ArcCatalog Tutorial 1

1.1 - Building a Catalog
1.2 - Viewing Images and Image Attributes
2.1 - Creating a GLOBE Image
2.2 - Creating a Shapefile
## 3.1 – Using High Spatial Resolution Aerial Photographic Reference Data

Tutorial Objectives: *Become familiar with several important ArcCatalog features, and create a GLOBE Landsat subscene for later use in Land Cover Mapping*

ArcCatalog is the ArcGIS tool used to access, browse, organize, and manage files (Fig. 1), a process called *building a catalog*. ArcCatalog is similar to the Windows Explorer applet, but specifically designed to manage *geographic* data. It is a simple but powerful program for digital geographic data management.

ArcCatalog also provides connections to networked databases (Database Connections Folder), Address Locators, GIS Servers (ArcGIS Servers, ArcIMS Internet Servers), and provides a folder where search results are stored. These connections are not addressed in this tutorial.

Before beginning this tutorial, it is necessary to have the appropriate Landsat scene and SAMB aerial photographs (see the Data Access Tutorial).
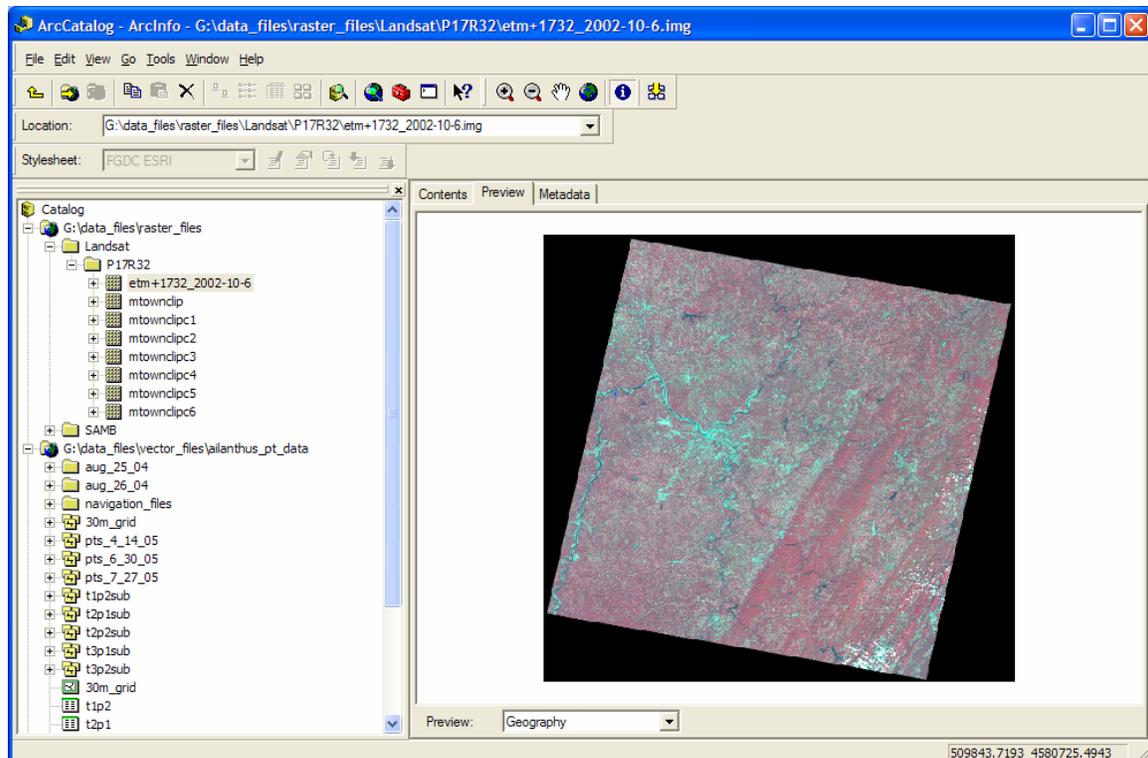
Figure 1. ArcCatalog, showing the *Catalog tree* in the left pane and a Preview of the Landsat image etm+1732_2002-10-6.img in the right pane.

