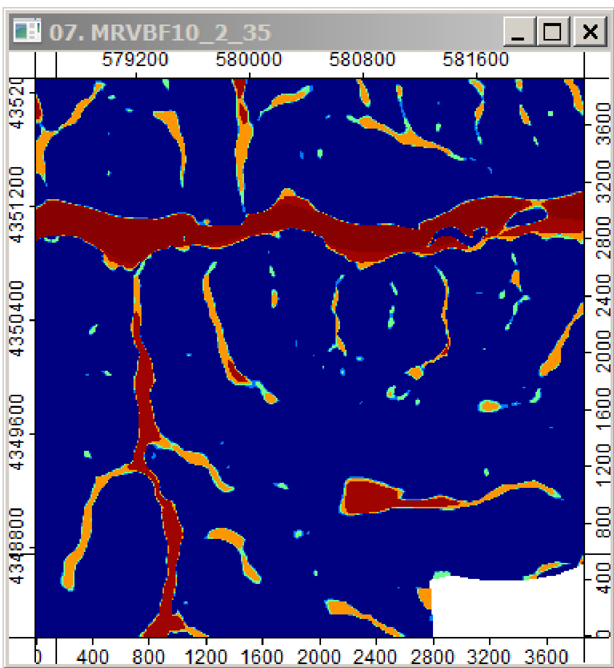


Default Elevation Percentile (Lowness) = **0.4**



Elevation Percentile (Lowness) = **0.3**

Notice some of the smaller valleys pinching off, or narrowing.

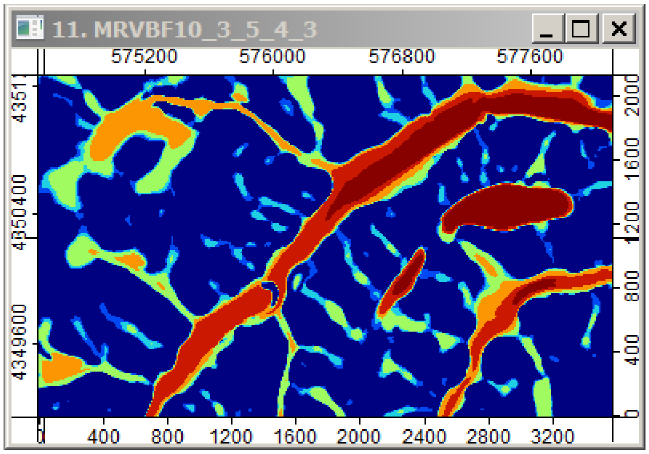


Elevation Percentile (Lowness) = **0.2**

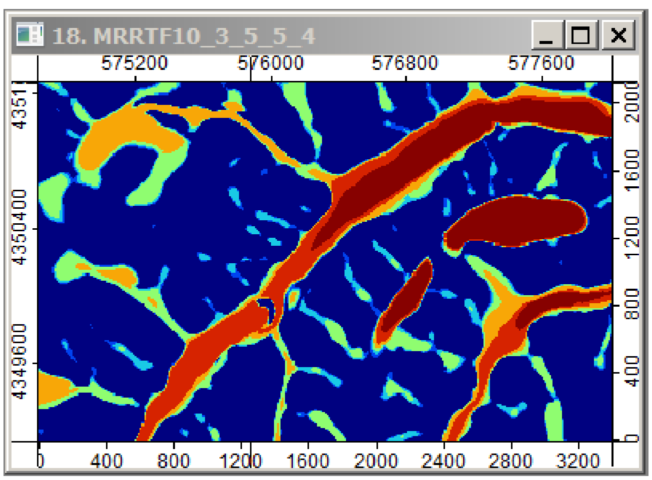
Notice most of the smaller valleys pinching off and narrowing.

Shape parameter function

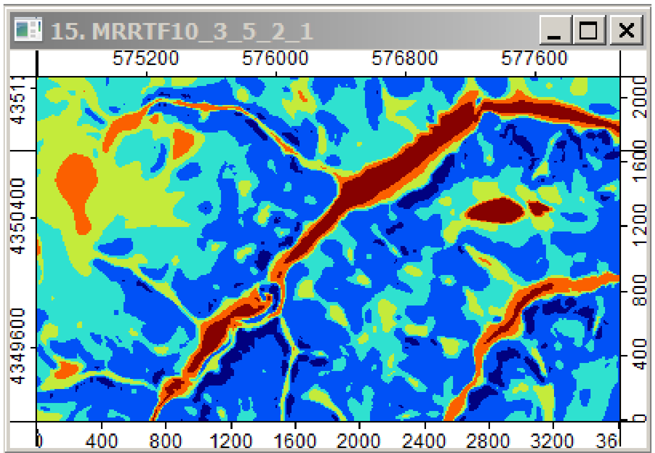
This controls the transition from valley to upland, with lower values being gradual and higher values being abrupt. **Based on experience, the default values of 4 and 3 are fine**. The following screen shots are examples for your information.



Default setting for shape parameter
Slope threshold = 10%
Threshold for elevation percentile (Lowness) = 0.3
Threshold for elevation percentile (Upness) = 0.5
Shape parameter for slope = **4**
Shape parameter for elevation = **3**



Increasing shape parameter – subtle, but results in more abrupt transition from valleys (brown) to "non-valley" (blue)
Slope threshold = 10%
Threshold for elevation percentile (Lowness) = 0.3
Threshold for elevation percentile (Upness) = 0.5
Shape parameter for slope = **5**
Shape parameter for elevation = **4**



Decreasing shape parameters – results in more gradual transition from valleys (brown) to "non-valley" (blue)

Slope threshold = 10%

Threshold for elevation percentile (Lowness) = 0.3

Threshold for elevation percentile (Upness) = 0.5

Shape parameter for slope = **2**

Shape parameter for elevation = **1**

Carbonate Normalized Difference Ratio

This ratio is the same conceptually as the carbonate ratio from the Soil Enhancement Ratios, but normalized instead of a simple ratio. Normalized difference Landsat band ratio using band 3 (VIS-red) and band 2 (VIS-green). (See Soil Enhancement Ratio description)

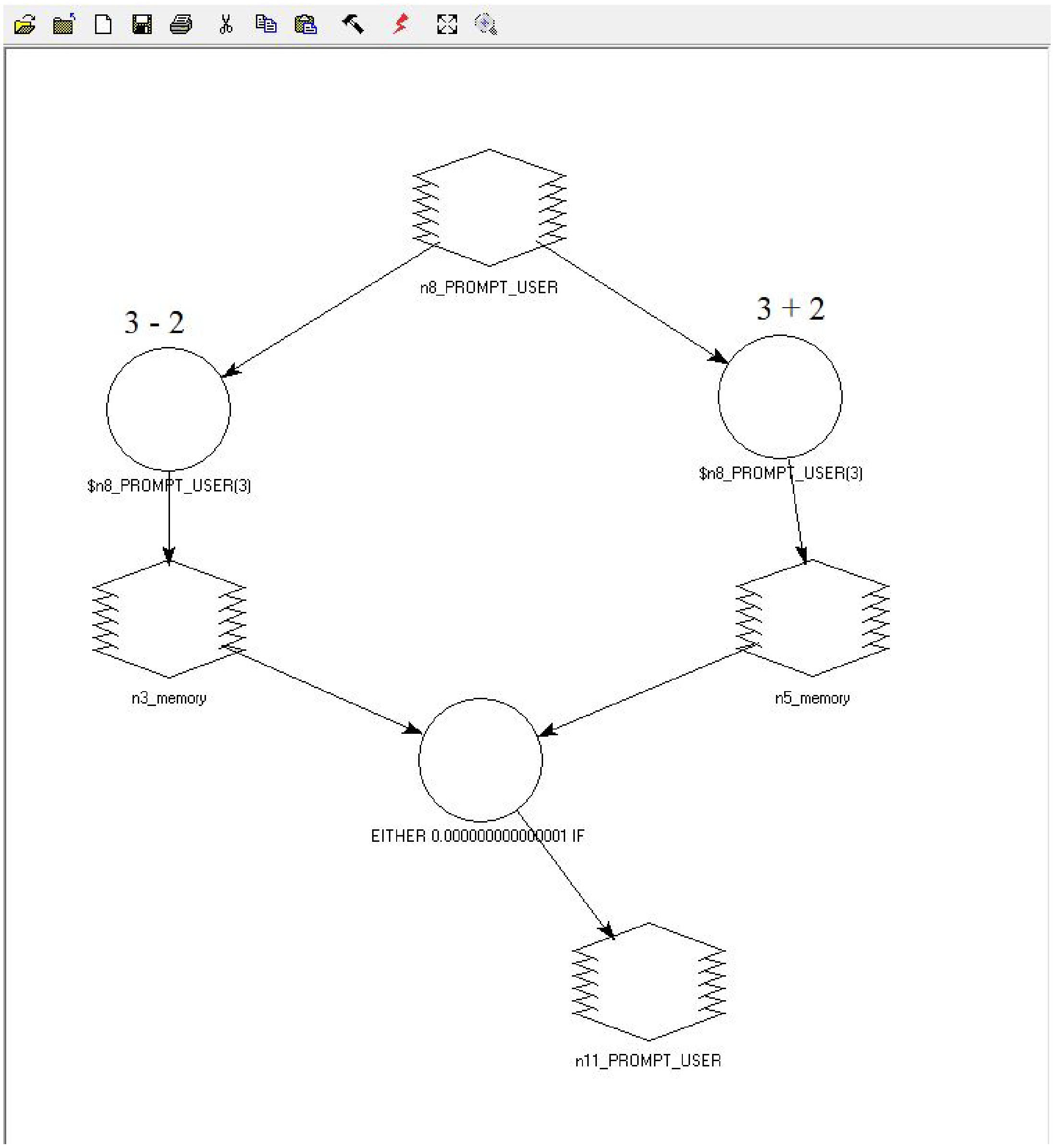
File: carbonate_index_3_2.gmd (ERDAS Imagine Model, v. 2013 and prior)**

Input: Landsat image containing at least bands 2,3*

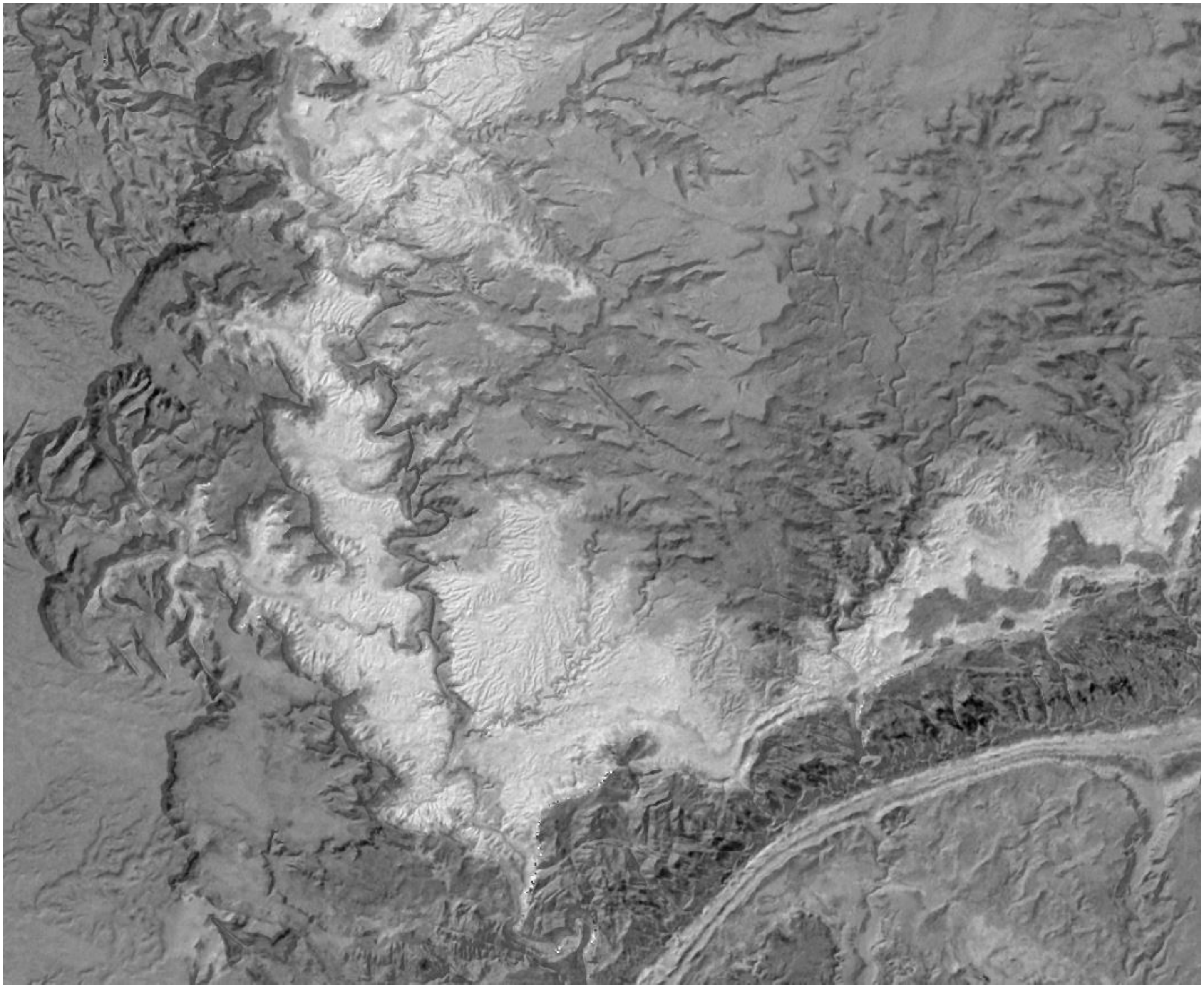
Output: single layer image with data values ranging from -1 to 1; Pixels with higher values can be interpreted as pixels with higher possibility of carbonate influence.

You will be prompted for the input image, and to name the output image.

The model open in ERDAS Imagine Model Maker:



Example normalized carbonate ratio output image over hillshade. Bright pixels have higher data values and can be interpreted as areas of potentially higher carbonate influence:



*Remember ratios must always be calculated on images which have had an atmospheric correction or image standardization applied. Resources for image standardization:
http://earth.gis.usu.edu/imagestd/

**These are .gmd ERDAS Imagine Model Maker files. They can easily be opened in ERDAS 2013 and all previous versions of Imagine. Starting with ERDAS 2014, they must be converted to the new Spatial Modeler interface and file type.

While these ratios were developed for Landsat bands, they could be modified and used with data from any sensor with bands capturing the same part of the electromagnetic spectrum as the indicated Landsat bands.

Enhanced Thematic Mapper Plus (ETM+)	**Landsat 7**	**Wavelength (micrometers)**	**Resolution (meters)**
	Band 1	0.45-0.52	30
	Band 2	0.52-0.60	30
	Band 3	0.63-0.69	30
	Band 4	0.77-0.90	30
	Band 5	1.55-1.75	30
	Band 6	10.40-12.50	60 * (30)
	Band 7	2.09-2.35	30
	Band 8	.52-.90	15

Clay (hydroxyls) Normalized Difference Ratio

This ratio is the same conceptually as the clay (hydroxyls) ratio from the Soil Enhancement Ratios, but normalized instead of a simple ratio. Normalized difference Landsat band ratio using band 5 (shorter wavelength SWIR) and band 7 (longer wavelength SWIR). (See Soil Enhancement Ratio description)

This ratio has also been successfully used in arid environments to indicate areas potentially influenced by gypsum (Nield et al., 2007).

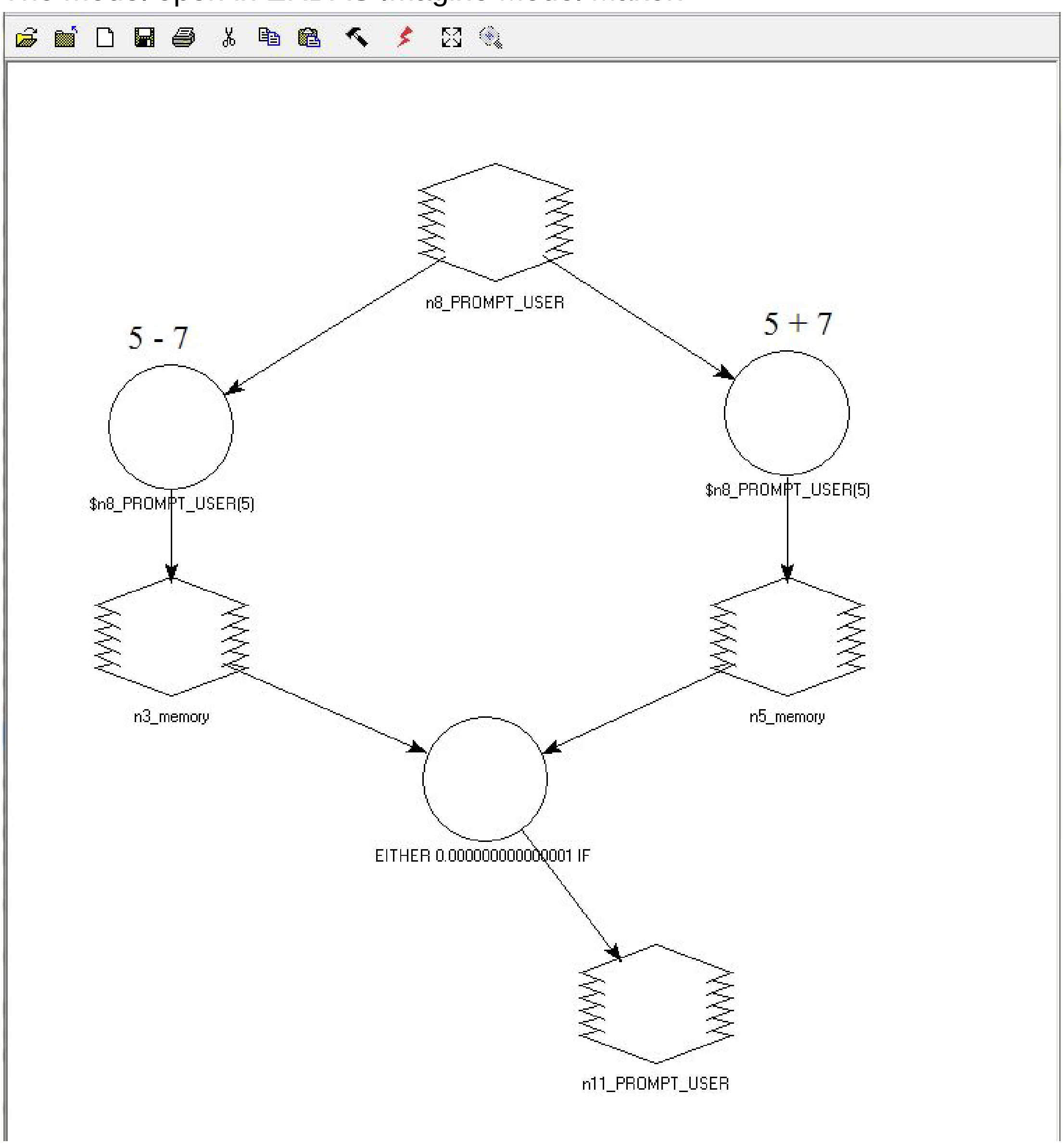
File: iron_index_5_7.gmd (ERDAS Imagine Model, v. 2013 and prior)**

Input: Landsat image containing at least bands 5,7*

Output: single layer image with data values ranging from -1 to 1; Pixels with higher values can be interpreted as pixels with higher possibility of clay (hydroxyls) influence.

