<u>CCPI Methodology and Advice</u>
By: Scott Gibson

This guide is written as a document that lays out the steps necessary for creating the landcover layers used in the CCPI grant to rank watersheds in the St. Joseph watershed. It is written in a very informal manner and is not intended to be of scientific publication quality, but instead to provide the reader with user-friendly instructions and advice. Although I have some familiarity with ArcGIS 9.X, I try to be as explicit as possible. As a result, there are probably times when I go into too much detail on simple tasks – this is not to insult anyone's intelligence, but simply to make sure that people with even a moderate amount of GIS experience can use it. In addition, I fully admit that there may be better methods for doing some of the tasks outlined. If this is the case, by all means, use them. This guide is in no way all inclusive or exhaustive. If I have omitted something, made an error, or if you simply have questions please feel free to contact me: gibson_sc@hotmail.com

<u>Data Acquisition</u>

Before you begin, you will need to acquire three different GIS datasets: 2001 National Landcover Dataset (NLCD), National Hydrography Dataset high resolution stream layers, HUC sub-watershed boundaries

<u>2001 NLCD</u> – This raster layer will provide you with all of the land cover data. This data is put out through the Multi-Resolution Land Characteristics Consortium (MRLC) - http://www.mrlc.gov/
Once at the site you should be able to download the data you need by going to the Access Data tab (It should be noted that at the time of this writing, the 2001 NLCD is the most current land cover available. However, the 2006 NLCD is supposed to come online sometime in late 2008 or 2009). There are two different ways to download data: by zone or using a Seemless Data Download. I strongly recommend that you use the Seemless Data Download tool, as grabbing them by zone will likely lead to gaps or require some nifty work to merge the grids into a single file later. Once you select Seemless Data Download, you need to select the MRLC Consortium Viewer. This will give you a map of the U.S. in which you can use the tools on the left-hand side of the viewer to zoom in to the approximate area you want to grab. Once you have the area in the view, you can draw a rectangle around the area you want to download data for using the 'define' tool under the 'Download' section. Because it will be easier in the long run to work with one landcover grid that covers the entire area you wish to work with, I strongly suggest that you grab everything you need in one download (for example if you are going to be working across the entire state of Indiana, make sure to grab the entire state in one hit). Once you grab the landcover grid, you will probably notice that the grid file does not have the names for the habitats embedded in the file. Instead you will likely only have fields in the table for the habitat "value" and the count of how many cells have that value. I recommend you add a field and fill in the table with the correct habitat names. This information can be found the website listed above.

NHD - The national hydrography data set is run by the USGS and provides a good set of stream data (among other things) for most of the U.S. (including all of Indiana and the Midwest). The national hydrography dataset streams layers were chosen for this project over the tiger-line data stream files for two reasons: 1) a visual inspection of the two layers overlaid on an aerial photo suggested that the NHD high resolution layer covered more of the smaller open waterways than the tiger data (this was qualitatively assessed) and 2) tiger line files include water bodies such as ponds and lakes (which is not only problematic when creating buffers, but these features are not something we want to include in the analysis), while the NHD files only show ditches, streams, and rivers. The NHD website is:
http://nhd.usgs.gov/
Once at the website, data can be downloaded by clicking on the 'data' tab and navigating to the NHD viewer. From this point, the navigation is similar to the Seemless Data Download tool explained with the landcover data. There are instructions on how to extract the data under the 'extract functions' area on the left. Make sure to extract 'NHD High Res' shapefiles. At the time of this writing there were several ways to extract data, but the only way to get High Res shapefiles was to extract them by 'subbasin' or '109th Cong. Dists.' At one time you could extract by county, but that has been disabled. The Subbasins correspond to the HUC 8 boundaries and are logical choices for extracting the data. Once you have the data you will notice that there are lots of different features that come with the file. The shapefile that shows the streams is called 'NHDFlowline.'