Lesson 1. Working with Digital Geographic Data in ArcCatalog

### 1.1 Creating a Catalog - Connecting to Drives and Folders

Browse to and boot ArcCatalog using the Start button (Start/Programs/ArcGIS/ArcCatlog) View the two default screens (Fig. 1)

Click on the *Connect To Folder* icon 🖳 and the Connect to Folder dialog box appears (Fig. 2).
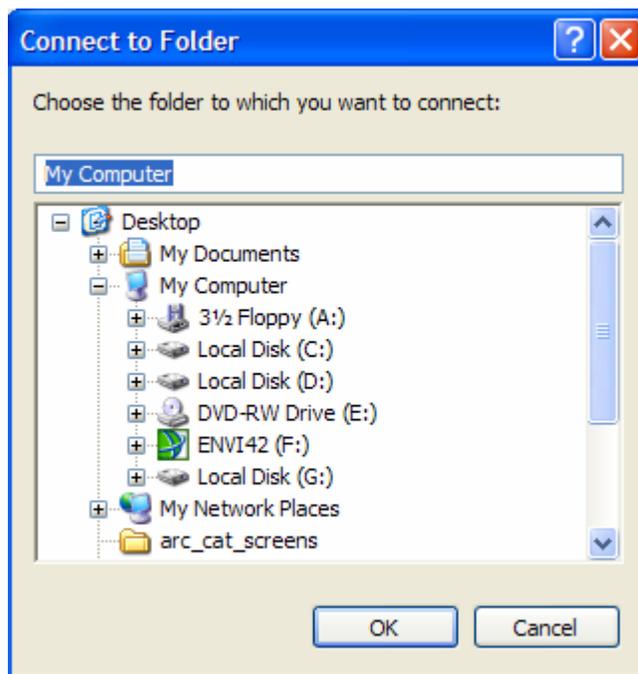
Figure 2. The *Connect To Folder* dialog box.

Once the Connect To Folder dialog box pops up, browse to the location of your data as you would using Windows Explorer or My Computer. Select the respective folder, then click on the 'OK' tab, creating a *connection*. That location is now accessed directly (Fig. 3). Note that *ArcCatalog does not permit connections to individual files*, only to drives and folders. Files are viewed simply by selecting (single-click highlighted), followed by clicking on the respective tabs labeled **Contents**, **Preview**, and **Metadata**.
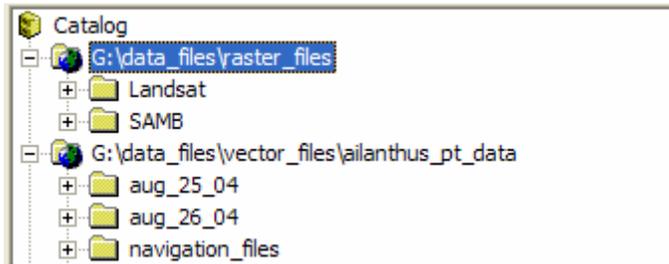
Figure 3. The *Catalog tree* showing connections to two folders.

You can also remove folders (disconnect) by highlighting the folder and clicking .
There are many other tasks possible once a connection is established, such as copying and pasting your data to other drives and directories, creating shape files (Lesson 2 or 3), and other GIS functions.

1.2 Viewing Data and Data Properties

Now that we've created an ArcCatalog connection, let's examine the image data. Browse to and click on the folder containing the Landsat file (etm+1732_2002-10-6.img), then select the image (single-click highlight). Note that the file's location is provided in the location window (Fig. 4).
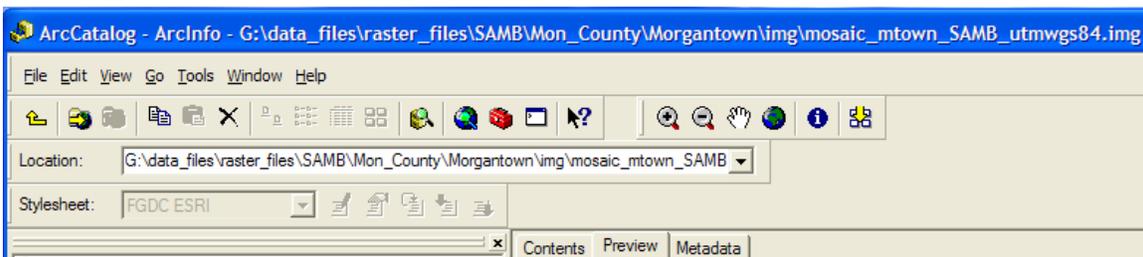
Figure 4. The location of the selected file is provided in the 'Location' window.