You will be prompted for the input image, and to name the output image.

The model open in ERDAS Imagine Model Maker:



Example normalized clay (hydroxyls) ratio output image over hillshade. Bright pixels have higher data values and can be interpreted as areas of potentially higher clay (hydroxyls) influence:

Reference for gypsum ratio: Nield, S.J., Boettinger, J.L., and Ramsey, R.D., 2007. Digitally mapping gypsic and natric soil areas using Landsat ETM data. Soil Science Society of America Journal 71: 245-252.

*Remember ratios must always be calculated on images which have had an atmospheric correction or image standardization applied. Resources for image standardization: http://earth.gis.usu.edu/imagestd/

**These are .gmd ERDAS Imagine Model Maker files. They can easily be opened in ERDAS 2013 and all previous versions of Imagine. Starting with ERDAS 2014, they must be converted to the new Spatial Modeler interface and file type.

While these ratios were developed for Landsat bands, they could be modified and used with data from any sensor with bands capturing the same part of the electromagnetic spectrum as the indicated Landsat bands.

Enhanced Thematic Mapper Plus (ETM+)	**Landsat 7**	**Wavelength (micrometers)**	**Resolution (meters)**
	Band 1	0.45-0.52	30
	Band 2	0.52-0.60	30
	Band 3	0.63-0.69	30
	Band 4	0.77-0.90	30
	Band 5	1.55-1.75	30
	Band 6	10.40-12.50	60 * (30)
	Band 7	2.09-2.35	30
	Band 8	.52-.90	15

Ferrous Minerals Normalized Difference Ratio

This ratio is the same conceptually as the ferrous minerals ratio available in ERDAS Imagine Spectral Indices Tool, but normalized instead of a simple ratio. Normalized difference Landsat band ratio using band 5 (shorter wavelength SWIR) and band 4 (NIR).

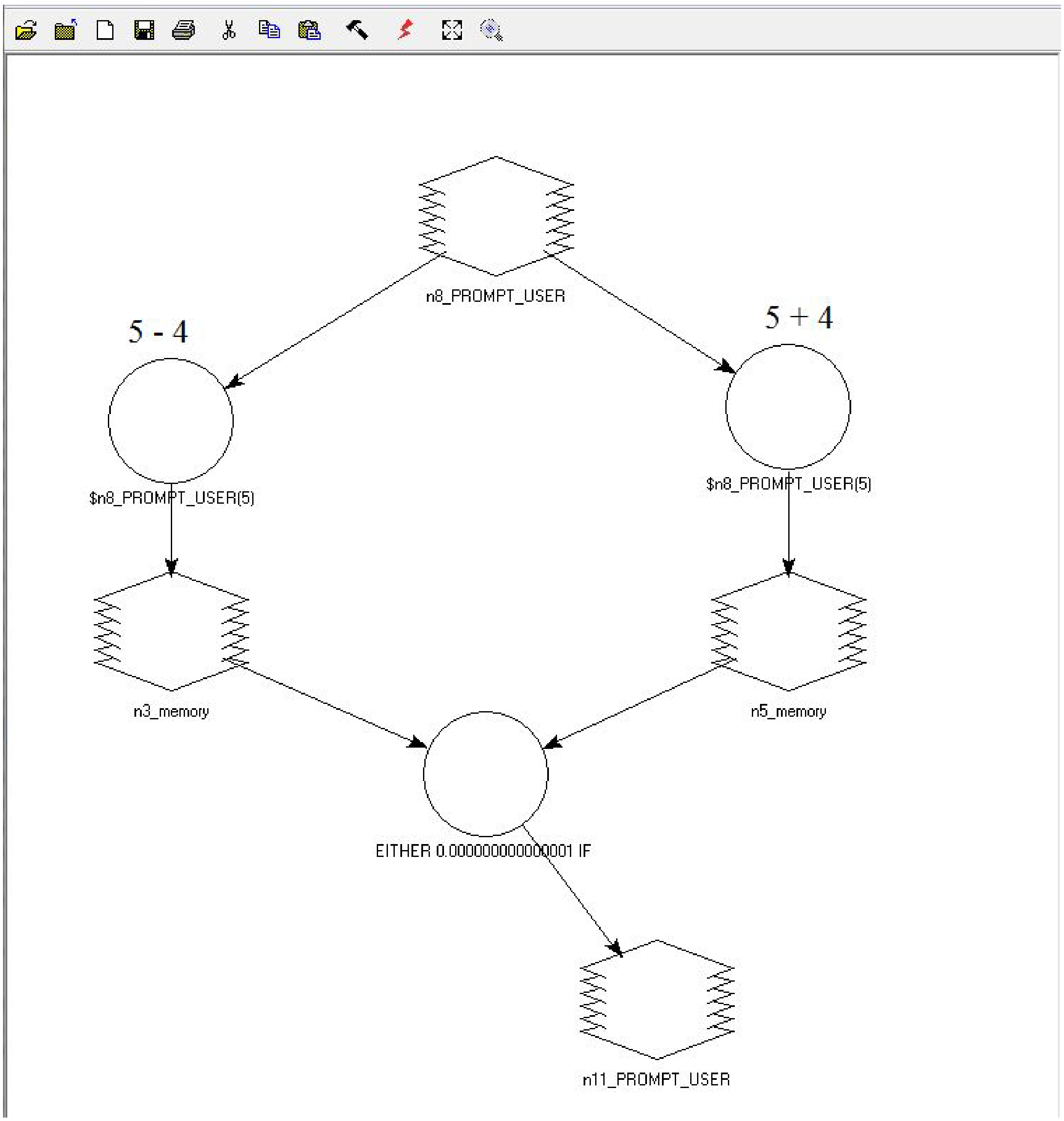
File: ferrous_index_5_4.gmd (ERDAS Imagine Model, v. 2013 and prior)**

Input: Landsat image containing at least bands 4,5*

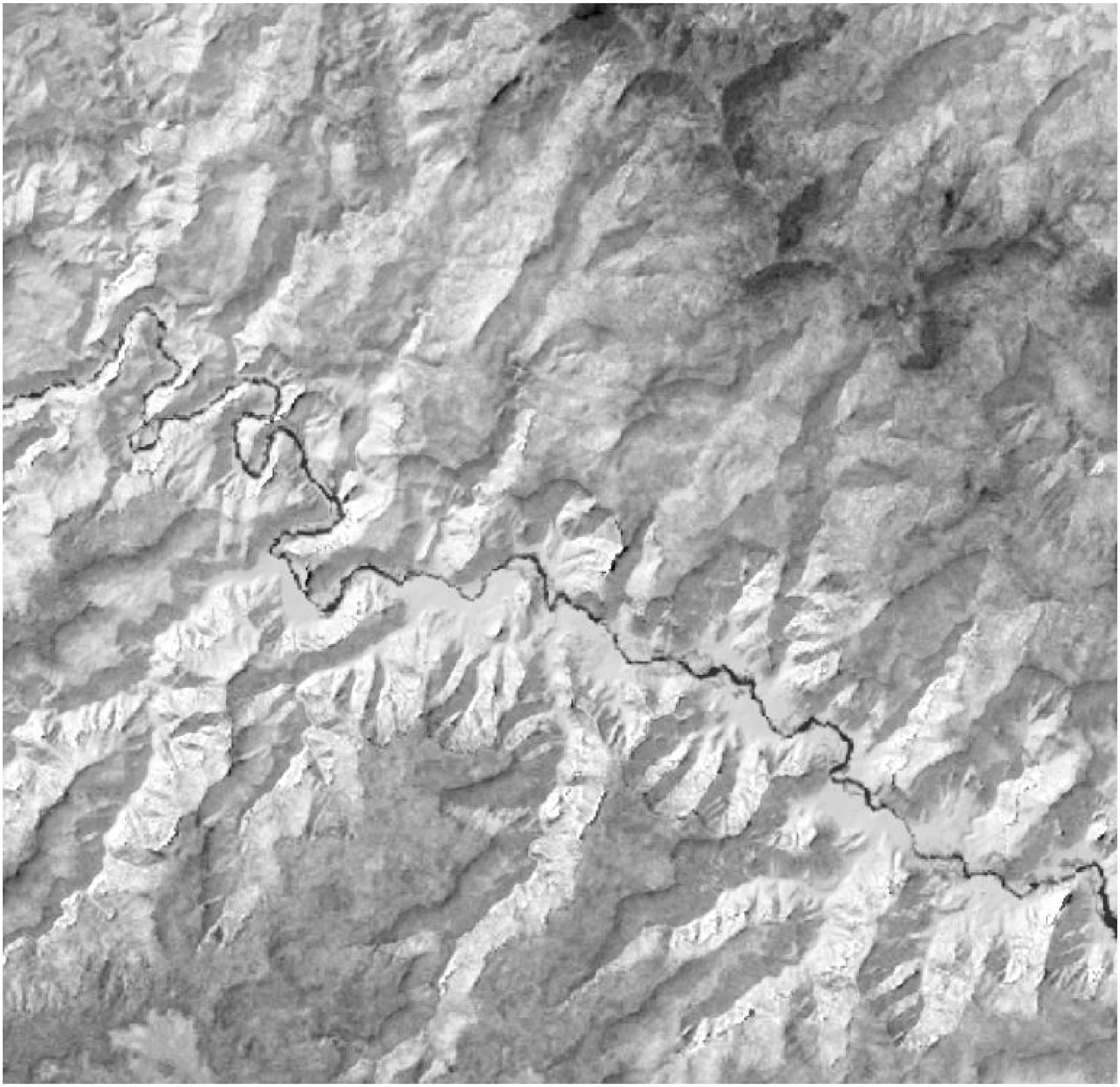
Output: single layer image with data values ranging from -1 to 1; Pixels with higher values can be interpreted as pixels with higher possibility of ferrous minerals influence.

You will be prompted for the input image, and to name the output image.

The model open in ERDAS Imagine Model Maker:



Example normalized ferrous minerals ratio output image over hillshade. Bright pixels have higher data values and can be interpreted as areas of potentially higher ferrous minerals influence:



*Remember ratios must always be calculated on images which have had an atmospheric correction or image standardization applied. Resources for image standardization: http://earth.gis.usu.edu/imagestd/

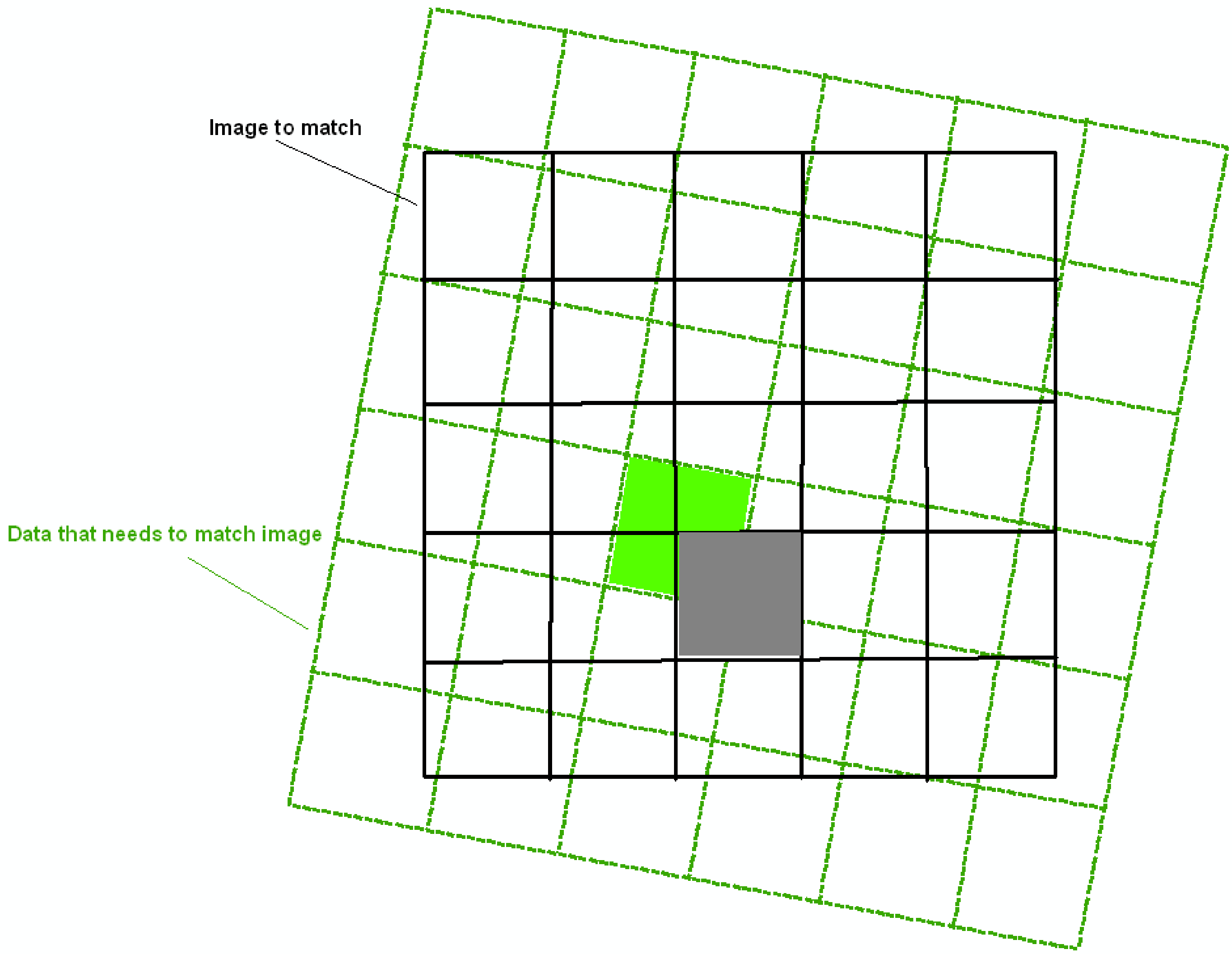
**These are .gmd ERDAS Imagine Model Maker files. They can easily be opened in ERDAS 2013 and all previous versions of Imagine. Starting with ERDAS 2014, they must be converted to the new Spatial Modeler interface and file type.

While these ratios were developed for Landsat bands, they could be modified and used with data from any sensor with bands capturing the same part of the electromagnetic spectrum as the indicated Landsat bands.

Enhanced Thematic Mapper Plus (ETM+)	**Landsat 7**	**Wavelength (micrometers)**	**Resolution (meters)**
	Band 1	0.45-0.52	30
	Band 2	0.52-0.60	30
	Band 3	0.63-0.69	30
	Band 4	0.77-0.90	30
	Band 5	1.55-1.75	30
	Band 6	10.40-12.50	60 * (30)
	Band 7	2.09-2.35	30
	Band 8	.52-.90	15

Grid resample purpose:

When one has raster data from disparate sources, grid cells are often not aligned.



For example, the black grid represents a georeferenced TM image, while the green grid represents the surface of an interpolated soil property. These grids should be co-registered to perform operations like correlation, regression, co-kriging or any exploratory data analysis or modeling operation. The grid_resample model performs this operation.

Iron Normalized Difference Ratio

This ratio is the same conceptually as the iron ratio from the Soil Enhancement Ratios, but normalized instead of a simple ratio. Normalized difference Landsat band ratio using band 3 (VIS-red) and band 7 (SWIR). (See Soil Enhancement Ratio description)

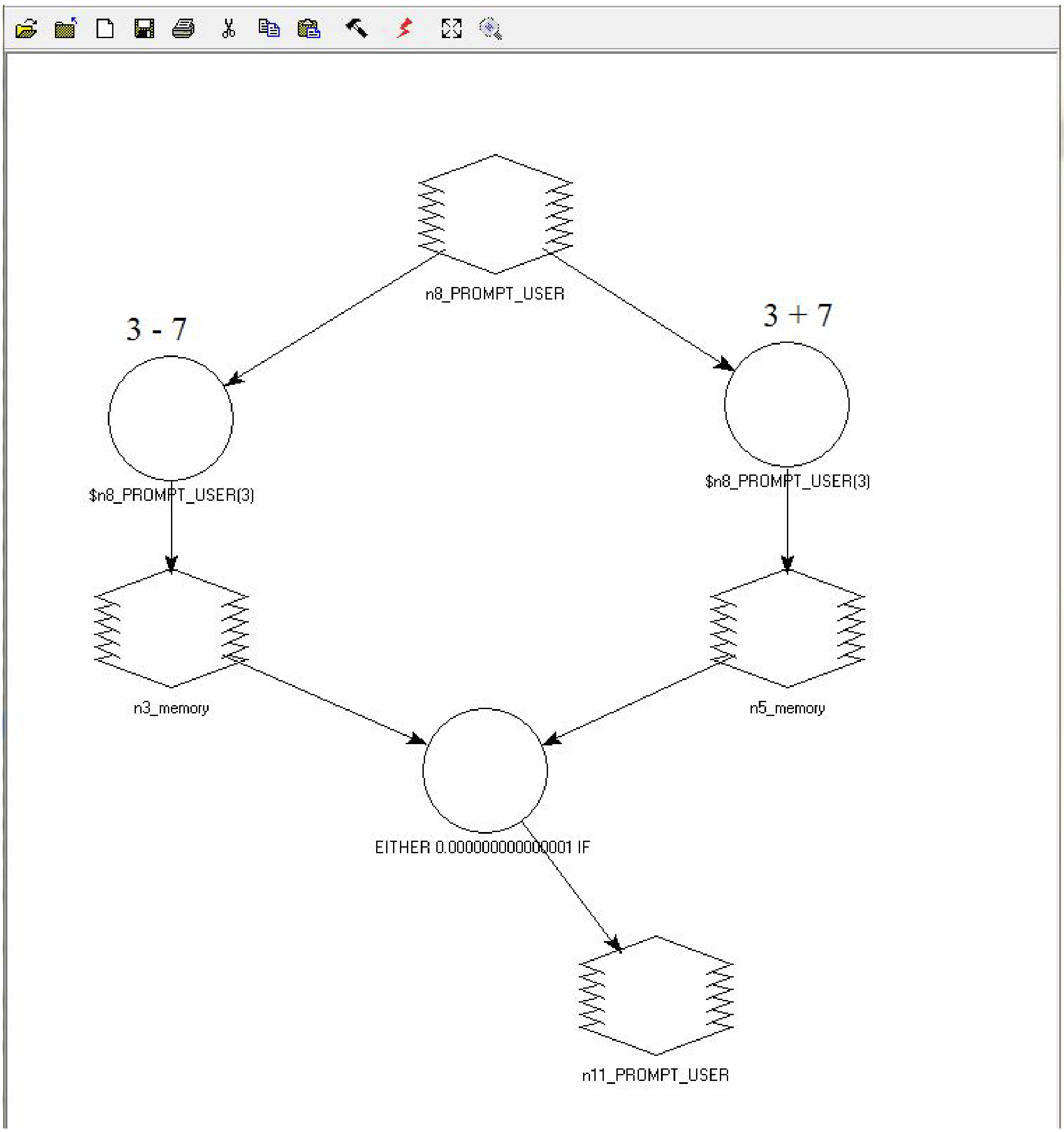
File: iron_index_3_7.gmd (ERDAS Imagine Model, v. 2013 and prior)**