Sub-watersheds – For the CCPI grant we used HUC 11s and HUC 14s (for more info on Hydrologic Unit Codes please refer to http://water.usgs.gov/GIS/huc.html). However, there seems to be a shift to using HUC 12s by many agencies. HUC 12s often reflect the identical boundaries of HUC 14s but occasionally merge two (or portions of two) HUC 14s together. If you are working with areas in Indiana, HUC 11 and HUC 14 data can be found at the Indiana GIS Data Atlas -
http://129.79.145.7/arcims/statewide_mxd/index.html
The task of finding these in other states is challenging. At the time of this writing, Ohio data could be found here: http://www.oh.nrcs.usda.gov/technical/14-digit/download.html
Additionally, HUC 11 and 14 for *some* of the Michigan watersheds could be found here:
http://www.michigan.gov/cgi
It should be noted, however, that the HUC boundaries you get from different states may not line up exactly with each other (they may be very close but might have slight gaps between borders and lines). Because of this, if you will be doing a multistate analysis, you may want to consider using HUC 14s which can be obtained from the NRCS/USDA Geospatial Data Gateway - http://datagateway.nrcs.usda.gov/GatewayHome.html

Data Manipulation

Projecting the data - Depending on where you sourced the data, you will likely have to make sure that the data is all projected the same way (for Indiana I recommend UTM NAD 1983 Zone 16).

- The 2001 NLCD comes projected in Albers conical equal area. To project this layer in ArcGIS 9.X you can do it in ArcMap or ArcCatalog. In either case, you need to open ArcToolbox (click on the icon that looks like a red toolbox). Once there, select 'Data Management Tools', 'Projections and Transformations', 'Raster,' and 'Project.' This will bring up the 'Project' dialog box. Simply select the layer you wish to project (the input coordinate system should automatically fill in – if not you will need to select it), tell it where you want it and give it a name. NOTE: When you project a raster file such as the land cover layer, there is some shifting that occurs because of the resampling technique used by the program (you can read about the different resampling techniques in the project dialog box and there is also quite a bit of info online if interested). Simply stated, ArcGIS takes the grid and twists it to a different projection so you will not wind up with exactly the same number of grid cells classified as a certain habitat in the output grid as the original (however, the habitats should remain in the same location – e.g. if the original grid has one cell classified as deciduous forest at point X, it may have two cells classified as deciduous forest in the output grid, but they will still be centered on point X). To see what happens, all you need to do is compare your output grid with the original. To my knowledge there is no way around this, but such discrepancies are minor.

-The procedure for projecting the NHD and HUC layers are similar to the NLCD data except that they are not raster files. You still use ArcTool box, except you select 'feature' under 'Projections and Transformations' instead of 'Raster.' Once you do that, the instructions are similar. The NHD data is not projected when you download it, so the input coordinate system should be GCS North American 1983. The HUC data will likely vary in what projection it comes in based on where you got it (the files from the Indiana GIS Data Atlas should already be in UTM83 Zone 16).

The overall goal of this project is to get two kinds of data: **1)** landcover within the sub-watershed boundaries, and **2)** landcover within 30m of the streams within sub-watershed boundaries. I will now provide detailed steps on how to get there.

Assuming that you have all of the data you will need downloaded and projected, the next step you will want to take is to make individual shapefiles for each of the sub-watershed boundaries you want to work with. In general when you download the HUC data you will simply get one shapefile that has sub-watersheds as different features of the same shapefile (e.g. the whole shapefile might be all of Indian broken into HUC 14s) and we want to split these apart. As with most things in GIS there are many ways to do this (some quicker than others), but the easiest seems to be to simply highlight the respective sub-watersheds and export them as individual files. To do this, you can either highlight the targeted HUC in the table or on the screen using the 'select features' tool. It should turn blue when selected. You then right click on the layer in the table of contents and select 'Data,' 'Export Data' and make sure 'Selected Features' is selected in the 'Export' box and pick a file location and name. This will save whatever the highlighted feature was as its own shapefile. You then proceed to do that for all of the sub-watersheds that you will be working with. As I mentioned, there are other ways to do this, and some may be faster if you have many to do, so don't feel obligated to follow this method – the goal is simply to have individual shapefiles for each subwatershed you want to work with.