First, we'll view folder and file information using the 'Contents' tab.

Contents Tab:

The Contents Tab provides a list of items contained in each folder or file

Click on 'Contents', note the name (number of **bands**), **type,** and **projection. NLAPS** stands for National Landsat Archive Processing System and is a type of product generation system (we needn't worry about this).
Click the '+' to the left of the file name, expanding the raster bands (double-clicking the file will produce the same result). This will show each band within the file. Click on band 1, projection information is provided.
Click the '-', the bands collapse.
While the file is selected, right-click and draw down to 'Properties…', left-click and the raster dataset properties dialog box appears.

ArcCatalog Tutorial 1.0   page 3
West Virginia GLOBE, West Virginia View

Note: Data Source, **Raster** Information, Extent, **Spatial Reference**, and Statistics (we will examine each of these)

Preview Tab:

The Preview Tab shows raster images, coverages (vector files), and TINs

Select the Landsat scene and click on Preview. The full scene should appear. Zoom in and out, pan around the image using the 'hand' tool, and view the entire image by clicking the earth icon ("full extent").

Zoom in to see individual pixels. The 'i' icon provides information on the intensity value for each pixel in each of the three channels. Zoom in so that a single pixel fills most of your screen, and place the cursor in the upper left corner of the pixel. Note the coordinates at the bottom right of the screen; they should be close to xxxxxxx.00 meters in both the X (**Eastings**) and Y (**Northings**) dimensions. This is because the NLAPs processing system assigns pixel coordinate values to the upper left corner of each pixel. What would be an alternative way to assign coordinates?

Metadata Tab:

Metadata is information (data) about the data

The Metadata tab provides three types of information, available via a Description tab, a Spatial tab, and an Attributes tab. Accurate, complete metadata is critical because it provide the data's history - how the data was developed and processed, by whom, and when, among other things.
For our purposes, the 'Spatial' information is most critical. We need to be working in the UTM grid (we're in zone 17 North) and the WGS84 **ellipsoid** and **datum**.

Click on the 'Metadata' tab, then on 'Spatial'. Check the coordinate system of our file. It should read **UTM Zone** 17N, WGS84. If it doesn't, we need to reproject it (see Defining and Projecting, p. 8).

We can do the same with the SAMB high spatial resolution aerial photos (see Data Access Tutorial).

Lesson 2. Manipulating Digital Geographic Data

2.1 Creating a Georeferenced GLOBE Landsat Scene in ArcToolbox

Now that we're somewhat familiar with ArcCatalog, let's create the satellite image we'll be using in the Land Cover – Biometry protocol. Recall that GLOBE standards require that the scene is 15km square and centered on the GLOBE school of interest. Given that Landsat imagery is 30m spatial resolution, 500 pixels would measure exactly 15km. Yet the actual requirement for a subscene is 512 pixels x 512 pixels, or 15.36km square.

GLOBE must have a reason for the 512 x 512 pixel standard (perhaps this provides a sufficient buffer?).

Why do we want a **georeferenced** subscene? A georeferenced subscene is not only a highly accurate map (that is, an **orthorectified**, geographically correct digital image with north at the top and west to left), but it also has geographic coordinate information for every specific location "on" the map's surface. We can "GPS" a point (for example, the center of our Biometry sample location), and find its location on the Landsat image and on **aerial photographs** we'll use for **ground reference data** collection. So, you can understand why a georeferenced image is so important to the GLOBE Land Cover mapping protocol. It allows one to determine their exact geographic location "on" the image.

Before we create the subscene, we need to find the GLOBE school. Let's use South Middle School in the Monongalia County school district. Use the zoom tool to zoom in to the building. Now, record the UTM coordinates of South. 'Y" is *northings*, 'X' is *eastings* and increase from west to east.

Eastings ___589715.00____, **Northings** ___4385187.00_____

Now that we have the location of the school, we need to establish the boundaries of our GLOBE image. We know we need a 512 pixel by 512 pixel subscene, so we divide 512 by 2 in each direction from center and we get our dimensions in pixels.