Input: Landsat image containing at least bands 3,7*

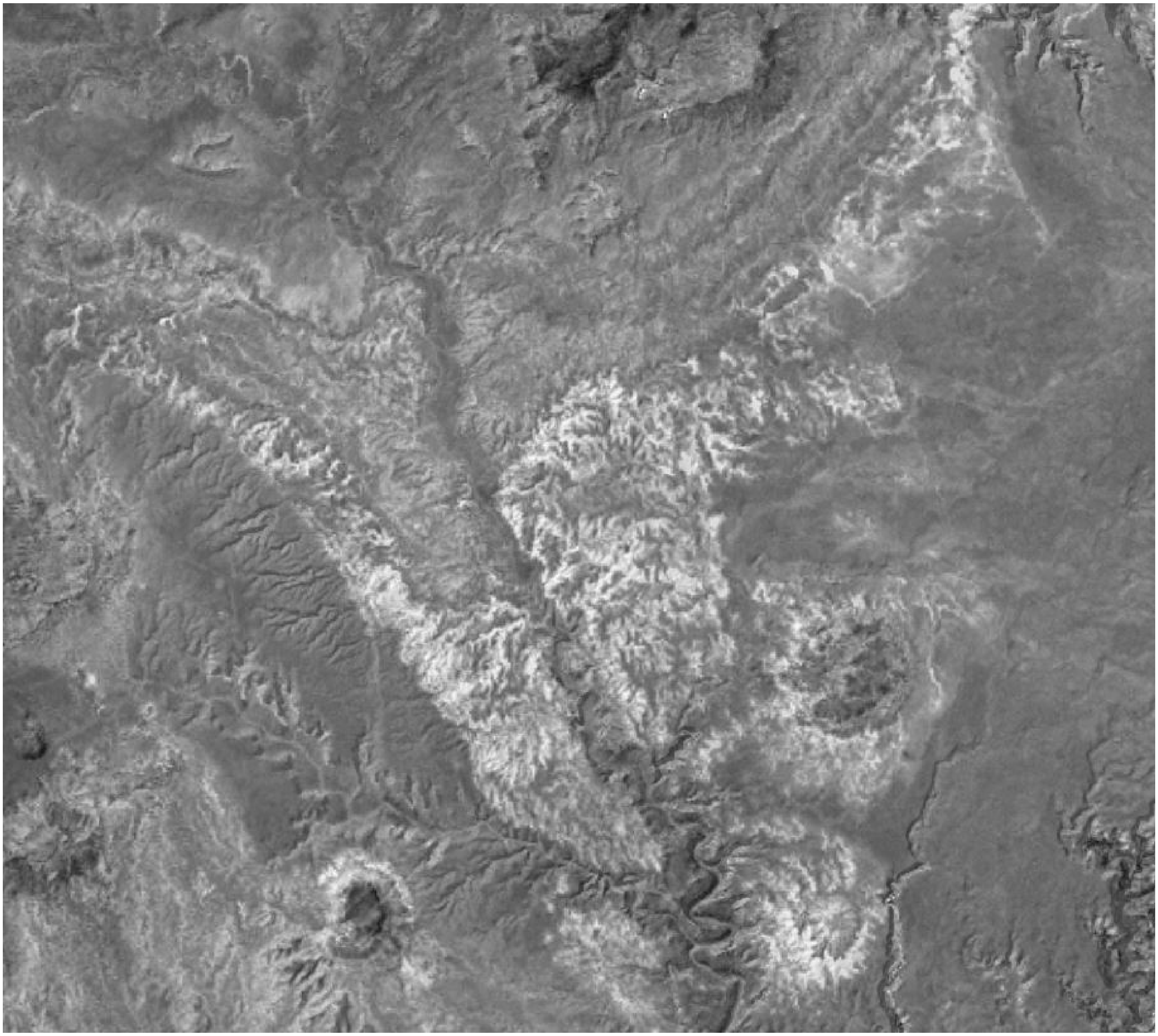
Output: single layer image with data values ranging from -1 to 1; Pixels with higher values can be interpreted as pixels with higher possibility of iron influence.

You will be prompted for the input image, and to name the output image.

The model open in ERDAS Imagine Model Maker:



Example normalized iron ratio output image over hillshade. Bright pixels have higher data values and can be interpreted as areas of potentially higher iron influence:



*Remember ratios must always be calculated on images which have had an atmospheric correction or image standardization applied. Resources for image standardization: http://earth.gis.usu.edu/imagestd/

**These are .gmd ERDAS Imagine Model Maker files. They can easily be opened in ERDAS 2013 and all previous versions of Imagine. Starting with ERDAS 2014, they must be converted to the new Spatial Modeler interface and file type.

While these ratios were developed for Landsat bands, they could be modified and used with data from any sensor with bands capturing the same part of the electromagnetic spectrum as the indicated Landsat bands.

Enhanced Thematic Mapper Plus (ETM+)	**Landsat 7**	**Wavelength (micrometers)**	**Resolution (meters)**
	Band 1	0.45-0.52	30
	Band 2	0.52-0.60	30
	Band 3	0.63-0.69	30
	Band 4	0.77-0.90	30
	Band 5	1.55-1.75	30
	Band 6	10.40-12.50	60 * (30)
	Band 7	2.09-2.35	30
	Band 8	.52-.90	15

Optimum Index Factor

The optimum index factor (OIF) is a method for determining the three-band combination that maximizes the variability in a particular scene and was developed by Chavez et al (1982; 1984). The optimum three-band combination is useful for visualization purposes, but may also be useful in analysis. An OIF is calculated for each of the 20 possible three-band combinations that can be created from the six bands of Landsat TM data (excluding the thermal band). Although the OIF method was developed for Landsat TM data, it is applicable to any multispectral remote sensing dataset. Kienast-Brown and Boettinger (2010) found utility in applying OIF to multiple data types; terrain derivatives, as well as remote sensing data.

OIF is based on the amount of total variance and correlation within and between the possible band combinations. The three-band combinations are ranked, and those with the largest OIF values contain the most information (measured by variance) with the least amount of duplication (measured by correlation) and are the best choices for maximizing the variability in a particular scene. For viewing purposes, the analyst must decide what color channel (red, green, or blue) to assign each band of the three-band combination. The following equation is used to calculate OIF for any three-band combination:

$$OIF = \frac{\sum_{k=1}^{3} s_k}{\sum_{j=1}^{3} Abs(r_j)}$$

where s_k is the standard deviation for band k, and r_j is the absolute value of the correlation coefficient between any two of the three bands being evaluated. Standard deviation values are obtained using the Metadata button in Imagine, and the correlation coefficients for a particular scene are derived using Model Maker in Imagine. An Excel spreadsheet programmed with the equation above is provided and can be used to calculate and rank OIF values.

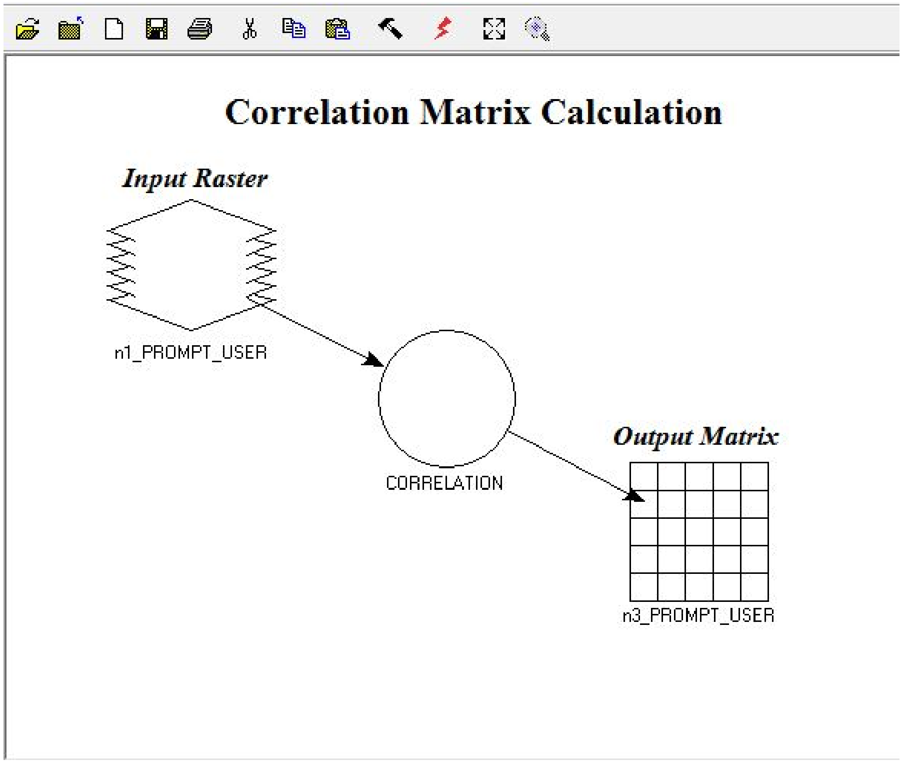
File: OIF_correlationmatrix.gmd (ERDAS Imagine Model, v. 2013 and prior)** and OIF_worksheet.xlsx

Input: 6-band Landsat image (or any 6-band image) for correlation matrix model; correlation matrix and standard deviation values for image bands for worksheet

Output: Model outputs the correlation matrix for input image; Worksheet uses correlation matrix and standard deviation values to calculate OIF values for the 20 possible three-band combinations for input image, and rank them largest OIF value to smallest.

You will be prompted for the input image, and to name the output matrix when running the model.

The model open in ERDAS Imagine Model Maker:



OIF worksheet for example dataset. The absolute value of the correlation coefficients must be used (see equation above). Standard deviation values for each band of the input image are obtained using the Metadata button in Imagine. After OIF values are calculated, you must sort descending to get the ranked band combinations and associated OIF values:

OIF- Worksheet
Example Dataset

Correlation Matrix

TM band	1	2	3	4	5	7
1	1.00	0.99	0.99	0.91	0.80	0.83
2	0.99	1.00	0.99	0.89	0.74	0.79
3	0.99	0.99	1.00	0.93	0.79	0.83
4	0.91	0.89	0.93	1.00	0.88	0.89
5	0.80	0.74	0.79	0.88	1.00	0.99
7	0.83	0.79	0.83	0.89	0.99	1.00

Standard Deviation

TM band	1	2	3	4	5	7
	35.566	35.28	44.152	35.089	50.753	45.168

Band Combo	OIF	Ranked Band Combo	OIF
1,2,3	38.73	3,5,7	53.60
1,2,4	38.00	2,5,7	52.09
1,2,5	48.03	2,3,5	51.46
1,2,7	44.53	1,3,5	50.64
1,3,4	40.67	1,5,7	50.30
1,3,5	50.64	3,4,5	49.86
1,3,7	47.20	2,4,5	48.10
1,4,5	46.93	1,2,5	48.03
1,4,7	44.10	2,3,7	47.76
1,5,7	50.30	4,5,7	47.41
2,3,4	40.73	1,3,7	47.20
2,3,5	51.46	1,4,5	46.93
2,3,7	47.76	3,4,7	46.91
2,4,5	48.10	2,4,7	45.01
2,4,7	45.01	1,2,7	44.53
2,5,7	52.09	1,4,7	44.10
3,4,5	49.86	2,3,4	40.73
3,4,7	46.91	1,3,4	40.67
3,5,7	53.60	1,2,3	38.73
4,5,7	47.41	1,2,4	38.00

Sheet1 Sheet2 Sheet3

References: Chavez, P. S., G.L. Berlin, and L. B. Sowers. 1982. Statistical method for selecting Landsat MSS ratios. Journal of Applied Photographic Engineering. 8(1): 23-30.

Chavez, P.S., S.C. Guptill, and J. A. Bowell. 1984. Image processing techniques for Thematic Mapper data. Proceedings, ASPRS Technical Papers. 2: 728-742.

Kienast-Brown, S. and Boettinger, J.L. 2010. Applying the Optimum Index Factor to multiple data types in soil survey. *In*Boettinger, L.L., Howell, D.W., Moore, A.C., Hartemink, A.E., and Kienast-Brown, S. (eds.) Digital Soil Mapping: Bridging Research, Environmental Application, and Operation. Progress in Soil Science. Vol. 2. Springer, Heidelberg.

**These are .gmd ERDAS Imagine Model Maker files. They can easily be opened in ERDAS 2013 and all previous versions of Imagine. Starting with ERDAS 2014, they must be converted to the new Spatial Modeler interface and file type.

Rock Outcrop Normalized Difference Ratio

This ratio is a normalized difference ratio of Landsat 7 ETM+ bands 5 (SWIR) and 2 (VIS-green) to identify limestone outcroppings. Limestone and dolomite have greater reflectance in band 5 relative to band 2, while andesite and other igneous materials have greater reflectance in band 2 relative to band 5. This ratio can essentially be used to distinguish sedimentary from igneous parent material.

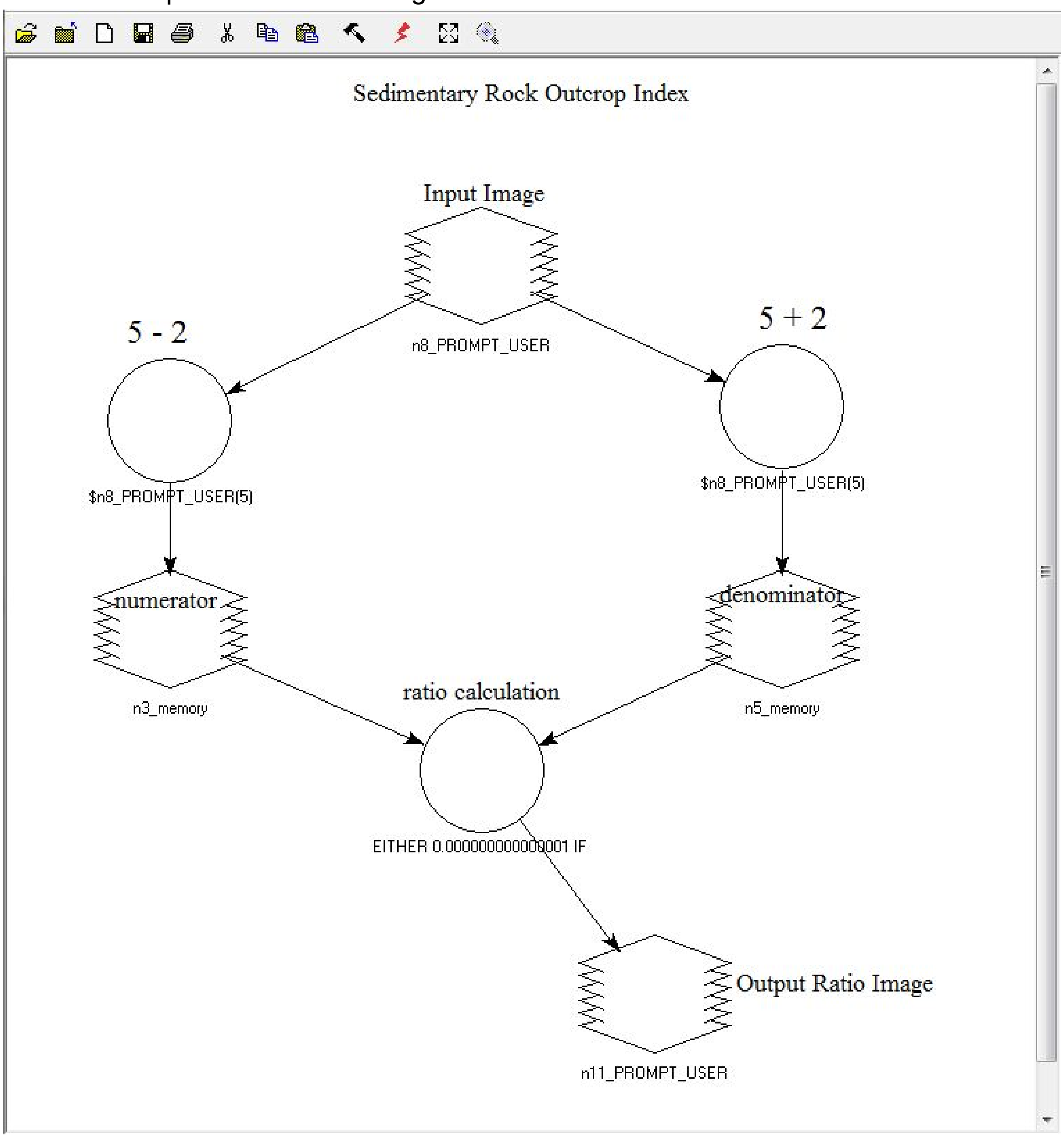
File: rock_outcrop_5_2.gmd (ERDAS Imagine Model, v. 2013 and prior)**

Input: Landsat image containing at least bands 2,5*

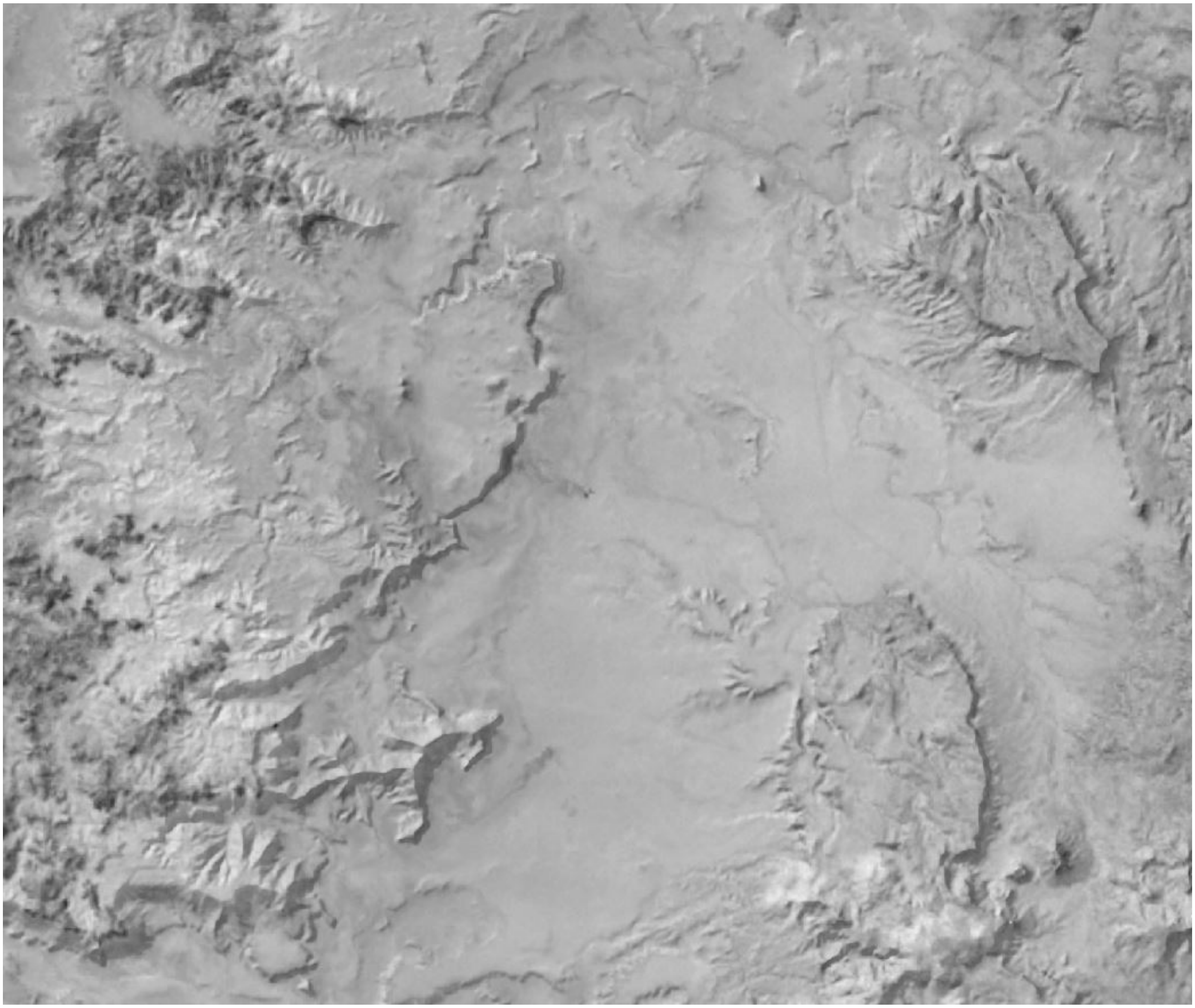
Output: single layer image with data values ranging from -1 to 1; Pixels with higher values can be interpreted as pixels with higher possibility of limestone or dolomite influence and pixels with lower values can be interpreted as pixels with higher possibility of igneous material influence.

