

# Spatial Analysis Tip Sheet – Preparing for Analysis

Barbara Parmenter, Tufts University  
 Updated March 5, 2013

This tip sheet covers two important set up steps before performing any analysis:

- Getting all your data sets into the same coordinate system
- Setting the geoprocessing environment settings

## All data and the data frame should be in the same coordinate system:

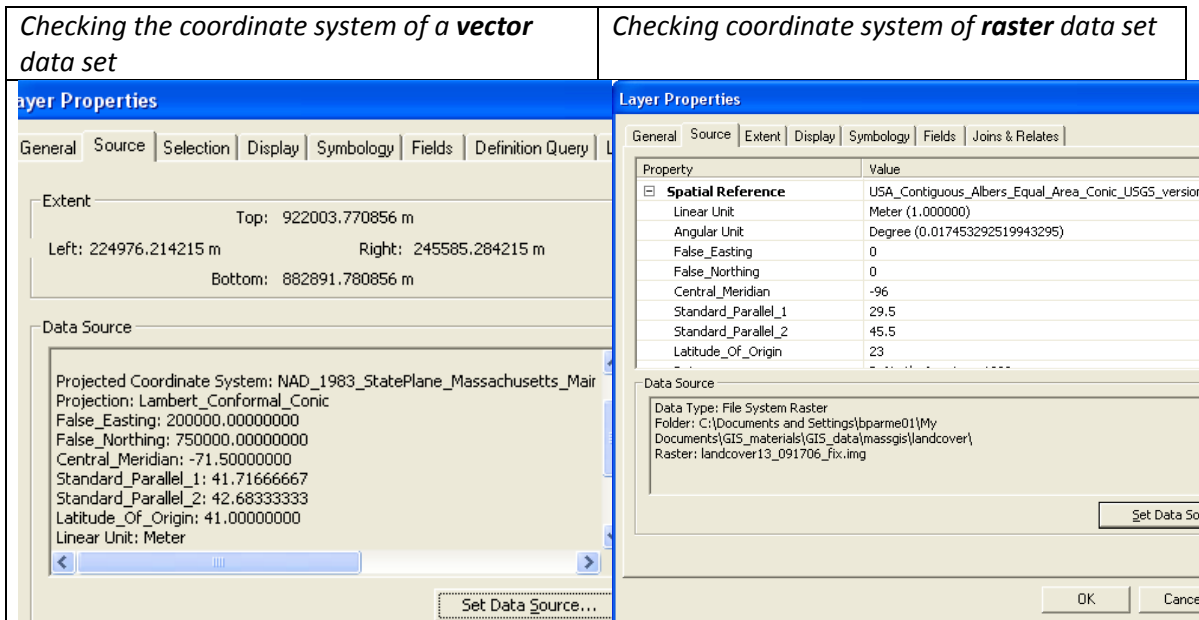
Before doing any analysis functions in ArcGIS, make sure that all the data sets that you will be using in your analysis are in the same map projection/coordinate system, and that the data frame is also in that coordinate system. For ANY analysis in ArcGIS, the layers you are using in your analysis MUST be in the same coordinate system, and this must match the coordinate system of the data frame. If you don't do this, you may either get errors, or the analysis function will appear to run, but nothing will happen.

For more information about coordinate systems, see the *Tip Sheets* regarding *Map Projections and Coordinate Systems* on the [Tufts GIS Center Tips Sheets web site](#).

Check the coordinate system of each data file in ArcMap:

1. **Right-click** on the data layer and choose **Properties**
2. Click on the **Source** tab

Scroll as needed to see the *coordinate system* information (in raster data sets, this is referred to as the *Spatial Reference*) – see the graphic examples below.



Note that in the example above, the two data sets are in two different coordinate systems. Before doing any analysis with these two data sets, you need to make sure they are in a common coordinate system. **Important:** note the **unit** of the coordinate system as well as the **name!** In both examples, the unit is meters.

**Important:** If the Source tab says your coordinate system is *UNDEFINED*, you must define the data set's coordinate system before proceeding. This requires that you know what the coordinate system of that data set is. Every data set is in some coordinate system. The metadata should tell you which one. Or you may have to search the web site of the data provider, write an e-mail to the data provider, phone, etc. Once you know what the coordinate system is, you should use the **Define Projection** tool in *ArcToolbox* – please see the *Tip Sheets* regarding *Map Projections and Coordinate Systems* on the [Tufts GIS Center Tips Sheets web site](#) for instructions on how to do this. **Very important:** You **cannot** simply define the data set's coordinate system as being the coordinate system you *want* it to be in, you have to define it as the coordinate system it is actually in, and to know that, you need metadata or confirmation from the data provider.

#### What target coordinate system should you use for your analysis?

There are two general guidelines for what coordinate system to use for your data:

- First, you should almost always be using a *projected* coordinate system with units either in meters or feet rather than a *geographic* coordinate system (GCS) in which the units are latitude and longitude.
- Second, if most of your data sets are already in a coordinate system that meets your needs (e.g., Mass State Plane, Mainland, NAD 83 – meters), then leave those as they are, and project any other data sets not in that coordinate system to the more common one.