512 pixels

256 pixels

256 pixels   256 pixels

256 pixels

512 pixels
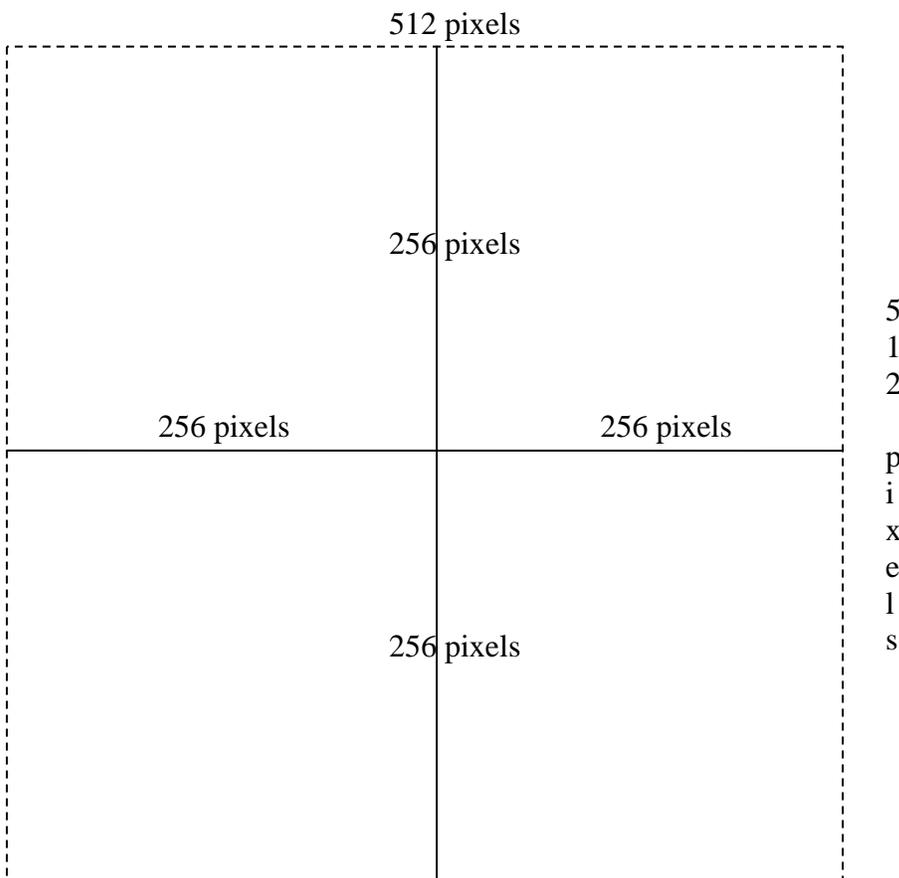
OK, but our dimensions are in pixels and the image coordinates are in UTM, which is in meters. Thus we have to convert our pixel dimensions to meters, using the center point (589715.00, 4385187.00). So, 256 pixels x 30 m / pixel = 7680 meters. Do the math, making sure we convert the Eastings and Northings to their respective image file X,Y coordinate values.

E 589715 + 7680 = 597365.00; we're establishing the eastern boundary.
E 589715 – 7680 = 582035.00; western boundary extent.
N 4385187 + 7680 = 4392865.00; northern boundary extent.
N 4385187 – 7680 = 4377505.00; southern boundary extent.

Now that we have our values we can perform the clipping process:

First, click on the small red 'toolbox' located in the standard menu, starting the ArcToolbox program. ArcToolbox is a very powerful and useful application used to process both raster and vector data (note that it can be launched from ArcCatalog or ArcMap). Open the Data Management Tools, then click on 'Raster'.