You will be prompted for the input image, and to name the output image.

The model open in ERDAS Imagine Model Maker:



Example normalized rock outcrop ratio output image over hillshade. Bright pixels with higher pixel values can be interpreted as areas with a higher possibility of limestone or dolomite influence and dark pixels with lower values can be interpreted as areas with higher possibility of igneous material influence:



Reference: Bodily, J.M., 2005. Developing a digital soil survey update protocol at the Golden Spike National Historic site. M.S. Thesis, Utah State Univeristy, Logan.

*Remember ratios must always be calculated on images which have had an atmospheric correction or image standardization applied. Resources for image standardization: http://earth.gis.usu.edu/imagestd/

**These are .gmd ERDAS Imagine Model Maker files. They can easily be opened in ERDAS 2013 and all previous versions of Imagine. Starting with ERDAS 2014, they must be converted to the new Spatial Modeler interface and file type.

While these ratios were developed for Landsat bands, they could be modified and used with data from any sensor with bands capturing the same part of the electromagnetic spectrum as the indicated Landsat bands.

Enhanced Thematic Mapper Plus (ETM+)	**Landsat 7**	**Wavelength (micrometers)**	**Resolution (meters)**
	Band 1	0.45-0.52	30
	Band 2	0.52-0.60	30
	Band 3	0.63-0.69	30
	Band 4	0.77-0.90	30
	Band 5	1.55-1.75	30
	Band 6	10.40-12.50	60 * (30)
	Band 7	2.09-2.35	30
	Band 8	.52-.90	15

Soil Enhancement Ratios

The BLM developed the Soil Enhancement Ratios, which are three simple Landsat band ratios for carbonates, iron, and clay. Carbonates can be indicated using band 3 (VIS-red) and band 2 (VIS-green); iron using band 3 (VIS-red) and band 7 (SWIR); and clay (hydroxyls) using band 5 (shorter wavelength SWIR) and band 7 (longer wavelength SWIR). In this model, the three band ratios are calculated and stacked into a three-layer output image. Although, the exact physical relationships of these ratios to surface materials may not apply for all study areas, this set of band ratios can represent and help quantify mineralogical diversity of parent material across the landscape.

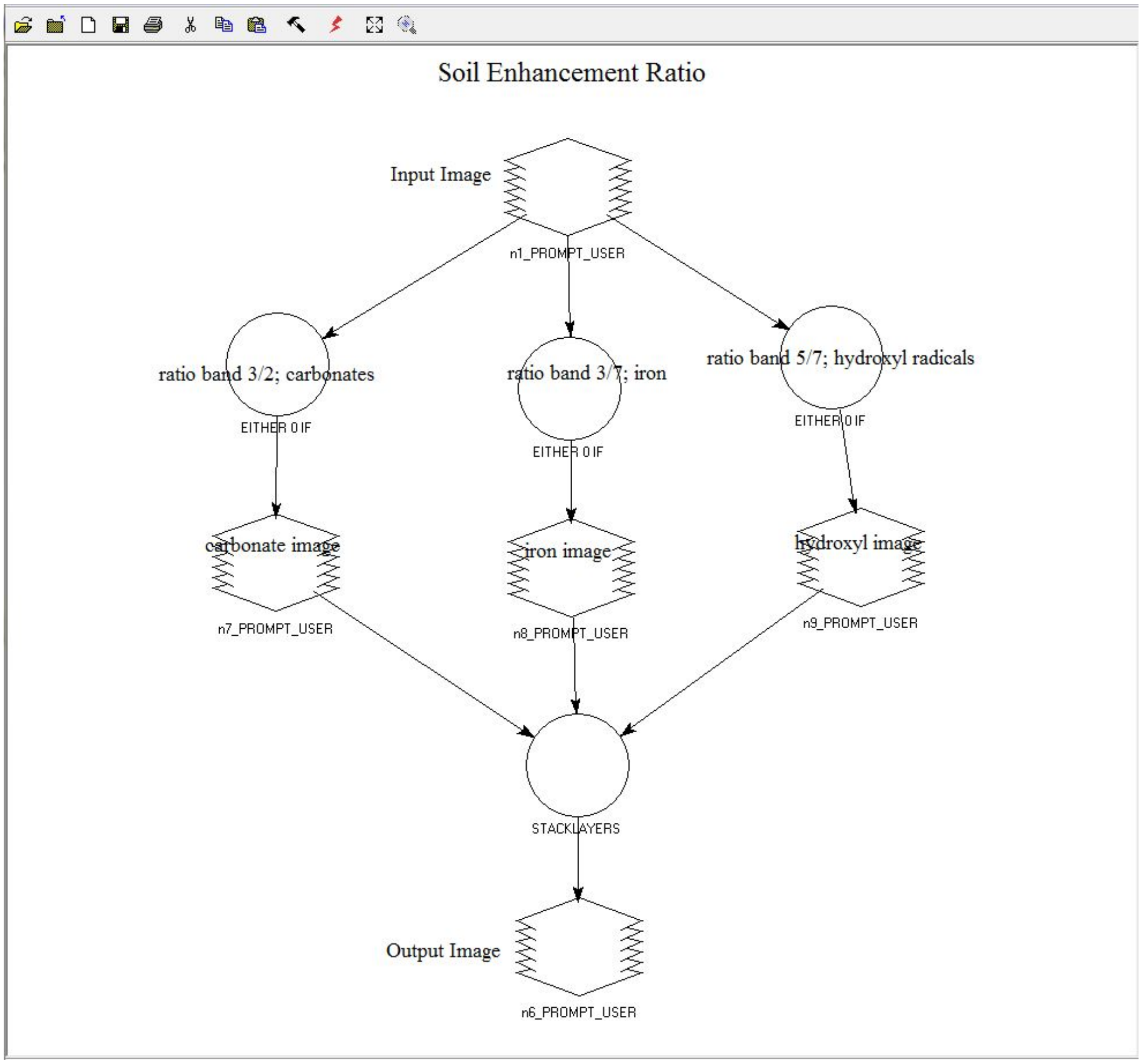
File: soilenhancement.gmd (ERDAS Imagine Model, v. 2013 and prior)**

Input: Landsat image containing at least bands 2,3,5,7*

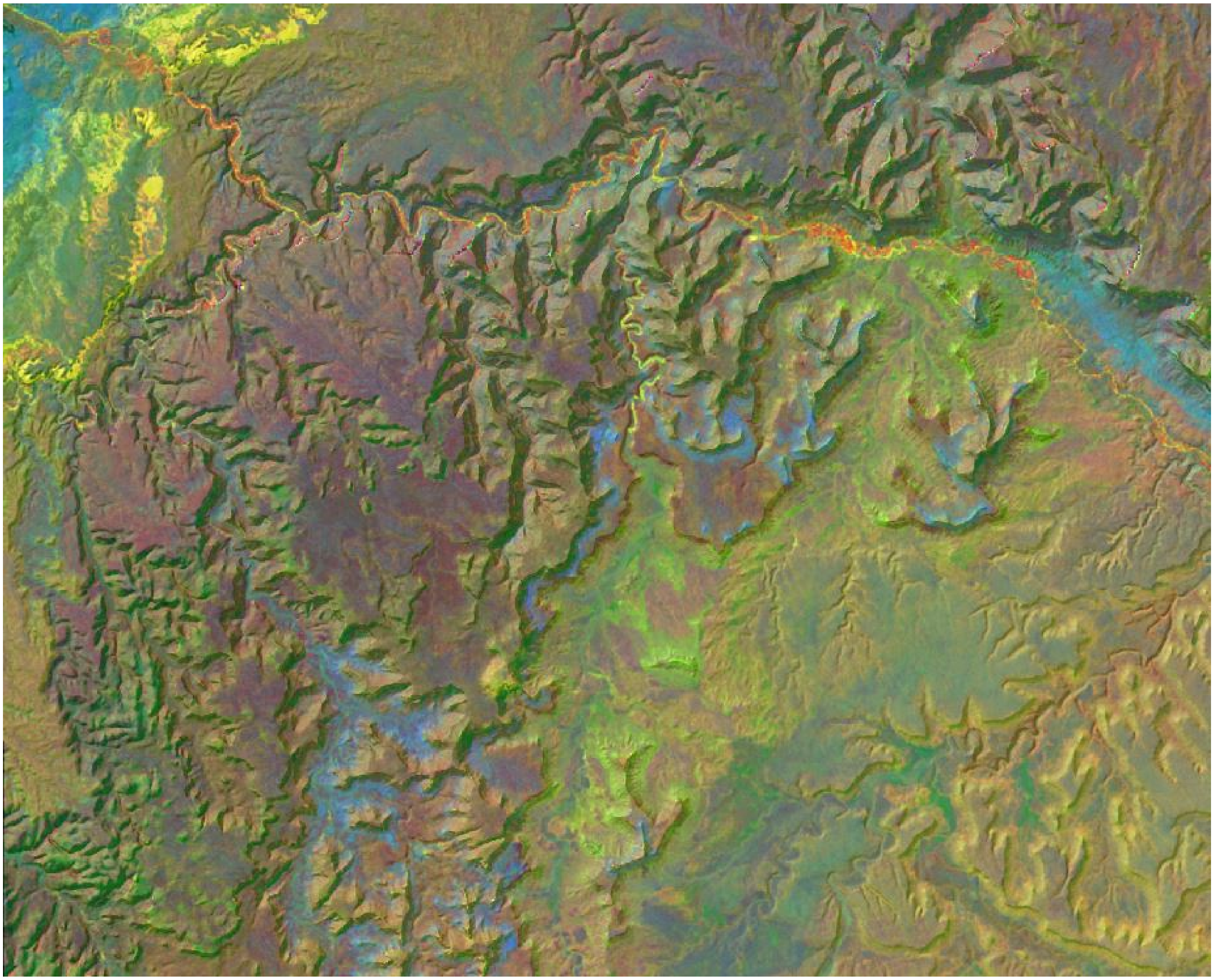
Output: carbonate image, iron image, clay image (single layer images from simple ratios); three-layer image containing clay ratio (layer 1), iron ratio (layer 2), carbonate ratio (layer 3)

You will be prompted for the input image, and to name the output carbonate, iron, and clay (hydroxyl) images, and the three-layer output image.

The model open in ERDAS Imagine Model Maker:



Example three-layer output image over hillshade:



*Remember ratios must always be calculated on images which have had an atmospheric correction or image standardization applied. Resources for image standardization: http://earth.gis.usu.edu/imagestd/

**These are .gmd ERDAS Imagine Model Maker files. They can easily be opened in ERDAS 2013 and all previous versions of Imagine. Starting with ERDAS 2014, they must be converted to the new Spatial Modeler interface and file type.

While these ratios were developed for Landsat bands, they could be modified and used with data from any sensor with bands capturing the same part of the electromagnetic spectrum as the indicated Landsat bands.

Enhanced Thematic Mapper Plus (ETM+)	**Landsat 7**	**Wavelength (micrometers)**	**Resolution (meters)**
	Band 1	0.45-0.52	30
	Band 2	0.52-0.60	30
	Band 3	0.63-0.69	30
	Band 4	0.77-0.90	30
	Band 5	1.55-1.75	30
	Band 6	10.40-12.50	60 * (30)
	Band 7	2.09-2.35	30
	Band 8	.52-.90	15

Processing DEMs with the ArcSIE extension

Background

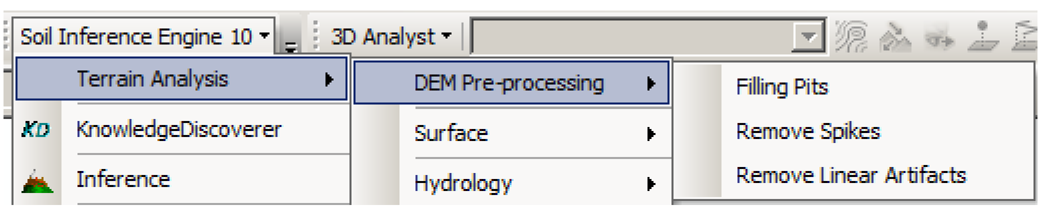
This job aid is to inform users of the availability of the ArcSIE Extension and provide a brief overview of unique features and a link to the User Guide. The ArcSIE software is free and CCE certified for USDA use.

ArcGIS is the GIS software of choice for USDA-NRCS. There has been a licensing agreement in place since 2002 that covers much of the ESRI product line for many USDA agencies, including NRCS. Use of the software for conservation, engineering, and soil survey applications is well established across the Agency. ArcGIS has limitations regarding the development of terrain derivatives. Some of these limitations have been overcome by the ArcSIE software extension that operates within the familiar ArcGIS framework.

ArcSIE is primarily used as a knowledge-based reasoning tool for soil mapping activities. It also provides the capability for processing DEMs and development of terrain derivatives using parameters unavailable with standard ArcGIS software. The ability to specify size and shape of a neighborhood, independent of raster resolution, is a significant enhancement available with the ArcSIE extension. Please refer to the ArcSIE User Guide (AUG) for more detailed information (available at: http://www.arcsie.com/Download.htm).

Preprocessing DEM Tools

Note: Refer to Chapter 3, Terrain Analysis, in AUG.

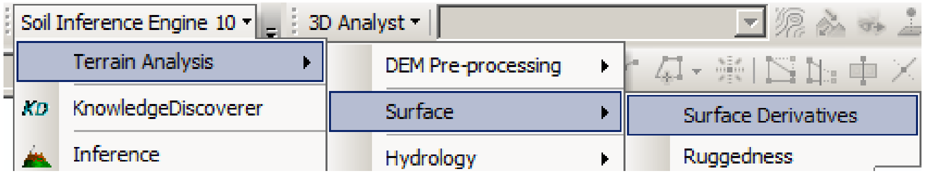


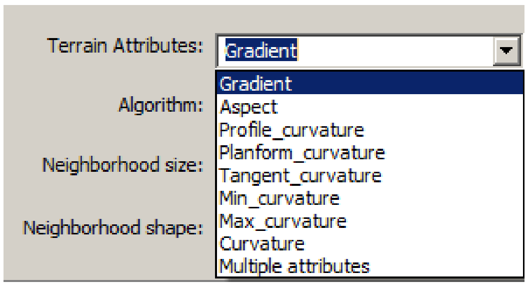
Filling Pits. This tool is analogous to the Fill command in ArcGIS.

Remove Spikes. This tool provides the option of removing artifacts or vegetative spikes automatically or with a specified shapefile of spike locations. The user may also confine the removal to areas below a specified elevation. There is no directly available tool in ArcGIS for this operation.

Remove Linear Artifacts. This tool is designed to remove the linear “bumps” of roads from elevation data. A shapefile of the linear features is required as an input. There is no directly available tool in ArcGIS for this operation.

Developing Terrain Derivatives





Processes Available in ArcSIE and Not in ArcGIS

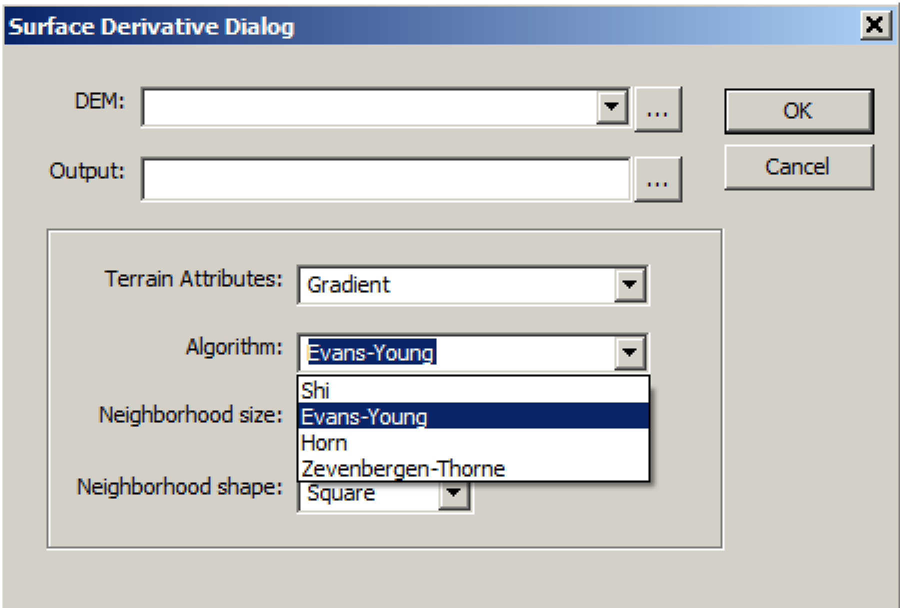
- Tangent Curvature
- Minimum Curvature
- Maximum Curvature
- Multiple attributes (batch process selected set of available derivatives)

Available Parameters

Note: There is a job aid on resolution and neighborhoods that provides additional explanations for these parameters.

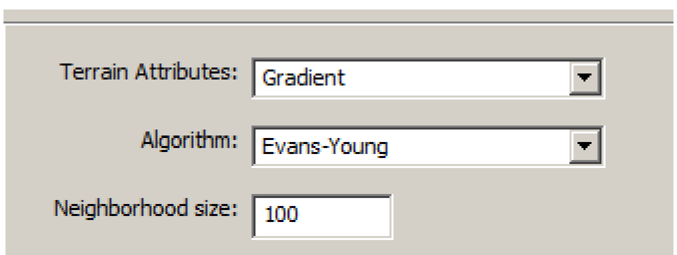
Interpolation Algorithms. Four are provided (some are grouped in one):

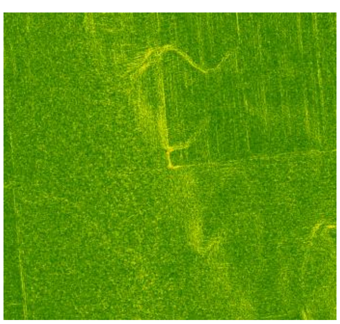
- Preferred method for noisy DEMs
- Evans-Young or Horn
- Preferred method for smooth DEMs
- Shi or Zevenbergen-Thorne



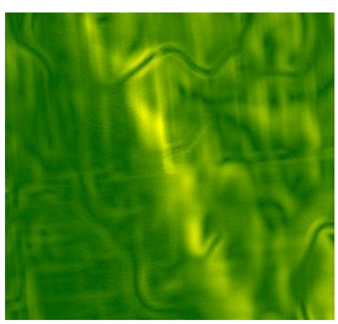
ArcGIS uses the Evans-Young algorithm exclusively.