<u>Landcover within the HUC boundaries</u> – it really doesn't matter what level HUCs you use (e.g HUC 11, HUC 12, etc.), the steps are the same regardless. There are two ways to go about doing this: 1) keeping the landcover as a grid file or 2) converting it to a shapefile first. I recommend the second method (converting it to a shapefile first) as you will need to do this for the 30 m buffer landcover anyway and it also prevents some shifting that occurs when you 'clip' a grid. I will, however, explain both methods (I recommend you try both to see some of the differences):

> -<u>Keeping the NLCD data as a grid</u>: First make sure that you have the HUC watersheds you want to work with in ArcMap as well as the landcover grid. You are merely going to clip the grid to the shape of each HUC boundary. However, unlike shapefiles, you can not simply use the 'clip' function – you have to go through Spatial Analyst and use the raster calculator. To do this, Spatial Analyst must be turned on – go to the 'Tools' menu, select 'Extension' and make sure the box next to 'Spatial Analyst' is checked. This should bring the Spatial Analyst box into your view. To "clip" a grid to the shape of a HUC, you must first tell spatial analyst what you will use as a mask as well as a few other parameters. To do this, click on the Spatial Analyst drop down list that is now at the top of your view and go to 'Options.' This will bring up the Options dialog box. First go to the 'General' tab. In the 'Analysis Mask' box you want to select the HUC boundary that you want to "clip" the grid to (i.e. the sub-watershed you want to cut the landcover grid into the shape of) . You shouldn't have to worry about the other boxes. Next go to the 'Extent' tab and set the Analysis extent to be the same as whatever HUC polygon you used as a mask. Last, go to the Cell Size tab and set it so it is the same as the grid you are using (this should be 30 x 30 m cells). You are now ready to make the "clip." To do this, go back to the Spatial Analyst drop down menu and select 'Raster Calculator.' This will bring up a dialog box that looks like a calculator or one used for queries. In the 'Layers' list you should see the grid you want to clip. Double click the grid and it should show up in the box at the bottom. All you need to do now is hit 'Evaluate.' This will "clip" the grid to the shape of the mask you plugged into the options dialog box. However, the output grid you create is not actually saved anywhere and will likely show up in your view as "calculation." It is temporary and will need to be converted to a grid. Before you do this though, I recommend that you do a join with the original grid so that the new grid you are making will have the landcover names in the table (you can also do this after the grid has been made permanent). To do a join, right click on the "calculation" layer and go to 'Joins and Relates' and select join. When the table comes up, select to join based on 'value' and select the original grid with the landcover/habitat names already in the table as the layer to join to. Lastly, make sure you select 'value' as the field to base the join on and click OK. The table of your temporary output grid should now have the fields from the original grid. Unfortunately, it also brings over the "count" field

which tells you how many grid cells the original grid or landcover layer had. You can delete this after you have made the grid permanent. To make it a permanent grid, I recommend exporting it (though you can also use the "make permanent function" but I have had some issues with that). Right click on the temporary "calculation" grid and go to 'Data', 'Export Grid,' and you will get a big dialog box. You should only have to worry about the bit at the bottom: select a location where you want it saved, pick the 'Format' as "grid" and give it a name (there are some naming restrictions for grids, so don't be surprised it gets rejected if your name is too long). Hit save. The new grid should now be in the table of contents. At this point you can simply delete the temporary "calculation" grid. Once you have done that, I recommend going into the table for the new grid and removing the fields that are not applicable so you don't get confused later: specifically you will want to get rid of the field that gives you the cell count for the old grid (likely named "count_1"). At this point you should have all the info you need to calculate how much of each landcover type is found within the border. To do this you will need to export the table: with the attribute table open go to 'Export' and work through the dialog box (it will save as a .dbf file). You can then open the table in excel and manipulate the data to calculate the area of each landcover type. Since each grid cell is 30 x 30 meters we know that each cell represents 900 square meters. By simply multiplying the 'count' of each landcover type by 900 you can calculate the total number of square meters of the respective habitat types within the HUC unit. From there you can convert to other units of area measurement if need be.