Double-click on 'Clip' icon Clip , enter the coordinates, name the output and note file location.
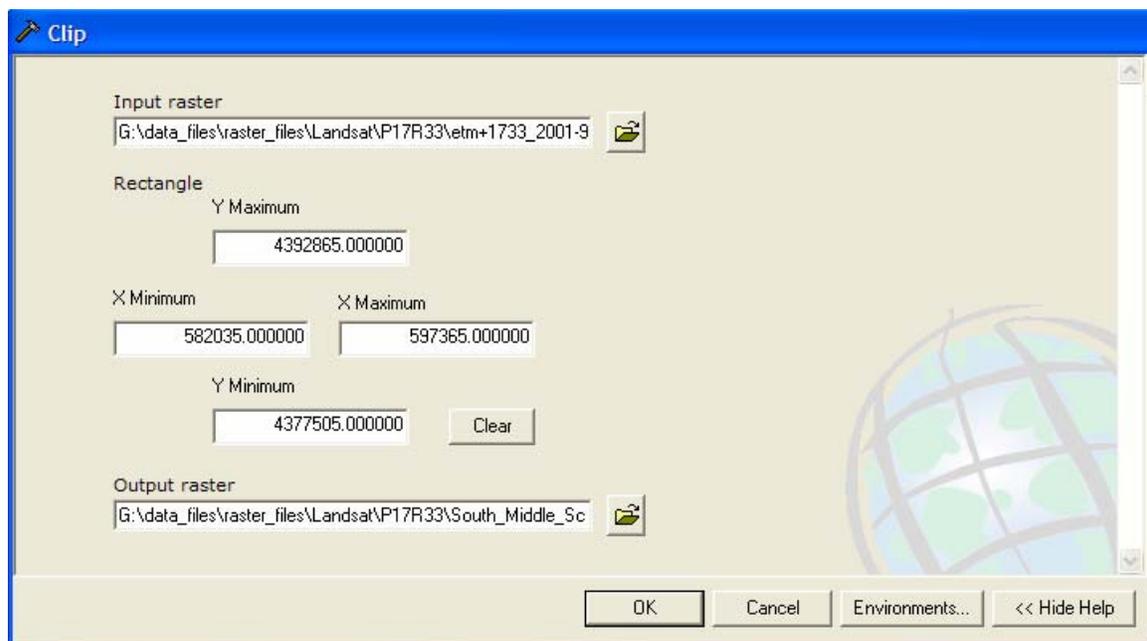
Figure 5. The ArcToolbox > Data Managament Tools > Raster > Clip screen

Check the resulting clipped image in ArcCatalog. Select it, right-click and draw down to 'Properties', check 'Raster Information'. It *should* show 512, 512 for columns and rows, and should contain all 6 bands. Also, check to see that the image retained its Spatial Reference (UTM zone 17N, WGS84). If so, that's it! We've got our official GLOBE-standard image.

2.2 Initiating the Shape File Creation Process

In order to perform the Land Cover GIS mapping exercise, we'll need a "shape file". Shape files are an ArcView **vector** file standard, a commonly used, indeed almost universal, file format that MultiSpec and many other GIS and remote sensing programs can handle.

First, we need to create a folder location for the shape file (alternatively, we could use an existing folder).

Browse to the folder data_files\vector_files\shape_files. Create a connection here is you haven't already done so (see 1.1 for details)

Now, select the location, right-click, draw down to 'New', then Shapefile…
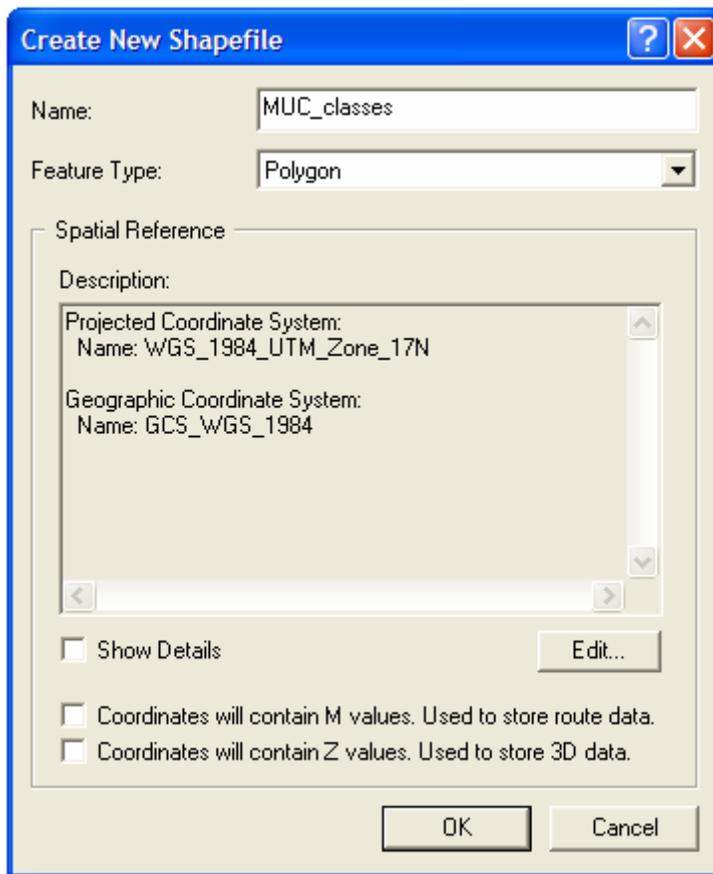
Figure 6. The Create new Shapefile dialog box.