Neighborhood Size. Any size may be specified for the calculation neighborhood. The neighborhood size is not dependent on the cell resolution of the DEM. Neighborhood size is a critical parameter to control when using high-resolution data. ArcGIS restricts its neighborhood to the traditional 3 x 3 window, resulting in noisy output when using high-resolution data.



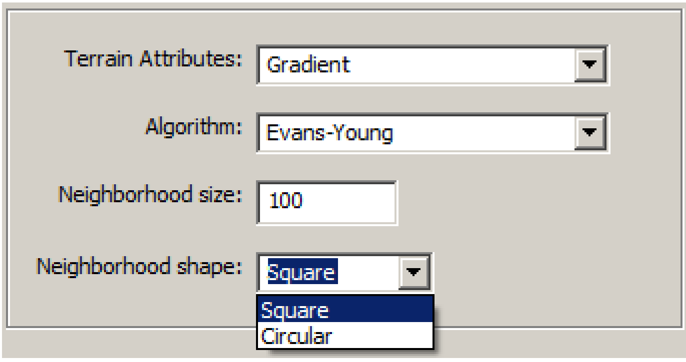


Above: Slope from ArcGIS.



Above: Slope from ArcSIE with 100' neighborhood.

Neighborhood Shape. Both square and circular neighborhoods are supported.



Research suggests that circular neighborhoods are more accurate than square neighborhoods (Shi et al., 2007; Shi et al., 2012; Smith et al., 2006).

ArcGIS is confined to a 3 x 3 window.

Additional Derivatives

Ruggedness. This tool implements the topographic ruggedness index (TRI) developed by Riley et al. (1999). TRI is the rooted mean squared difference between the elevation of a cell and the elevations of the cells in its neighborhood. The same routine is available in SAGA and the xTerrain Toolbox.

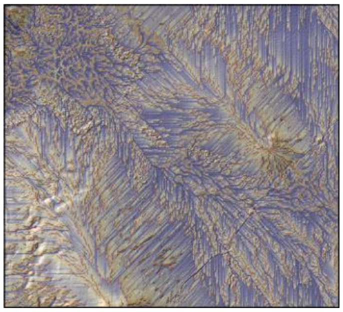
Flow Accumulation. This tool determines the upslope contributing area for each cell and assigns this area to the value of the cell. The unit assigned is the same as the horizontal units of the layer, e.g., square feet or square meters. This differs from ArcGIS, which assigns a cell count to the value of the cell.

Both uni-path and multipath flow paths are supported. The uni-path algorithm assumes water flows only to one of the cells in the eight cardinal directions. The multipath algorithm routes water to all neighboring cells of a lower elevation, using the Quinn et al. (1991) algorithm. The algorithm in ArGIS is confined to uni-path.

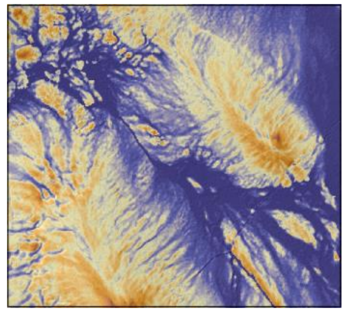
Wetness Index. This is also known as Compound Topographic Index and Topographic Wetness Index and is calculated as:

w = ln(Flow Accumulation/Slope Gradient)

The uni-path, multipath, and multipath smoothed flow path algorithms are supported. The multipath algorithm overcomes the unnatural patterns typical with the uni-path output. The multipath smoothed algorithm helps to overcome the “dry valley” problem common in broad, flat flood plains.



Above: Wetness index implemented with ArcGIS Unipath.



Above: Multipath wetness index with ArcSIE.

Ridgeline. This tool identifies cells that occupy ridgetops.

Broad – Narrow Ridgeline. This tool differentiates broad and narrow ridges.

Topographic Classification. This tool implements the method developed by Iwahashi and Pike (2007) for characterizing terrain. The method uses slope gradient, local convexity, and surface texture with a set of rules to dissect the mapping area into a user-specified number of classes (8, 12, or 16).

Zimmerman's Relative Position. This tool is based on an algorithm by Zimmermann (1999) and generates continuous measurements of relative slope positions, from ridgetop or peak to valley bottom. It runs at various spatial scales and then hierarchically integrates the results at different scales into a single layer.

Streamlines. This tool derives streamlines from a DEM. The user can select one of three methods: (1) O'Callahan and Mark, (2) Skidmore, or (3) Peuker and Douglas. The preferred method is O'Callahan and Mark. An option to order streams according to the following methods—Simple, Shreve, Strahler, Unique—is also available.

References

Iwahashi, J., and R.J. Pike. 2007. Automated classifications of topography from DEMs by an unsupervised nested-means algorithm and a three-part geometric signature. Geomorphology 86:409–440.

Quinn, P.F., K.J. Beven, P. Chevallier, and O. Planchon. 1991. The prediction of hillslope flowpaths for distributed modelling using digital terrain models. Hydrological Processes 5:59-80.

Riley, S.J., S.D. DeGloria, and R. Elliott. 1999. A terrain ruggedness index that quantifies topographic heterogeneity. Intermountain Journal of Science 5:23-27.

Shi, X., A-X. Zhu, J. Burt, W. Choi, R-X. Wang, T. Pei, and B-L. Li. 2007. An experiment with circular neighborhood in the calculation of slope gradient from DEM. Photogrammetric Engineering and Remote Sensing 73(2):143-154.

Shi, X., L. Girod, R. Long, R. DeKett, J. Philippe, and T. Burke. 2012. A comparison of LiDAR-based DEMs and USGS-sourced DEMs in terrain analysis for knowledge-based digital soil mapping. Geoderma 170:217-226.

Smith, M. P., A. Zhu, J.E. Burt, and C. Stiles. 2006. The effects of DEM resolution and neighborhood size on digital soil survey. Geoderma 137(1):58-69.

Zimmerman, N. 1999. http://www.wsl.ch/staff/niklaus.zimmermann/programs/aml4_1.html

Using SAGA to Develop Terrain Derivatives

What is SAGA?

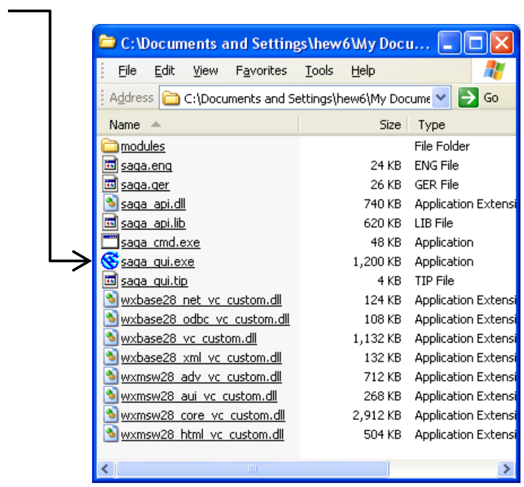
The System for Automated Geoscientific Analysis, or SAGA, was developed at the University of Hamburg, Delaware, and released through the GNU public license. This system is an open-source, free, and CCE-certified software. It has more built-in functionality for terrain analysis than ArcGIS, and many terrain derivatives are only available in SAGA. The system has capabilities for image and data processing as well as analysis. More information on SAGA and the SAGA user's group is available at: http://www.saga-gis.org. Useful guides are available here: Volume 1 and Volume 2.

Obtaining SAGA

For a conventionally installed version of the latest CCE Certified version, contact your ITS specialist.

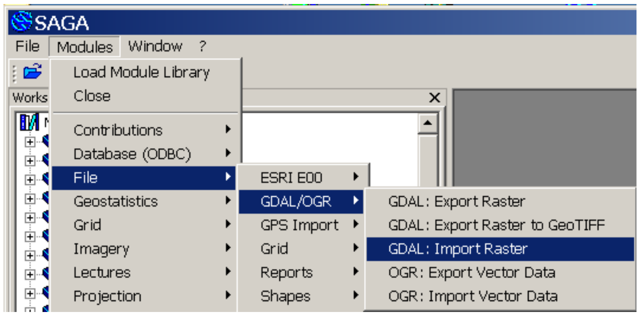
Opening SAGA

- Double-click on the shortcut saga or
- Navigate to the install directory (e.g., "C:\Program Files\saga_2.1.0_x64") and double-click on the executable program "saga_gui.exe."

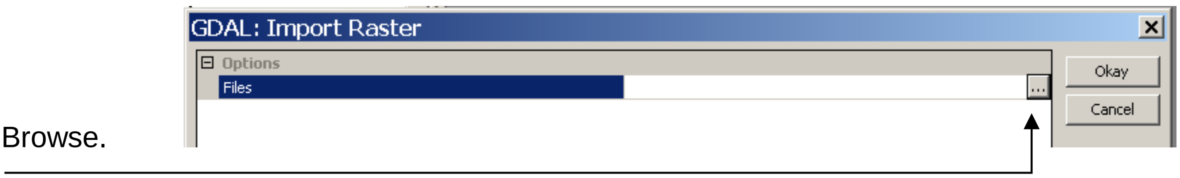


Importing Data

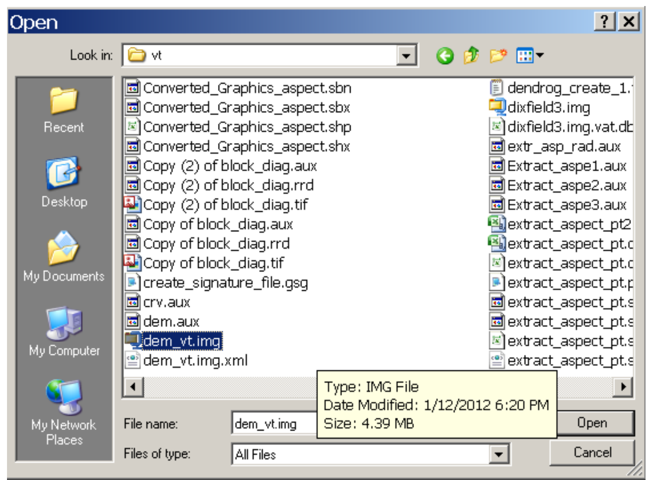
- Select a file to import with the GDAL module.



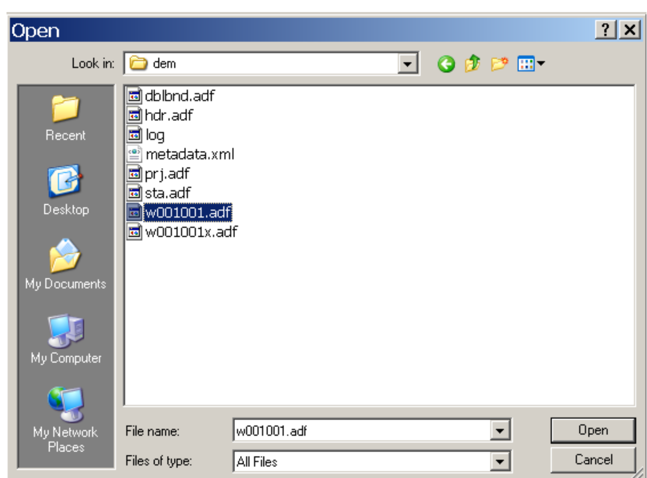
- Browse.



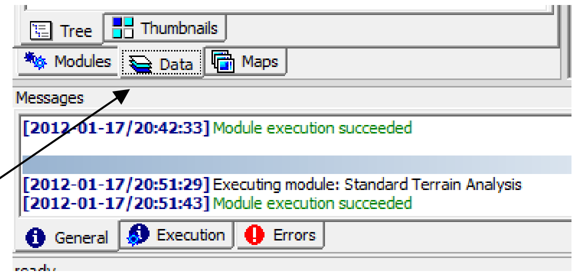
- Example of an Imagine file:



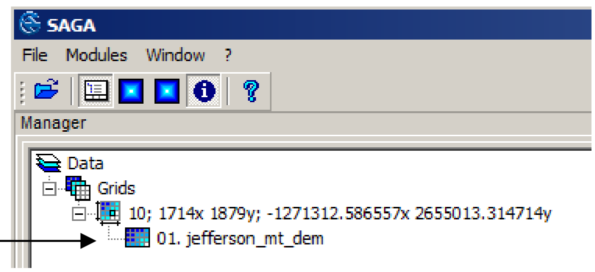
- Example of an ESRI GRID file:

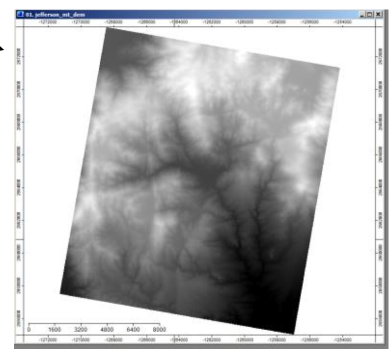


- To view data, click on the Data tab.



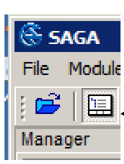
- Double-click on the layer to open a map.

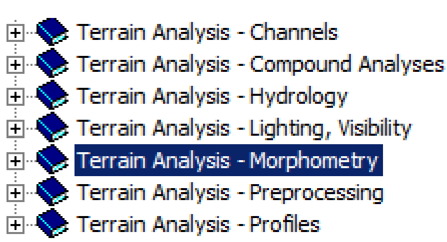




Terrain Analysis

The various modules are available from the Manager Window.





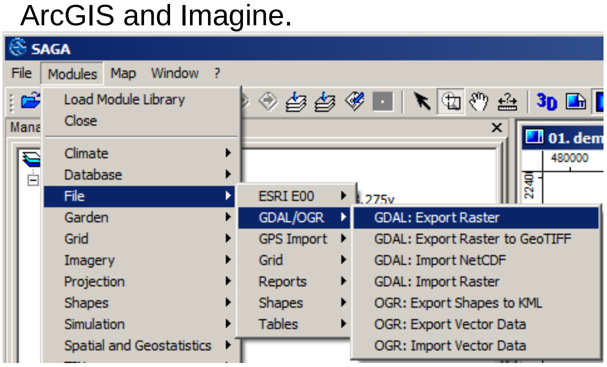
The following make up a selected list of derivatives used in digital soil mapping work:

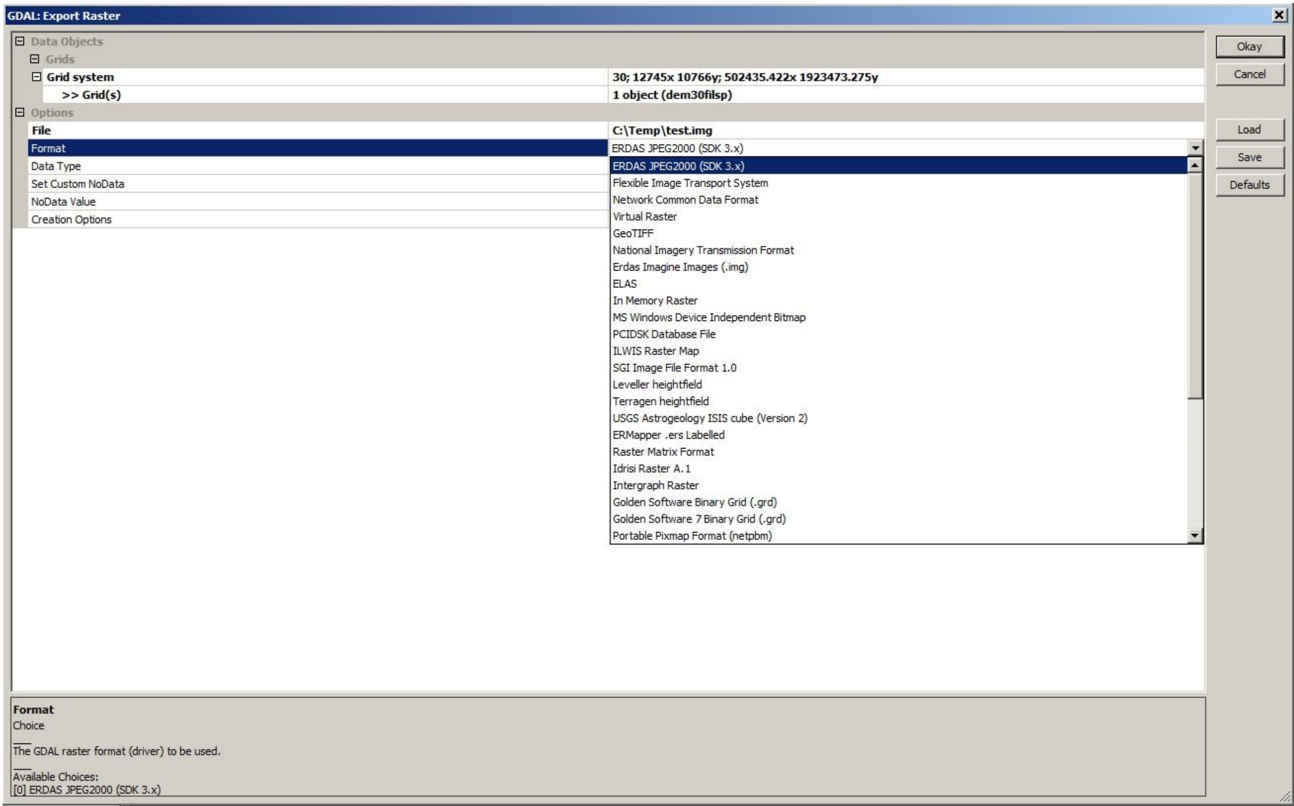
- Channel network base level (Terrain Analysis – Channels)
- Convergence index (Terrain Analysis – Morphometry)
- Downslope Distance Gradient (Terrain Analysis – Morphometry)
- Fuzzy Landform Elements from Schmidt & Hewitt (Terrain Analysis – Morphometry)
- Incoming solar radiation (Terrain Analysis – Lighting, Visibility)
- Morphometric Protection Index (Terrain Analysis – Morphometry)
- Multiresolution Index of Valley Bottom Flatness (Terrain Analysis – Morphometry)
- Overland flow distance to channel (Terrain Analysis – Channels)
- SAGA Wetness Index (Terrain Analysis – Hydrology)
- Slope length (Terrain Analysis – Hydrology)
- Standardized slope height (Terrain Analysis – Morphometry)
- Stream power index (Terrain Analysis – Hydrology)
- Terrain Surface Classification from Iwahashi & Pike (Terrain Analysis – Morphometry)
- Vertical distance to channel (Terrain Analysis – Channels)

The job aid “Multiresolution Index of Valley Bottom Flatness” is useful for mapping ridgetops and valleys. It is available at:
https://sharepoint.gru.wvu.edu/sites/digital_soils/GIS%20Guides/23_MRVBF_sheet.pdf.

Save the Output

- Right-click on any map grid output you want to keep for future use in SAGA and click on“Save As.”
- The output will be saved in .sgrd format, the native binary format for SAGA.
- Use the GDAL Export Raster to save in formats for use with other software, such as ArcGIS and Imagine.