NOTE: when you clip the landcover grid using the method above (i.e. keeping it as a grid) you will notice that the grid cannot be clipped to contour along the exact boundary of the watershed. This is because the grid cells are square and the program will delete ones with less than 50% in the watershed and vice-versa. As a result the edges will be jagged. Although converting to a shapefile first and then clipping will eliminate this, I have found the difference is negligible (usually much less than 0.1% difference). Additionally, whenever you "clip" a grid using the above method some minor shifting occurs. Specifically, the resulting output gird will often be shifted ~5 m in any direction. The amount and direction of the shift seems to be dependent on the size of the mask applied. The grid cells themselves do not change size, and the landcover type stays the same – they simply shift. This is not important for calculating the total area of each landcover type as all the grid cells shift evenly – in other words you still have the same number of cells, they just shift. This is, however, important if you were to then use this output grid again for some other calculation. As a result you should ALWAYS use the original grid for other calculations if you choose to use this method. This is another reason that you will probably want to convert to a shapefile first – there is no shifting involved.

-Converting to a shapefile first (recommended) – As stated above, there are several reasons that this technique seems to be better (can clip to the exact shape of the sub-watershed, no shifting, will need it for later anyway). The first step in this process is to convert the grid to a shapefile. Follow the directions above for turning on the Spatial Analyst extension if it is not on already. Once it is on, click on the Spatial Analyst drop-down menu and select 'Convert.' You will want to select 'Raster to Features.' This will give you a dialog box. First make sure that you select the correct input raster (the grid you want to convet). For the field, I recommend having it set at "value" which is the landcover code given to it by the MRLC – whatever you select here will be retained in the output shapefile and the most of the other fields deleted (assuming you have a field with the landcover names, you could select that field – however, I have found that doing makes the conversion run MUCH slower). Leave the output geometry as 'polygon' and uncheck the 'Generalize lines' box. If you don't, the output shapefile will not be an exact copy of the grid (it will instead try to round out the borders). Once you select the save location and name the shapefile, click OK. Depending on how big the grid is you are converting and how fast your computer is, this may take awhile. Once the conversion is completed, you will get your shapefile (and it should be an exact copy of your grid – no shifting). If you now open the table to your new landcover shapefile, you will notice that you now have many parts or features where before your grid had only a handful (one for each landcover type). This is because the shapefile has made an individual feature for each cluster of cells that share the same value or landcover type. Although you can later composite them when you export tables to excel, I recommend you regroup them in your shapefile. There is no "fast" way to do this that I know of. I usually start by running a dissolve (go to ArcToolbox >Data Management Tools>Generalization>Dissolve). A dissolve often groups many of these together for you, but there are often quite a few that remain separate. From here I generally to go in and manually select features in the table that share the same landcover type (or value) and do a 'merge.' To do this, you will first need to edit the layer - go to the 'Editor' drop down menu and select "Start Editing." Make sure you select the right folder and that the right layer is selected in the 'Target' box at the top of your view. Next, open the attribute table of the layer you want to merge and select all the features that share the attribute you want to merge (I usually do a sort first so they are sequential and easier to select). Then simply go to the 'Editor' drop down menu and select "Merge." This will bring up a dialog box with all the features you have selected and ask you which feature you want to merge the others with. It really doesn't matter which one you select – pick one and click OK. This will merge those features into one. Repeat this step for the other landcover classes (or values). When you are done, you should have a shapefile with the same number of features as the grid (i.e. one feature per landcover class).

The next step is to clip the landcover layer to the respective sub-watershed boundaries. Make sure you have the HUC boundaries you want to work with in the table of contents along with the new landcover shapefile. Go to ArcToolbox and 'Analysis Tools.' Under the 'Extract' heading you should see 'Clip.' Choose this. Fill in the dialog box: the input feature will be the landcover grid, the clip feature will be whatever sub-watershed you want to clip the landcover shapefile

to, and the output feature class is simply what you want to name the output shapefile. The resulting shapefile should be a clip of the landcover shapefile in the exact shape of the sub-watershed. However, in contrast to "clipping" a grid in the first method I wrote about, there is nothing in the attribute table to tell you the area of each landcover class. To do this you will need to use an extension. If you have access to ArcView 3.2 I recommend bringing the shapefile into this program and using X-Tools to calculate the area (X-Tools for 3.2 can be downloaded from various sites
e.g. http://www.odf.state.or.us/divisions/management/state_forests/XTools.asp.
The alternative extension for 9.X (which I don't like as well) is Hawth's Tools. You can get it here: http://www.spatialecology.com/htools/tooldesc.php
Once you calculate the areas, you can export the tables in the manner mentioned in the first technique and open them in Excel.