Now, name the file, set the 'Feature Type' to Polygon, click the 'Edit' tab and select the appropriate Spatial Reference system (UTM WGS84), and click 'OK'. That's it; you've created your shape file. We'll map the MUC classes in a later exercise using ArcMap (see ArcMap Tutorial 2.1 – Digitizing the MUC Land Cover Classe)s.

The next step is accessing and preparing our high resolution aerial photo reference data that we'll use along with the Landsat image for generating our digital land cover map. We'll generate the map itself in ArcMap, and we'll use a 5km by 5km Landsat subscene

(because we're time-limited), centered on South Middle School. However we will supplement the smaller Landsat image with the high spatial resolution SAMB aerial photographs. Both the 5km by 5km Landsat subscene and the SAMB aerial photograph are provided for you on disk.

Finally, before we move on to the MUC classification exercise, *keep in mind that the Land Cover – Biometry protocol has an important field exercise component*. We'll be using the aerial photos to "sample", or "ground reference" the areas that are not visited and sampled in the field using the biometry protocol. This is a standard technique in remote sensing.

3.1 Using High Spatial Resolution Aerial Photographic Reference Data

Let's face it, we don't have the time or support to "ground truth" each and every land use unit in our 15km x 15km area. That's OK, because there are alternative methods that we can use to generate our land use classification accuracy assessment map. What we need is accurate, high resolution digital air photos to check against the classified Landsat satellite image. The SAMB digital aerial photographs (see download instructions) are perfect for this exercise.

A. Defining the correct projection

Assuming we've downloaded the SAMB images, we need to perform one simple task before we can use them as a base layer for our Land Cover map. We need to 'define' the projection of each image. To do this, we browse to the folder where our SAMB images are stored, select an image, and view the 'Metadata' in the Display window. Select the 'Spatial' tab and note that there is no projection. However, this information exists in the metadata file associated with each SAMB image, but it hasn't yet been applied to the image. To do this, right-click on the image file in the Catalog Tree, go down t 'Properties…' and open the properties dialog box, scroll down to 'Spatial Reference' and you'll see <undefined>. Click the 'Edit' tab, click the select button, open the 'Projected coordinate Systems folder, open 'State Plane', then open NAD 1983 (Feet). Now, scroll down to West Virginia North FIPS 4701 (Feet).prj and select it so it shows up in the Name: box. Click 'Add", then 'Apply' and it should be shown in the Spatial Reference properties dialog box. Click 'OK', the 'Apply' and 'OK" once more, and that's it. Now, refresh the catalog Tree (View > Refresh, or F5 key), and now examine the 'Properties…' dialog box once more t make sure the image is projected.

B. Changing the image's projection

Again, we need to be working with data that is in the same projection. We therefore need to "reproject" the SAMB photos from NAD 83 to UTM WGS 84 (a GLOBE standard, and indeed a very common projection). We'll use Arctoolbox to perform this process. Boot Arctoolbox, scroll down to Data Management Tools', select and open 'Projections and Transformations', then 'Raster', then 'Project Raster'. Double-click on it and the dialog box will open. Browse to your SAMB data for 'Input Raster', select te image, 'Add' it, and note that it automatically names the new output raster but doen not provide a file type ending (it defaults to an ESRI GRID raster file). We want a .img file

(Multispec works well with these file types, as do many other image processing programs). So, rename the file the same exact name as its original, but place .img at the end (n-k08-19.img). Now select the output coordinate system (as we did before in the previous exercise, only this time select 'Projected Coordinate Systems', 'Utm', Wgs 1984, then scroll down to WGS 1984 UTM Zone 17N.prj, 'Add' it, 'Apply', 'OK', accept the remaining defaults, and the wizard should start. It'll take a minute, and then you can check its new projection as we did in the previous exercise.

At this point, and assuming we've downloaded all of our data, we're ready to create our Land Cover map in ArcMap.

GLOSSARY

Aerial Photograph (also air photo) – typically a nadir-looking photograph taken from either a fixed-wing or rotor aircraft. Aerial photographs are typically orthorectified and georeferenced and thus can be used like a map.