Tip for SAGA

SAGA stores the data and map layouts from session to session. If you would like to start SAGA with a “clean slate,” navigate to the location of SAGA and delete the file named “saga_gui.ini.”

A Quick Overview of the xTerrain_tools Toolbox for ArcGIS 10.1

By Jon Bathgate, Jon Bonjean, and Tom D'Avello

Introduction

This job aid is designed to show how to open the ArcToolbox window, add the Toolbox, and open the help panel for more information. Users need ArcGIS 10.1 and the xTerrain_tools_10.1_0.4.tbx toolbox.

Topography was proposed as one of the five soil-forming factors by Jenny (1941). Numerous papers cite the use of various terrain derivatives in support of soil mapping investigations. This suite of scripts was assembled to provide users an efficient means of developing some of the more commonly used derivatives. In addition, several scripts perform data pre-processing operations such as smoothing or data normalization.

Adding the Toolbox to ArcGIS

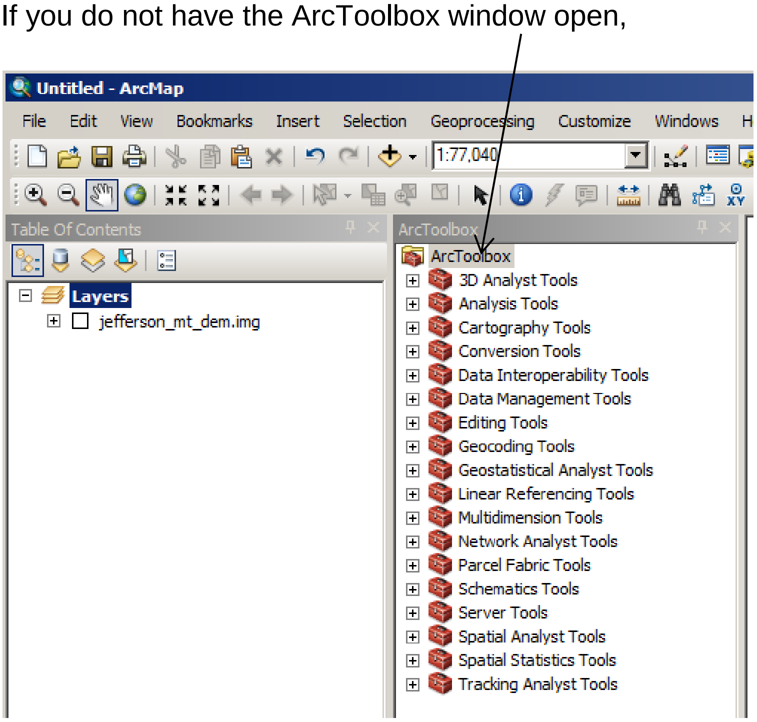
Objective

Show how to add the toolbox to ArcGIS.

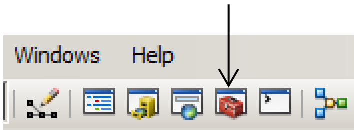
Procedure

Copy xTerrain_tools_10.1_0.4.tbx to a folder on your computer, for example, C:\workspace\scripts.

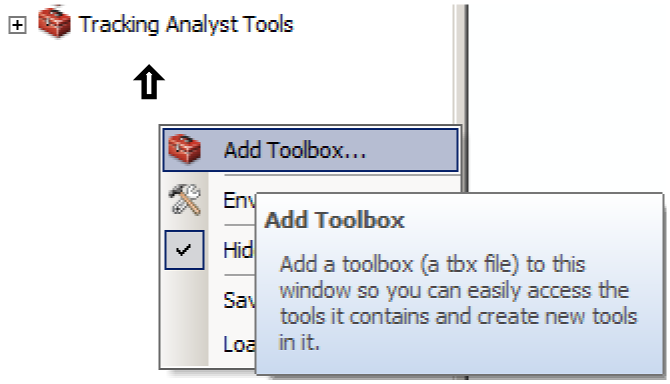
If you do not have the ArcToolbox window open,



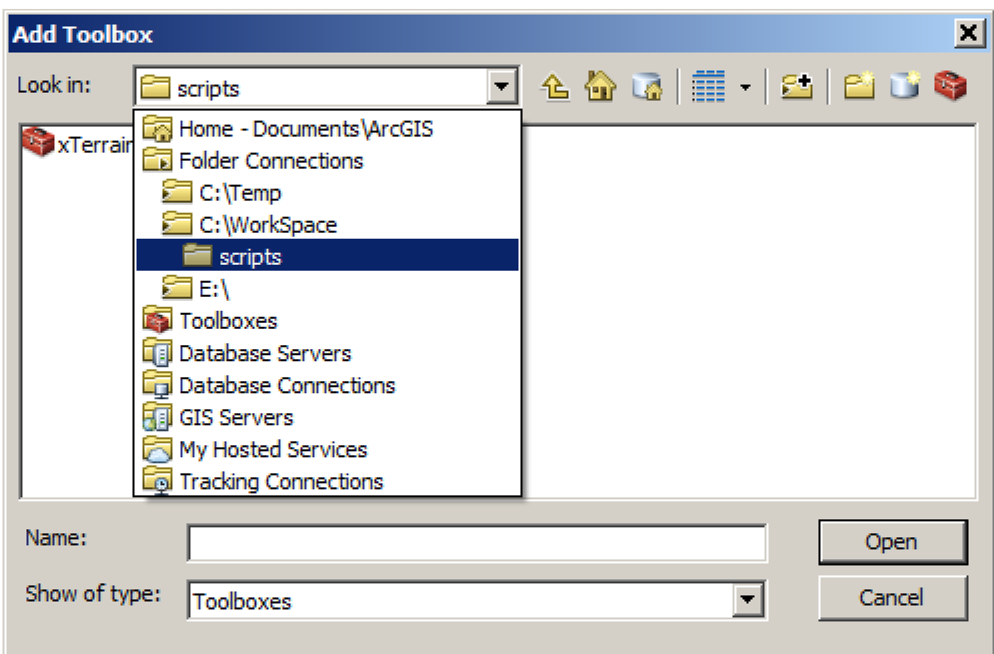
Click on the ArcToolbox button to add it to your GUI.



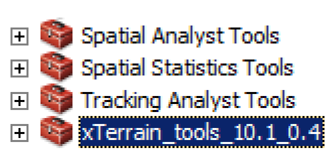
Right-click in the ArcToolbox window to add a toolbox.



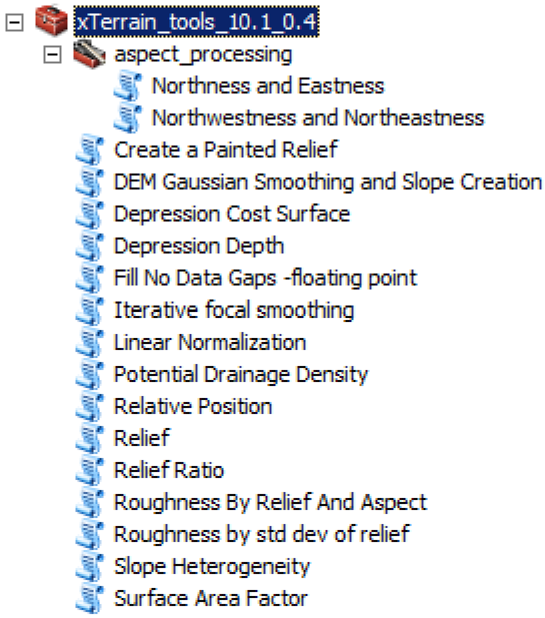
Navigate to the folder location, select "xTerrain_tools_10.1_0.4.tbx," and click on "Open."



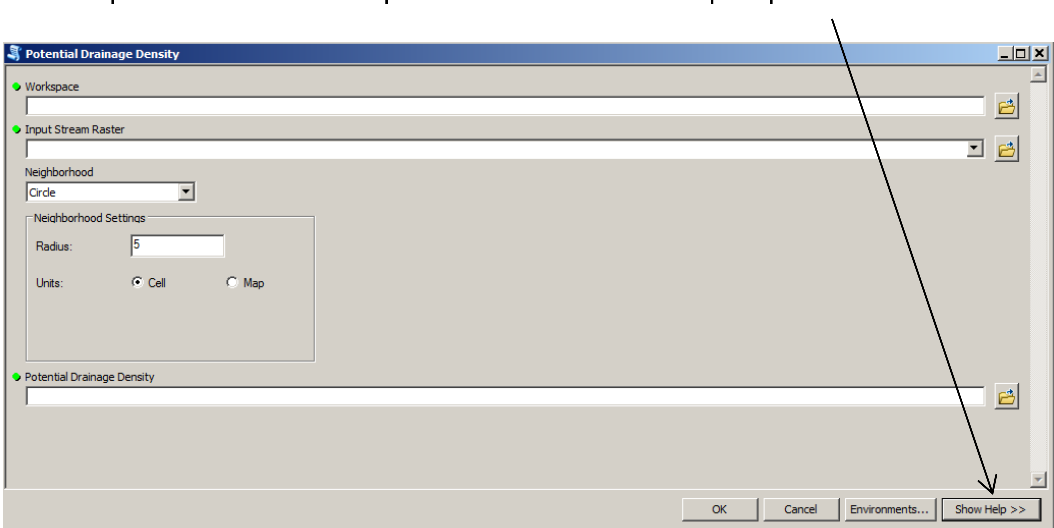
The toolbox is now added to the Arc Toolbox window.

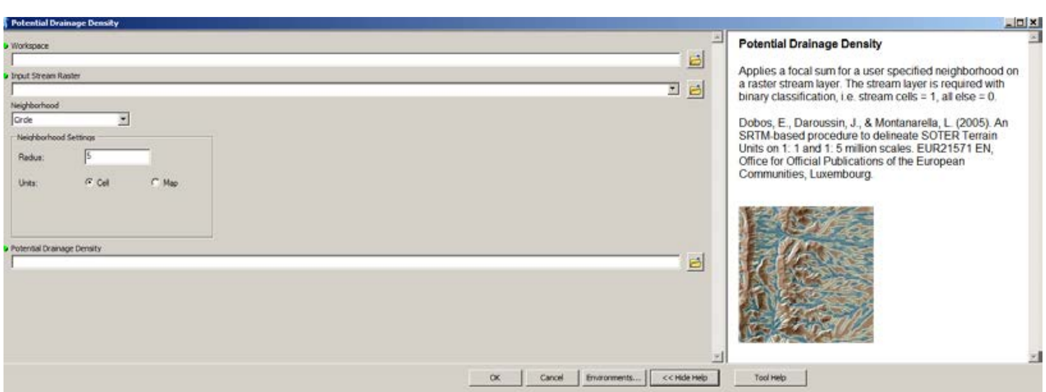


Contents of the Toolbox



A description of the tools are provided in the Show Help>> panel.





Citations

Jenny, H. 1941. Factors of soil formation. McGraw-Hill, New York.

Installation

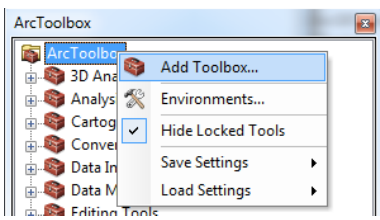
The Slope Class Script tool is an ArcGIS toolbox that contains a Python script. To execute the Python script via ArcGIS, the toolbox file (Slope Class Toolbox.tbx) must be added to the ArcGIS ArcToolbox or accessed through the ArcGIS Catalog browser.

The following link provides instructions for adding a Toolbox to the ArcToolbox Window:

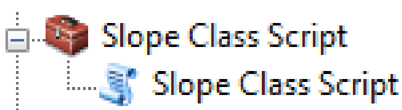
http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#//003q0000001m000000.htm

To add the toolbox into ArcToolbox so it consistently appears in new ArcMap workspaces, use ArcCatalog to install the toolbox.

1. Open ArcCatalog and click the 'ArcToolbox window' icon ()
2. The ArcToolbox window will open. Right-click on the primary ArcToolbox toolbox and click 'Add Toolbox'



3. In the file selection dialog, browse to Slope Class Script.tbx and click Open.
4. The toolbox will be added to the ArcToolbox window. Now the tool will persist within ArcToolbox for all future ArcCatalog or ArcMap sessions.



Instructions

The tool (Figure 1) accepts four required parameters and one optional parameter. Parameter error-checking is built into the tool, so if incorrect values or files are submitted, the tool will not run and the user will be notified of the error. The following table details the tool parameters and expected values:

Tool Parameter	Expected Value
Select Slope Raster (required)	Enter the path to a slope raster derived from a digital elevation model or LiDAR. The raster must be in a projected coordinate system or the tool will not run.
Set Slope Raster Reclass Values (required)	This table of values is used to reclassify the original slope raster. In the 'Old values' field, enter non-overlapping ranges of values. In the 'New values' field, enter numbers unique to each range of values.

Input Minimum Delineation Size (acres) (required)	Enter a number that represents the minimum delineation size (in acres) for slope class cell clusters. Groups of reclassified slope class raster cells with areas less than the minimum delineation size will not be included in the final output.
Select Geodatabase for Output (required)	Browse to and select a file geodatabase that will contain temporary files and store the final output. It is optimal to use a new or empty file geodatabase so files are not overwritten.
Delete Temporary Files? (optional)	If this option is selected, temporary files will be deleted from the file geodatabase. However, some important intermediate raster files will not be deleted.

When all parameters have been entered, click 'OK' on the tool dialog. The script will start running and progress will be indicated in the script dialog progress window. Custom messages are printed to the progress window. If an error occurs, a message with red text will appear in the dialog. Once the script is finished running, the final slope polygons feature class will be added to the ArcMap display. If the 'Delete Temporary Files?' parameter is checked, only four files should remain in the output database:

1. **SlopePolygons –** The final attributed feature class containing generalized slope polygons.
2. **FinalRaster –** The final raster from which the SlopePolygons feature class is derived
3. **Stage1_FinalRaster** – The raster file after initial Stage 1 generalization
4. **Reclassified_Slope** – The original reclassified slope raster using the values entered in the tool reclassification table

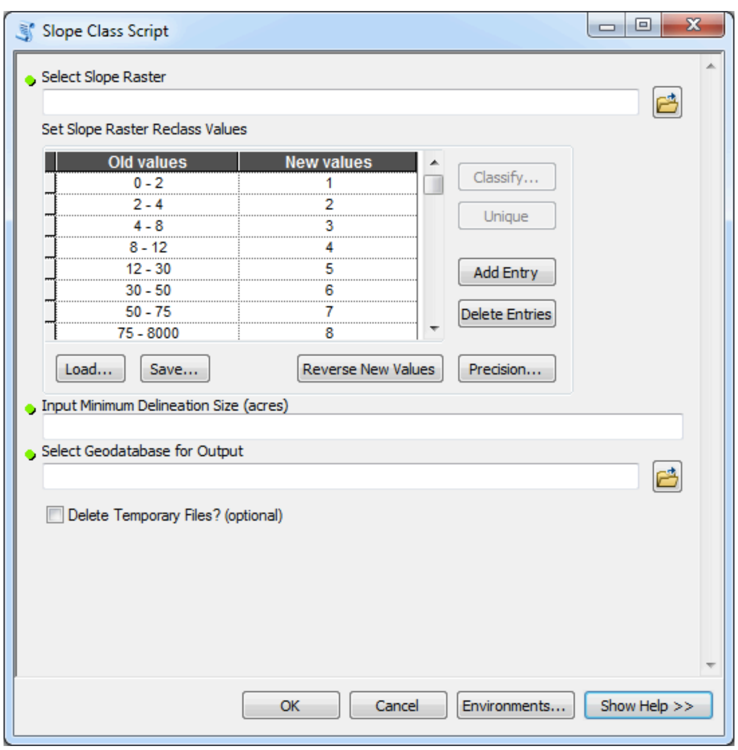


Figure 1. The GUI for the Slope Class Script

Tips

Dealing with high resolution data

High resolution elevation data derived from LiDAR varies widely in the amount of processing applied to the resulting product. Artifacts or anomalies that do not represent the bare-earth elevation may occur in LiDAR derived DEMs. It is common to perform neighborhood operations such as Focal mean on the DEM prior to the creation of slope maps to filter and smooth the resulting surface and minimize the effects of these artifacts.

A routine to reduce the effect of dead fall, dense understory or dense herbaceous vegetation like cattails has proven useful:

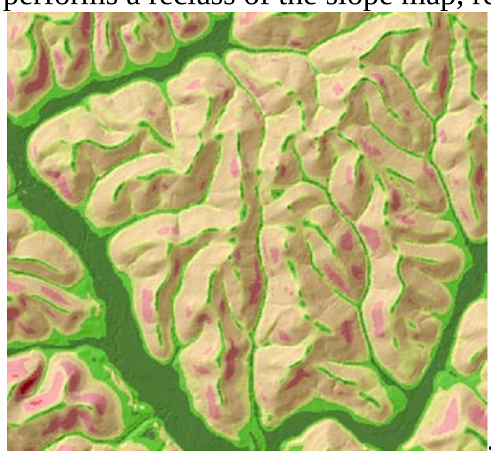
1) **Focal statistics –** minimum, using a two cell radius
2) **Focal statistics –** maximum, using two cell radius on the output of step 1
3) **Focal statistics –** mean, using two cell radius on the output of step 2

Processing time

Larger files may take 4-8 hours to process, depending on the CPU and available RAM. If the script crashes with no error, one potential solution would be to subset the slope file into smaller extents using a tool like Extract by Mask. This will require joining and editing of the resulting polygons from each subset.

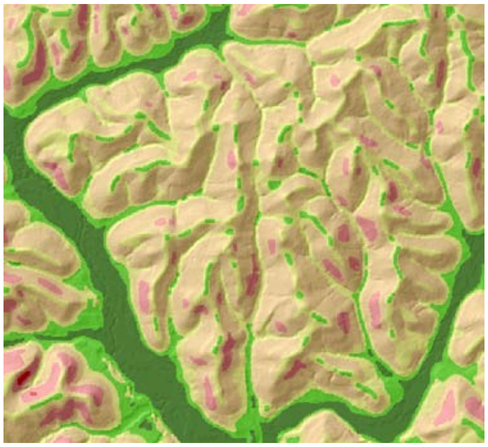
Appendix

There are 3 parts referred to in the script dialog progress window during execution of the script. Stage 1 performs a reclass of the slope map, resulting in a file named “Reclassified_Slope”(Fig. 2)



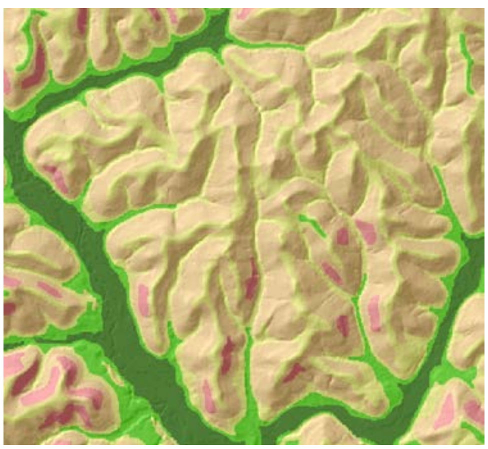
. Figure 2. Reclassified_Slope

A generalization routine is then performed on the entire layer using three progressively larger clusters of cells resulting in a file named **“**Stage1_FinalRaster” (Fig. 3).