You should now be able to complete the first part of the analysis: calculating the amount of each landcover type found in each sub-watershed.

Landcover within the 30m buffers – This portion involves first creating buffers around the streams and then calculating the amount of each landcover type that lie within the buffers.

The first step is to make sure you have your stream layers in order. If you downloaded the NHD data by sub-basin it is likely that you will have all of the stream data for an entire watershed. However, if you are going to be working in multiple watershed or don't have the coverage from one stream shapefile, I would recommend merging as many of the NHD layers together to cover what you need before you start anything else. You can do this with the merge tool in ArcToolbox (ArcToolbox > Data Management Tools > General > Merge). This is a different "merge" than I wrote about where multiple features of a single shapefile were merged into one. This type of merge takes two different shapefiles and merges them in to one. Before you do this, make sure you don't have any overlapping streams between the two (or more) stream layers – if you do, they will both wind up in the merged layer stacked on top of one another so you can't see them. This can create problems for later. I don't think this should happen if you download the NHD files by sub-basin, but it used to be a common occurrence when you could download them by county.

Once your stream layers are ready, you now need to make the buffers around the streams. There are multiple ways to make buffers in ArcGIS, but I prefer using the tool housed in ArcToolbox (ArcToolbox > Analysis Tools > Proximity > Buffer). Select the input feature (stream layer) and name the output shapefile. Since we are using 30 m buffers, set the distance as 30 and make sure the units are meters. You will want the 'side type' to be "full" and the 'end type' to be "round." The 'dissolve type' field dictates whether all of the resulting buffers features will be merged together (the program creates a buffer around each line *segment*). You will ultimately want the buffers to be merged, but I have found turning this on for the buffer creation causes the program to take MUCH longer to execute the buffer, so I leave it off and merge them later. Click OK and you should have your stream buffers. However, at this stage unless you said yes to dissolving them during their creation, you will notice that the layer is made up of many overlapping buffers. At this stage I usually merge them all into one feature. To do this, simply select them all in the attribute table and merge them using the technique previous discussed to merge the landcover features in the landcover shapefile (i.e. select them, edit them, use the merge feature under the editor drop down menu). This should condense all of the mulitpart buffers into one feature. The final step is to take the large buffer shapefile and clip it to the respective sub-watersheds that you will be working with (so that each sub-watershed has its own 30m buffered stream layer). To do this you can follow the steps outlined above for clipping the landcover shapefile to the shape of the sub-watersheds (except instead of using the landcover shapefile as the layer to be clipped you will use the buffer layer).

You should now be ready to clip the landcover shapefile to the buffers. If you haven't already made your landcover grid into a shapefile, you will need to do that before moving on. Please refer to the section on how to do this above. If you try to directly "clip" the landcover grid without converting it to a shapefile, it will not contour along the stream buffers as it will discard grid cells with less than 50% inside the buffer. Although this was not important for assessing landcover over an entire sub-watershed (which is thousands of acres) it would become a problem with the considerably smaller areas of the stream buffers. Once you have your landcover shapefile and the buffers for each sub-watershed in your view, you can begin to clip the landcover shapefile to the buffers. To do this, you simply need to use the same procedure you have already used with other shapefiles (ArcToolbox > Analysis Tools > Extract > Clip). Once you have you resulting shapefiles, you can calculate area as discussed previously and export the tables.

At this point, you should be done!

A note about proofing the outputs

With all the steps and data manipulation, it is easy to make mistakes. This is compounded by the fact that ArcGIS is a fickle program to begin with. You will undoubtedly encounter problems that I have not outlined above, and some that will seemingly have no explanation. Because of all this, I strongly recommend that you proof and double-check every layer that you create. For example, I usually calculate the areas of the sub-watersheds and the buffers and use that to check the areas of the landcover layers I create to make sure they match (keep in mind that they may not match exactly if you are directly clipping the grids without converting to shapefiles).