ArcCatalog – one of several ArcGIS programs, used primarily to organize digital geographic data.

ArcGIS – the current ESRI flagship Geographic Information System (GIS) software that includes a series of integrated programs to manage, create, process, analyze, and report geographic data.

ArcMap – the ArcGIS mapping program, also includes some data creation and analysis features.

ArcView – the suite of GIS programs that includes ArcCatalog, ArcMap and ArcToolbox.

Bands (also channels) – bands or channels are wavelength locations within the electromagnetic spectrum. Visible light is typically divided into three bands (because the human eye is only sensitive to these three colors). The "blue band" exists between approximately 450 to 550 nanometers (or .45 to .55 microns); the "green band" exists between 550 and 650 nm, and the "red band" exists from 650 to 750 or so.

Coordinate System – a geographic reference system, such as Universal Trans Mercator or Geographic Lat/Longs.

Datum (also geodetic datum) – a geodetic datum is a model of the earth. Datums define both the size and the shape of the earth, as well as orient coordinate systems that are used to map the earth. Referencing geodetic coordinates to the wrong datum can result in significant spatial error. See also **Map Projections** and **Coordinate Systems**.

Ellipsoid – a 3-dimentional surface, such as the earth, defined mathematically (can be a sphere, cigar-shaped, or pill-shaped). Think of it this way… take an ellipse and "spin it" about its major axis, and the resulting 3-D "object" is the ellipsoid (see http://en.wikipedia.org/wiki/Ellipsoid).

Georeferencing – the process of providing a coordinate system to an (orthorectified) image or map.

GLOBE – **G**lobal **L**earning and **O**bservations to **B**enefit the **E**arth. A educational program sponsored by NSF, NASA, NOAA and other organizations designed to teach science and environmental understanding to K-12 students.

GIS – Geographic Information System. Hardware (computers, GPS receivers, etc.), software (e.g. ArcView), and people (you and I) that work together as an integrated system to create, management, and disseminate digital geographic data for decision-making.

GPS – Global Positioning System. A system of satellites, receives, and software that uses trigonometry to determine locations on the earth's surface.

Ground Reference Data – information on a surface feature or features gathered either on-site, from nearby, or from another trusted source such as a recent, high-quality aerial photo.

Land Cover / Biometry Protocol – The GLOBE protocol designed to teach remote sensing, land use ecology, and elements of biometry (quantitative biology, measuring vegetation).

Landsat – Land remote sensing satellite program dating back to 1972 and currently managed by the USGS. Landsat 5 (launched in 1984) and 7 (in 1999) are still operational. **For a Landsat-specific glossary, see** http://landsat.usgs.gov/glossary.php

Orthorectified – ortho means 'straight, correct, right', Orthorectification is the process of removing distortions in the image due to interactions between the camera, lens, aircraft orientation, and the ground surface. Orthorectification is the process of removing these distortions and making an image or photograph geographically "correct", like a map.

Projection (also map projection) - the process of transforming (*projecting*) features such as points, lines, and areas or polygons, from a *spherical* (3-D) surface to a flat (map, 2-D) surface, with minimal distortion.

Raster (see also Vector) – a two dimensional array of pixels, typically expressed either as columns and rows or as x,y coordinates; any data structure expressed as a grid. A "black and white" digital photograph is an example of a panchromatic or grey-scale raster image. A natural color digital photographic is an example of a 3-band raster image (R,G,B or red, green, and blue).

Remote Sensing – the process of acquiring information (data) on the earth's surface from a distance (usually from an aircraft or satellite), processing that information, and analyzing the information for decision-making.

Satellite Image – a remotely sensed (usually digital) image of the earth's surface acquired by a sensor onboard a satellite.

Vector (see also Raster) – a geographic data structure using the point, or node and respective connecting segment, as the basic building block for representing geographic features. Points, lines, and polygons are the three basic types.

WGS 84 (also **WGS84**) - an earth-centered datum used by the Global Positioning System called the World Geodetic System 1984 or WGS 84. WGS84 was adopted as a world standard from a datum called the North American Datum of 1983 or NAD 83. For all practical purposes there is no difference between WGS 84 and NAD 83.

For a more general remote sensing glossary, see the Canada Centre for Remote Sensing web page at:
http://www.ccrs.nrcan.gc.ca/ccrs/learn/terms/glossary/glossary_e.html