 Figure 3. Stage1_FinalRaster

The Stage 2 process works by having adjacent slope classes generalized independently, increasing the likelihood that slope classes will be generalized into a similar, adjacent slope class. For example, the first process has 0-2 and 2-4 percent slope classes extracted and generalized using a specified cluster size, the second process has 2-4 and 4-8 percent slope classes extracted and generalized, etc.

Stage 3 runs another filter on the mosaiced output of Stage 2, producing a final, generalized slope class raster named "FinalRaster" (Fig. 4)

 Figure 4. FinalRaster

Finally, a raster to polygon operation is performed, producing a smoothed, attributed polygon file named "SoilPolygons" (Fig. 5)

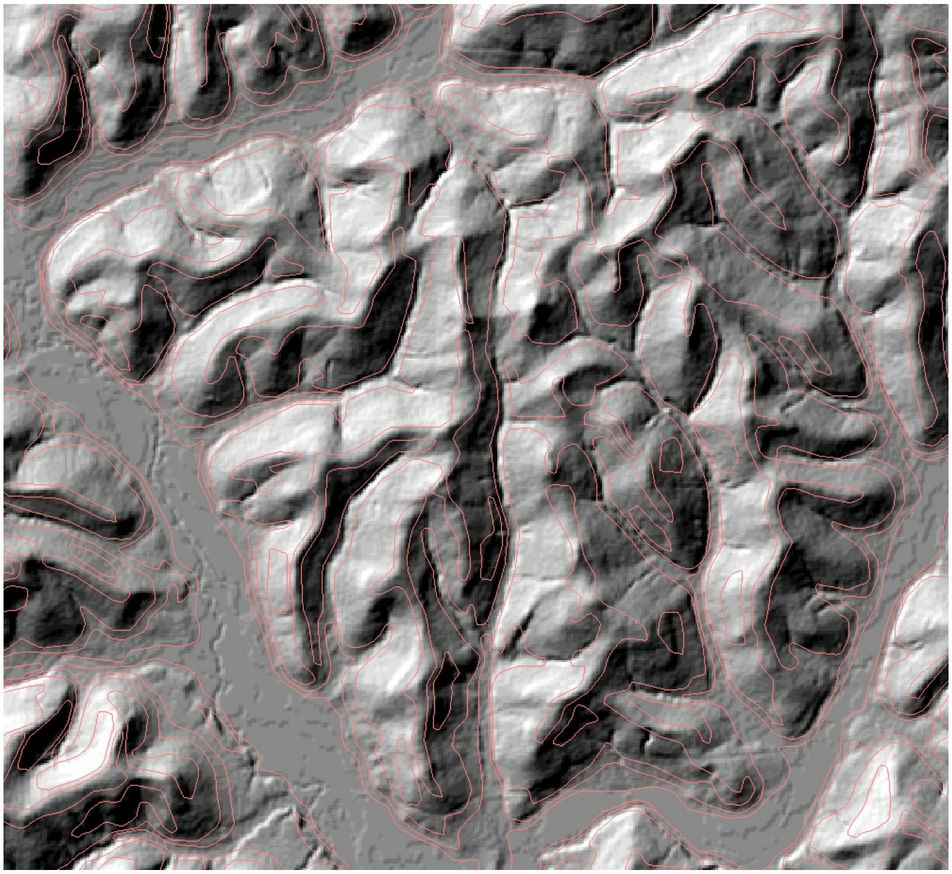


Figure 5. SoilPolygons

Fractional Vegetation

This takes the normalized difference vegetation index (NDVI) ratio and produces a normalized output expressed as a percentage ranging from 0 to 100.

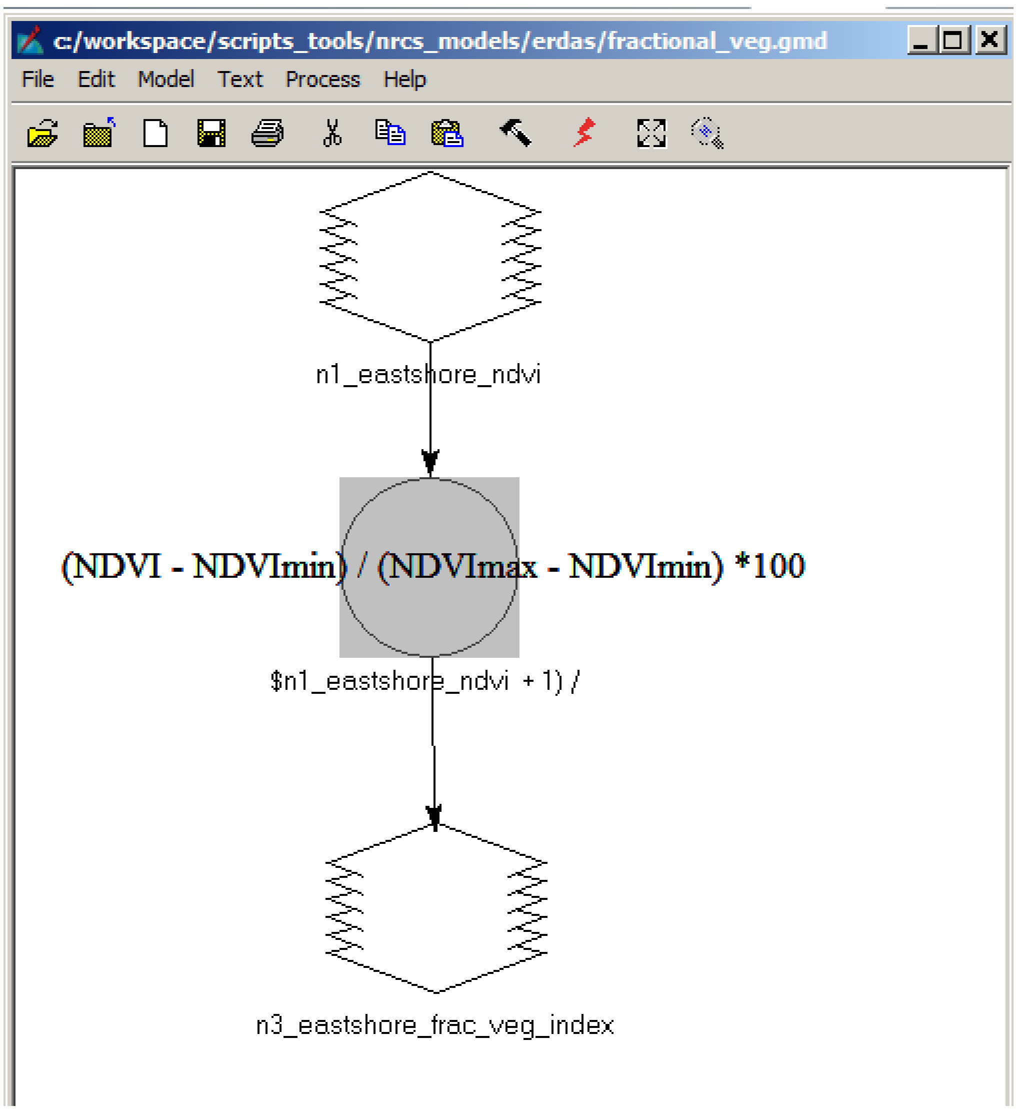
File: fractional_veg.gmd (ERDAS Imagine Model, v. 2013 and prior)**

Input: NDVI

Output: single layer image with data values ranging from 0 to 100; Pixels with higher values can be interpreted as pixels with greater vegetative cover.

You will be prompted for the input image, and to name the output image.

The model open in ERDAS Imagine Model Maker:



*Remember ratios must always be calculated on images which have had an atmospheric correction or image standardization applied. Resources for image standardization: http://earth.gis.usu.edu/imagestd/

**These are .gmd ERDAS Imagine Model Maker files. They can easily be opened in ERDAS 2013 and all previous versions of Imagine. Starting with ERDAS 2014, they must be converted to the new Spatial Modeler interface and file type.

While these ratios were developed for Landsat bands, they could be modified and used with data from any sensor with bands capturing the same part of the electromagnetic spectrum as the indicated Landsat bands.

Enhanced Thematic Mapper Plus (ETM+)	**Landsat 7**	**Wavelength (micrometers)**	**Resolution (meters)**
	Band 1	0.45-0.52	30
	Band 2	0.52-0.60	30
	Band 3	0.63-0.69	30
	Band 4	0.77-0.90	30
	Band 5	1.55-1.75	30
	Band 6	10.40-12.50	60 * (30)
	Band 7	2.09-2.35	30
	Band 8	.52-.90	15

Modifying Digital Elevation Models to Develop More Realistic Wetness Index Layers for Soil Survey Applications

By NRCS Soil Scientists Tom D'Avello, Joe Brennan, and Lynn Loomis.

Background

Performing additional processes on Digital Elevation Models is a common practice. Several basic techniques are reviewed in the May 2016 NCSS Newsletter (page 16). Hydrologically based covariates are grounded on the basis of watersheds and include:

- Flow accumulation
- Slope length
- Stream link
- Stream order
- Stream power index
- Upslope contributing area
- Watershed
- Wetness Index

Wetness Index (aka Compound Topographic Index, Topographic Wetness Index, or Topographic Index; Moore et al., 1988) is one of the most common and useful hydrologically based covariates used in soil survey applications (Gessler et al., 2000). Generally speaking, Wetness Index is the ratio of upslope contributing area and slope gradient. Manmade features like roads and railroads often confound the usefulness of this covariate by functioning as ridges or pits and creating micro-topographic noise. Figure 1 shows a northwest-to-southeast-trending linear feature that is a transportation right-of-way. The difference in cell values is apparent along the juncture of the right-of-way where artificial lows have been created (in the vicinity of "A"). Soil scientists ignore these minor features when creating polygon-based soil survey maps. When raster data is used for the mapping inputs and is the desired output format, the features may result in undesired anomalies. The Remove Linear Artifacts tool, available in ArcSIE, is one method described in the May 2016 NCSS Newsletter for minimizing the effect of the manmade features. This document describes another useful method.

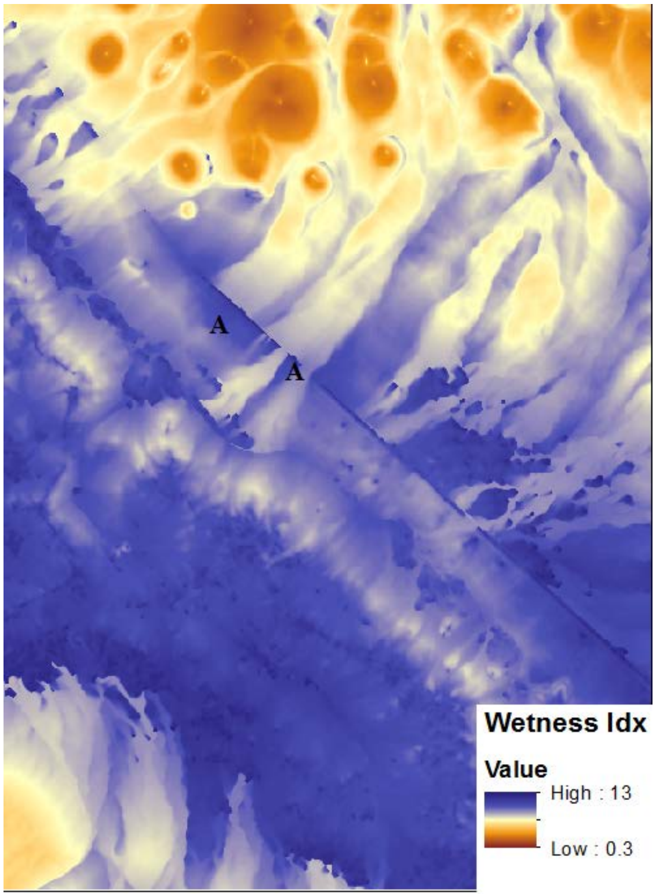


Figure 1.— A transportation feature bisecting a watershed.

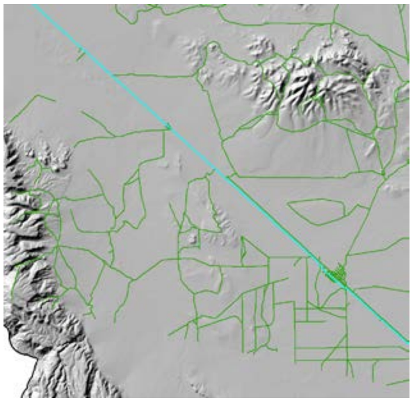
Preprocessing Checklist

1) Verify the data source
2) Verify the projection parameters
3) Verify the horizontal **and** vertical units
4) Verify the resolution
5) Verify the extent in terms of rows and columns
6) Verify raster statistics
7) Set a snap raster
8) Buffer the project area
9) Use complete watershed extents for hydrologically based derivatives
10) Perform a qualitative check using a hillshade
11) Display transportation layer
12) Process data
13) Store output in a common folder

Display the transportation vector layer over the DEM and make an inventory of transportation segments that are likely to introduce problems. The presence of smaller, local roads is usually of minor consequence and typically can be ignored. Large, multi-lane roadways and railroads, however, can create problems with hydrologically based covariates.

Steps for Minimizing Anthropogenic Features

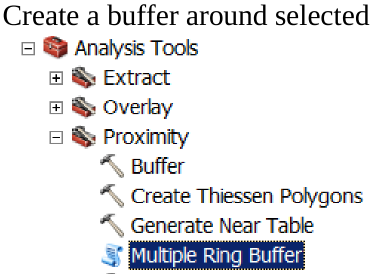
The example area, in the Basin and Range Province, indicates the presence of a U.S. highway with an adjacent railroad line (Figure 2). This feature, which bisects the watershed, will introduce anomalies that interrupt and confound the general landform trends, as revealed in Figure 1. The following process will remove the area represented by the transportation features, interpolate new elevations for a "roadless" DEM, and mosaic the "roadless" DEM back to the original DEM.

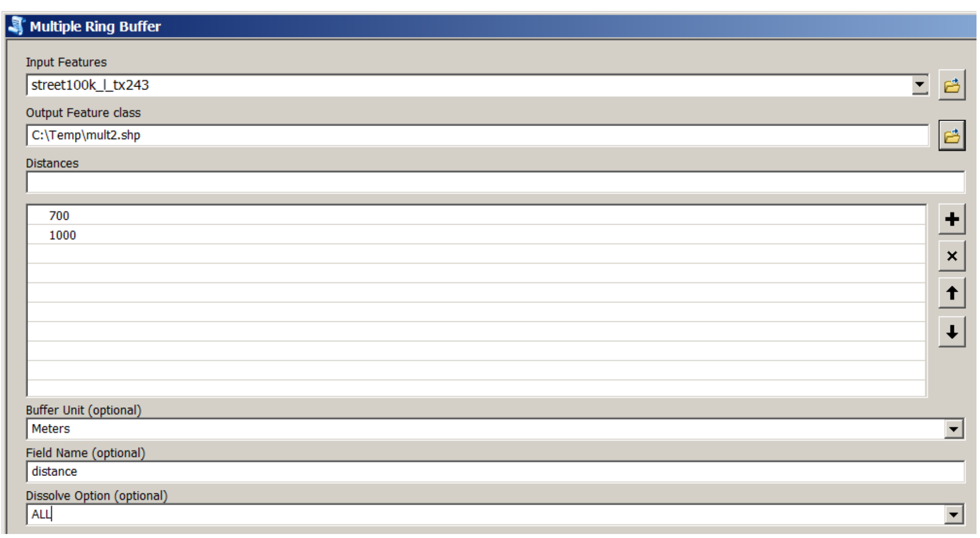

a.

b.

Figure 2. —U.S. highway highlighted with local roads (a), and a landscape image of the area (b).

1) Select the problem feature(s).
2) Create a buffer around selected feature(s) using the Multiple Ring Buffer tool:





This tool creates a polygon resembling a donut (see Figure 3). Use buffer distances large enough to be free of the effects of the transportation features.

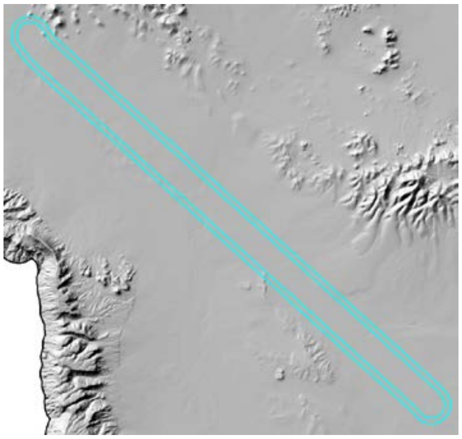
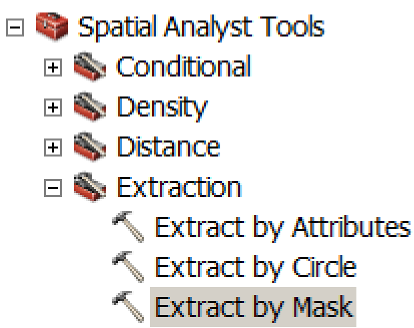
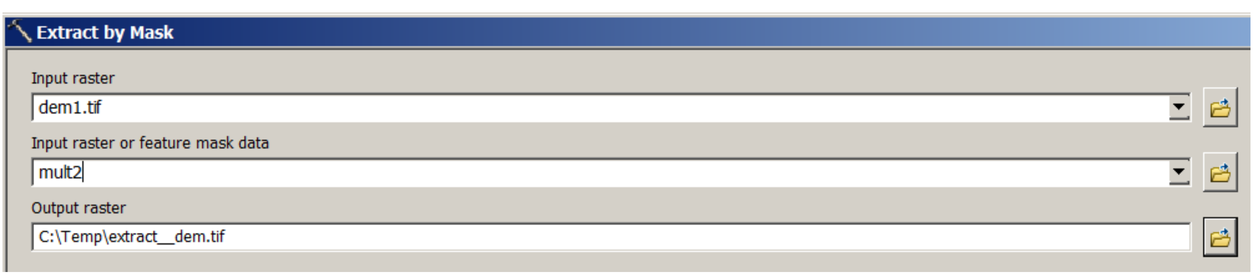


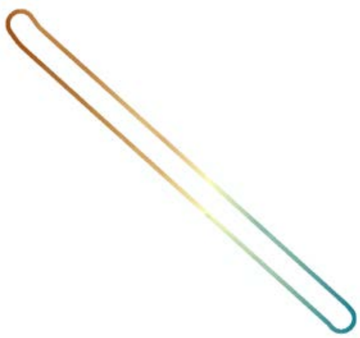
Figure 3. —Output of the Multiple Ring Buffer with outer ring selected.

3) Select the outer polygon ring.
4) Run Extract By MASK on your DEM of interest:

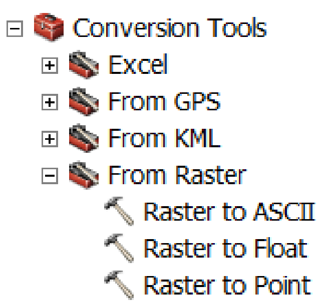




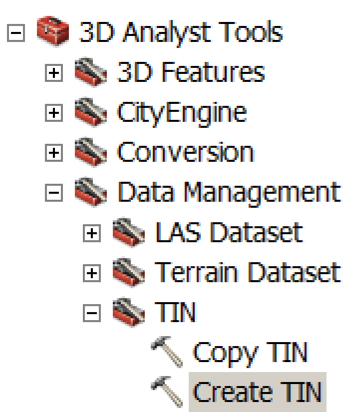
The result is a DEM representing the “donut” without the transportation features.