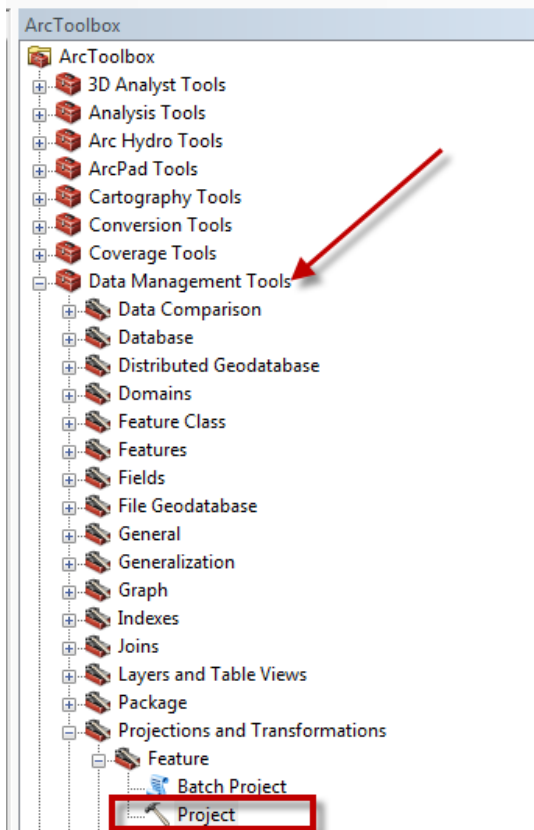
#### How to create a data set in the desired coordinate system:

OK, let's say all your data sets now have a defined coordinate system, but they have different coordinate systems. You have decided which target coordinate system you want to use, and which data sets need to be projected into that coordinate system. There are several ways to do this. But one critical rule....

**NEVER EVER EVER** use the *Define Projection* tool to *project* a data set into the *desired* coordinate system for your analysis. The *Define Projection* tool is for defining the *existing* coordinate system of the data set, not for projecting it into a *new desired* coordinate system.

Here is one simple way to project your data set into a *new desired* coordinate system:

Go to **Data Management Tools** in the *ArcToolbox*. Then click on **Projections and Transformations – Project**



Note there are two different **Project** tools – one under Features (for vector data ) and one under Raster for raster data.

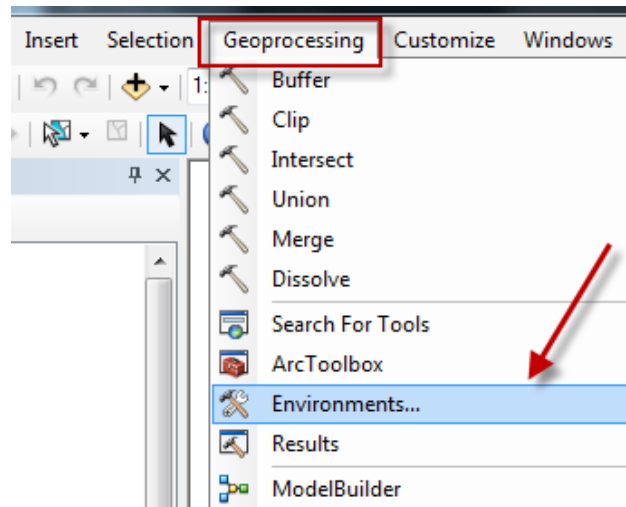
Note on the above process: you can create subset versions of larger data sets and project them into a desired coordinate system in the same step if you first select the features you want to be in the subset data set, then do the Data- Export Data function. For raster data, you make a subset data set not by selecting but by zooming into the area you want. For detailed instructions, see our *Tip Sheets* on the [Tufts GIS Center Tip Sheets web site](#) for creating subset vector and raster data sets under the *Creating Data* section.

## Specifying Geoprocessing Environment Settings

**(Important! This is required if you are using raster analysis – Spatial Analyst)**

For many analysis functions, and for ALL RASTER analysis functions, it is important to set the *Geoprocessing Environment Settings*

- Go to Geoprocessing – Environments... to open the Environment Settings box:



The critical items for you to set are:

- a *Workspace* (where files are going to be placed)
  - an *Extent* (the largest extent of your analysis area – e.g., Massachusetts)
  - a *Mask* – this is optional – to use an example, if you set both your extent and mask to a layer that consists of an outline boundary polygon for Massachusetts, the rectangular extent of your analysis would be a rectangle that encompassed Massachusetts, and any raster grid cells within that rectangle but outside the boundary polygon would be set to NO Data for a value, and would NOT be used by analysis functions
  - Note that if all your data layers are in the same coordinate system, you can leave the setting for coordinate system to the default value which is “same as input”
  - Cell size – for raster data – all your raster data should have the same cell size
1. Under the *Workspace*:
    - a. Set the *current workspace* to be where you are pulling data from (just click on a folder, not a file - the folder within which most of your GIS data is located)
    - b. Set the *scratch workspace* to a temp folder on your local computer or your H: drive that you have write access to. If you don't have write access, this will cause errors. Create a temp folder on the local computer under your user name if you have to.
  2. Set the *Output Coordinate System* to be your desired coordinate system
  3. Processing extent – this should be the largest extent of your study area – use an existing file that covers the study area or create a new one. This is critical for raster analysis – you must use the same extent for all your raster work the Extent will always be rectangular even if you use an irregular polygon area to set it
    - a. If you will be doing raster-based spatial analysis, it is good to set the *Snap Raster* under *Processing Extent* to an existing raster data layer that has the extent and cell size you want for your analysis. The extent that you set will then be snapped to this raster to ensure that cells align properly. This is often not available if you have set the extent. Presumably it snaps to the extent.

4. Under the *Raster Analysis* area:
  - a. Set the cell size (use this for all your raster analysis, so think about the cell size carefully)
  - b. Set a mask if you want (a polygon layer that represents the study area – the area outside this layer will not be included in the analysis – if you use an irregular polygon layer to set this, the mask will reflect that).