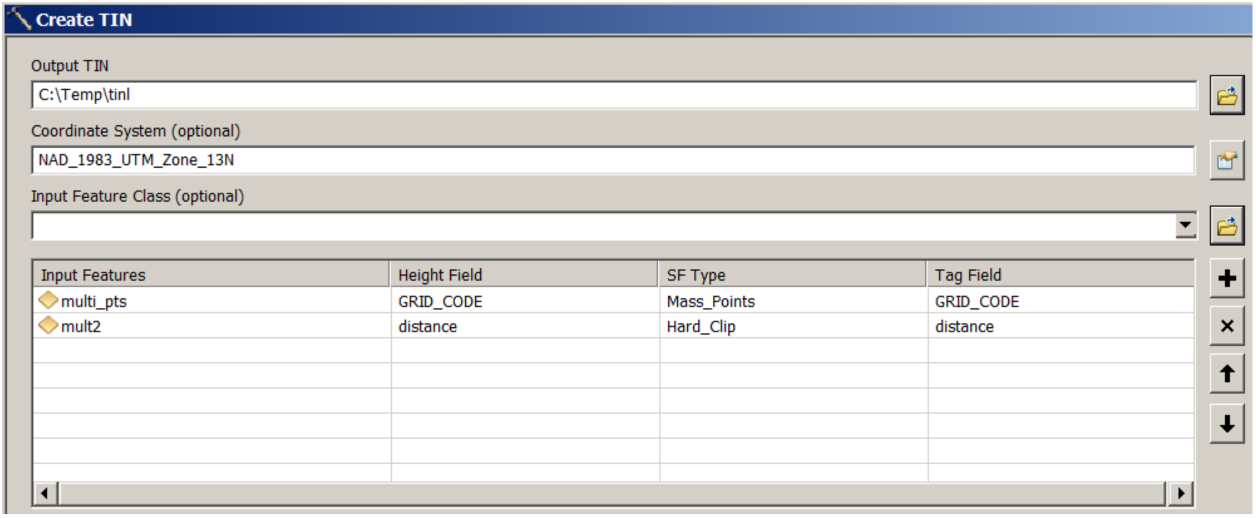
5) Convert this raster output to a point file using Raster to Point:



6) Create a TIN:

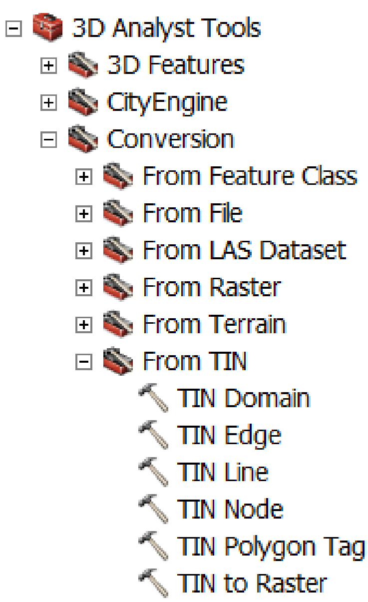


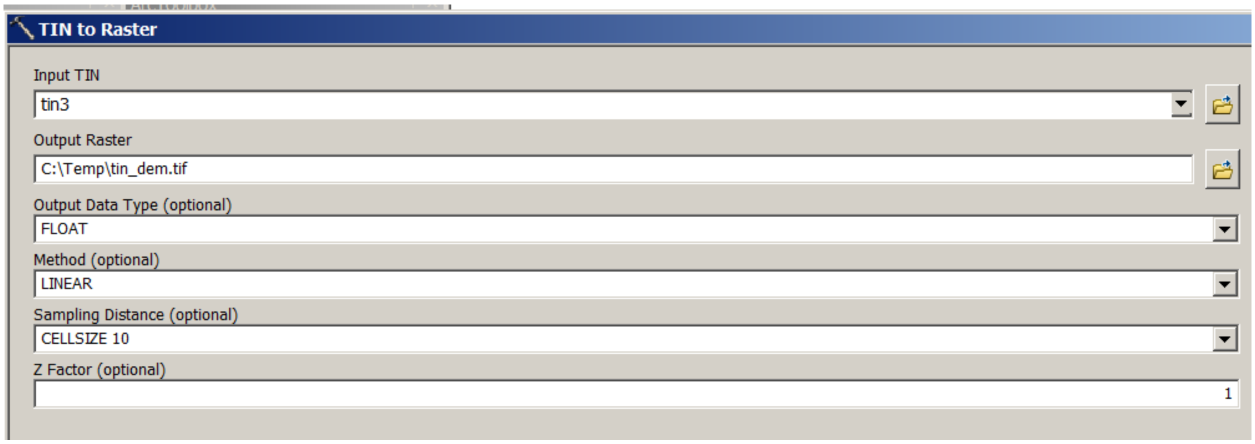
Use two input files to create the TIN: the point file from Step 5 as the Mass Points surface feature type and a polygon file representing the maximum buffer distance used in the Multiple Ring Buffer operation as the Harp Clip feature type.



The output shows a TIN with the complete surface for the “donut” and “hole” constrained to the outer bounds of the polygon used for the Hard Clip.

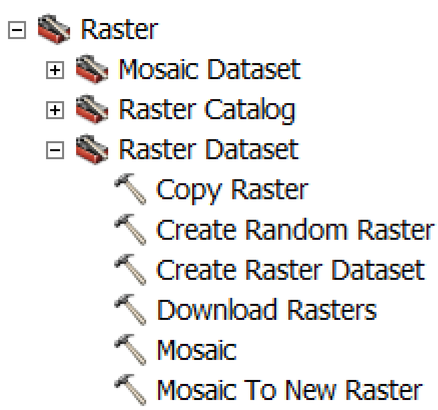
7) Convert TIN to Raster (be sure to specify correct cell resolution):

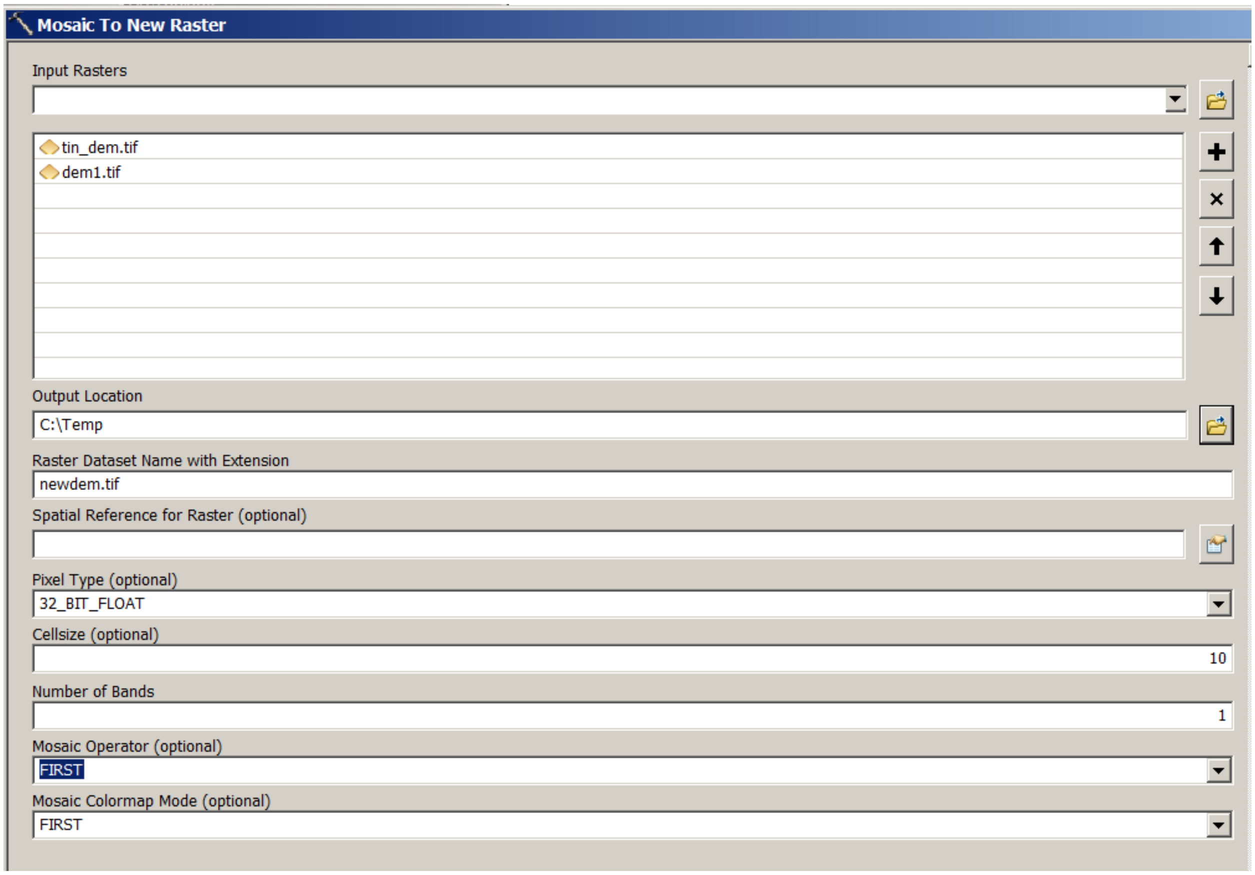




The output should be a 10 meter resolution raster file of the TIN from Step 6.

8) Mosaic the DEM from Step 7 back to the original DEM of interest with NO blending, using the Mosaic to New Raster tool in the Data Management Tool Box:

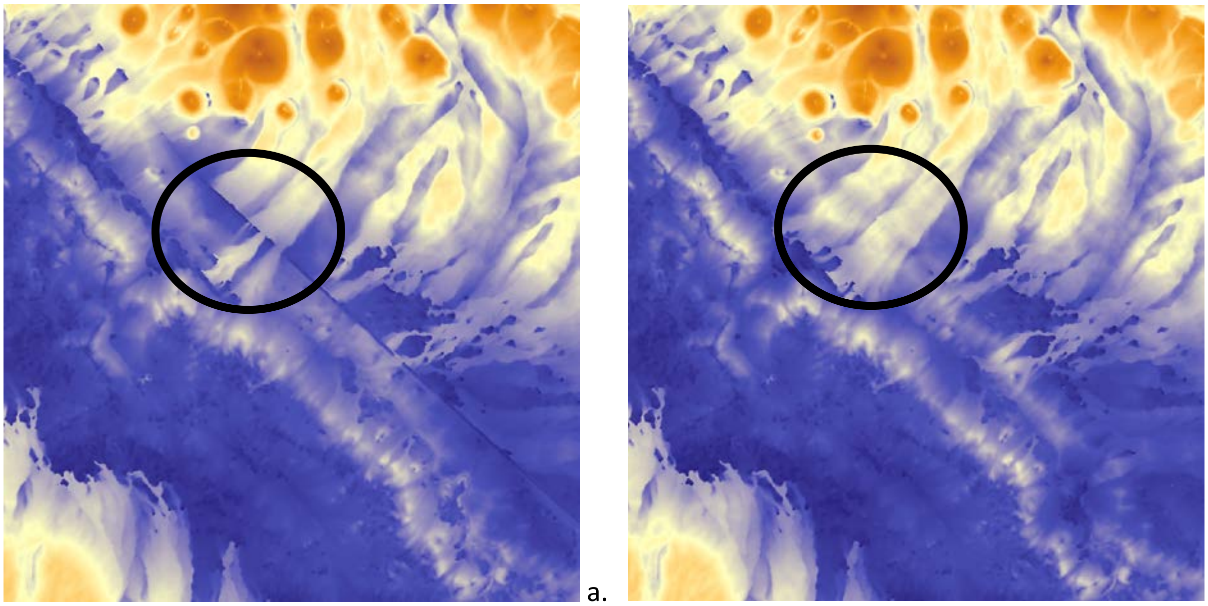




Selecting the FIRST Mosaic Operator ensures that cell values from the raster in Step 7 will be used for this overlapping area. The purpose of this modification is to create a DEM with transportation features eliminated or minimized.

Review Results

Wetness Index covariates were developed using the SAGA Wetness Index tool for the original and modified DEM. Viewing the area within the black ellipse in Figure 4, it is apparent that the fan deposits are interrupted by the transportation features on the original DEM in panel “a.” The modified DEM in panel “b” shows a more natural representation of the features and is more suitable for soil mapping purposes.



a. b.

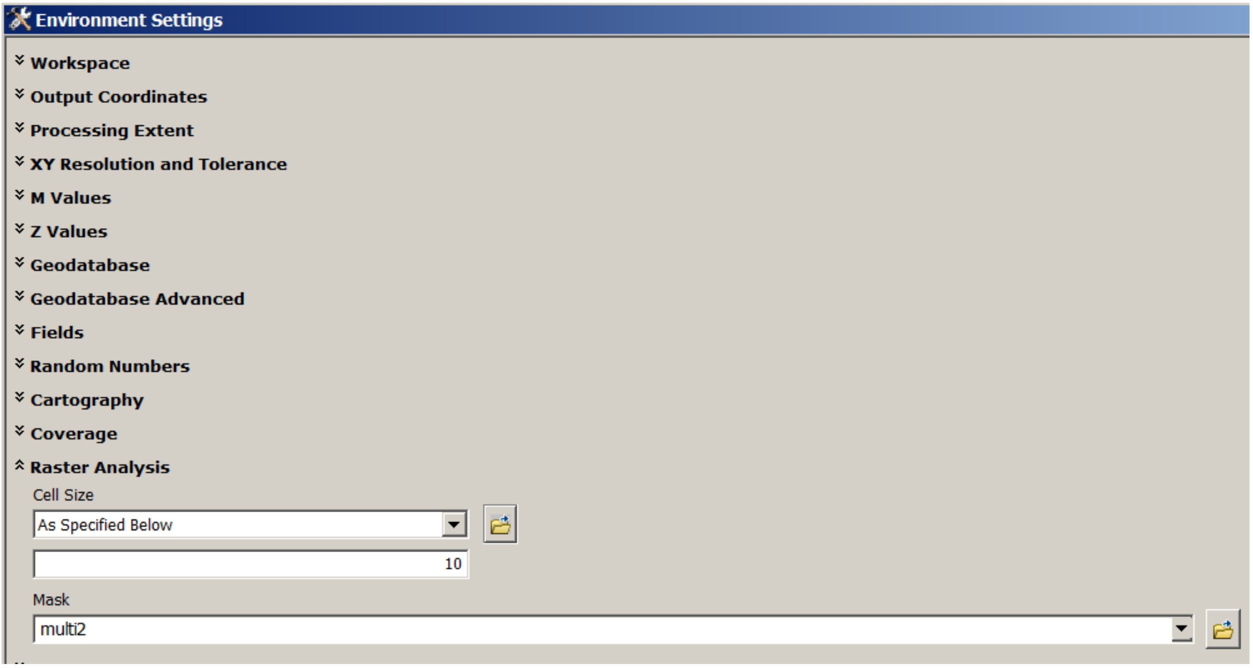
Figure 4. —Wetness Index from the original DEM (a) and modified DEM (b).

A desirable attribute of TINs is the direct relationship between point density and surface variability that may be established. TINs can have a higher resolution in areas with higher variability and a lower resolution in areas with lower variability. In this example, the assumption is a relatively uniform elevation surface within the created data "hole," which the TIN accommodates well. Using a Terrain data model will yield results similar or identical to those of the TIN data model. Both TIN and Terrain data model functions require the 3D Analyst Extension.

Miscellaneous

Interpolation techniques like Inverse Distance Weighting, Trend, Spline, Kriging, Local Polynomial, Global Polynomial, and Radial Basis Functions have also been tested. They were found to yield less satisfactory results and often require longer processing time. These tools require the Spatial Analyst or Geostatistical Analyst Extensions.

To be complete, the Natural Neighbor interpolation method is available with Spatial Analyst and produces output similar to TINs with short processing times. This is related to the similarities of the interpolation algorithms used by these two methods. Using Natural Neighbor would eliminate Steps 6 and 7 described above. If you use Natural Neighbor, you need to set a MASK in the Raster Analysis section of the Geoprocessing Environment Settings to confine the interpolation to the bounding polygon of the "clip" file from Step 2.



Although this job aid focuses on watershed-based derivatives, anthropogenic features can have undesirable results on terrain-based derivatives, such as relative position (aka relative elevation or normalized slope height). Figure 5 shows the effect of anthropogenic features on relative position.

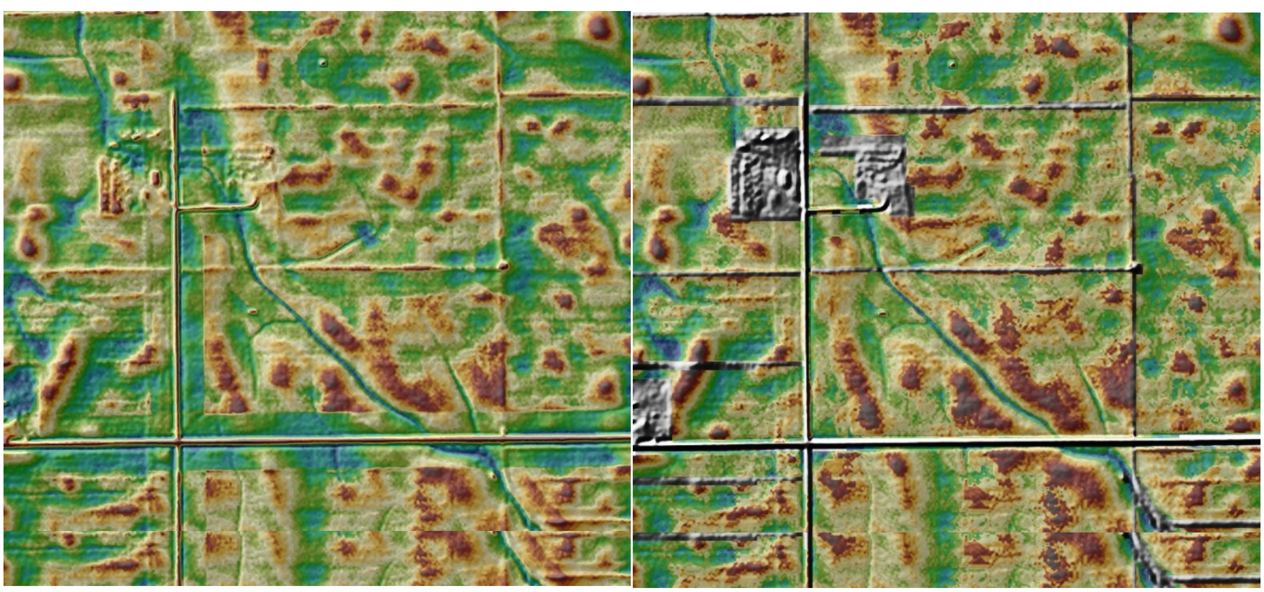


Figure 5. —Relative position calculated with (left) and without (right) anthropogenic features.

References

Gessler, P.E., O.A. Chadwick, F. Chamran, L. Althouse, and K. Holmes. 2000. Modeling soil–landscape and ecosystem properties using terrain attributes. Soil Science Society of America Journal 64:2046-2056.

Moore, I.D., G.J. Burch, and D.H. Mackenzie. 1988. Topographic effects on the distribution of surface soil water and the location of ephemeral gullies. Transactions of the ASAE 31(4):1098-1107.