TUTORIAL

Coastal and Marine Ecological Classification Standard Crosswalk Tool

Version 1.0

July 2015

NOAA Office for Coastal Management
(843) 740-1200
www.csc.noaa.gov

Table of Contents

Coastal and Marine Ecological Classification Standard Crosswalk Tool

Executive Summary

The process of converting thematic data attributed according to one classification system into another is known as *crosswalking*. This process is useful when integrating multiple data sets within an individual project or when integrating various local projects into a regional assessment. Crosswalking can be a tedious process and is vulnerable to some subjectivity. The Coastal and Marine Ecological Classification System (CMECS) Crosswalk Tool is intended to streamline the process and improve consistency in converting between classifications, while still allowing users flexibility.

The crosswalk tool is a geographic information system (GIS)-based tool designed to operate as a toolbox (*.tbx) within ArcGIS 10.2. The tool converts vector spatial data sets from one thematic classification system into multiple layers that reflect the CMECS components. This process relies on look-up tables containing units (attributes) from the source classification system and equivalent CMECS units by component.

The downloaded zip file that contains the toolbox and this document also includes several accompanying files. These files are look-up tables for three commonly used coastal and marine classification systems, as well as sample data to demonstrate how the tool functions.

Introduction

Successfully translating attributes between two classification systems is dependent on many factors. The crosswalk process requires that the relationship between units in each system be understood in order to establish equivalency. In some cases units in the two classifications are conceptually equivalent, have comparable definitions, share common thresholds, and can be easily translated. In other cases there may be no equivalent Coastal and Marine Ecological Classification System (CMECS) unit for a source category.

Look-up tables for three widely used landscape-level coastal and marine classification systems are provided with the crosswalk tool. The Florida System for Classifying Habitats in Estuarine and Marine Environments (SCHEME)[1] integrates well with other marine and terrestrial systems and has proven to be applicable in many areas outside of Florida. The Cowardin table is based on the Classification of Wetlands and Deepwater Habitats of the United States.[2] This system was originally national in focus and is itself a Federal Geographic Data Committee standard. The national estuarine research reserve table is drawn from the National Estuarine Research Reserve Classification System (NERRCS),[3] which is also national in scope and includes shallow aquatic, terrestrial, and extensive anthropogenic units. Each system is very amenable to the CMECS crosswalk tool because of its fixed code system and hierarchical structure.

If working with a system other than the ones provided with the crosswalk tool, users are advised to first conduct a conceptual crosswalk before applying the tool. This will avoid an incorrect crosswalk where unit relationships are ambiguous or emphasize different ecological elements. Appendix H of the CMECS standard document,[4] available at the CMECS website (*www.coast.noaa.gov/digitalcoast/publications/cmecs*), contains guidance and best practices for conducting the conceptual crosswalk. Conceptual crosswalk tables that can be modified by the user are also provided in the CCT.zip file. Once the conceptual crosswalk has been accomplished, the user can review the look-up tables to modify how source attributes translate to CMECS if necessary.

Look-up tables for several common classifications will be posted on the CMECS website. In cases where two

units are not equivalent, the tool will not attempt to assign a CMECS code in the output. Rather, users can assign a code later following further analysis, leave the NULL value assigned by the tool, or substitute a code of their choosing. The output attribute tables from the tool preserve the original source attributes and codes that assist users in deciding how to label the NULL units and in evaluating the automated results.

The crosswalk tool assigns the closest CMECS unit to the source classification unit regardless of its place in the hierarchy of either system. In some cases, equivalency will be at lower levels like Biotic Community and in others it will occur at the Biotic Class level.

Workflow

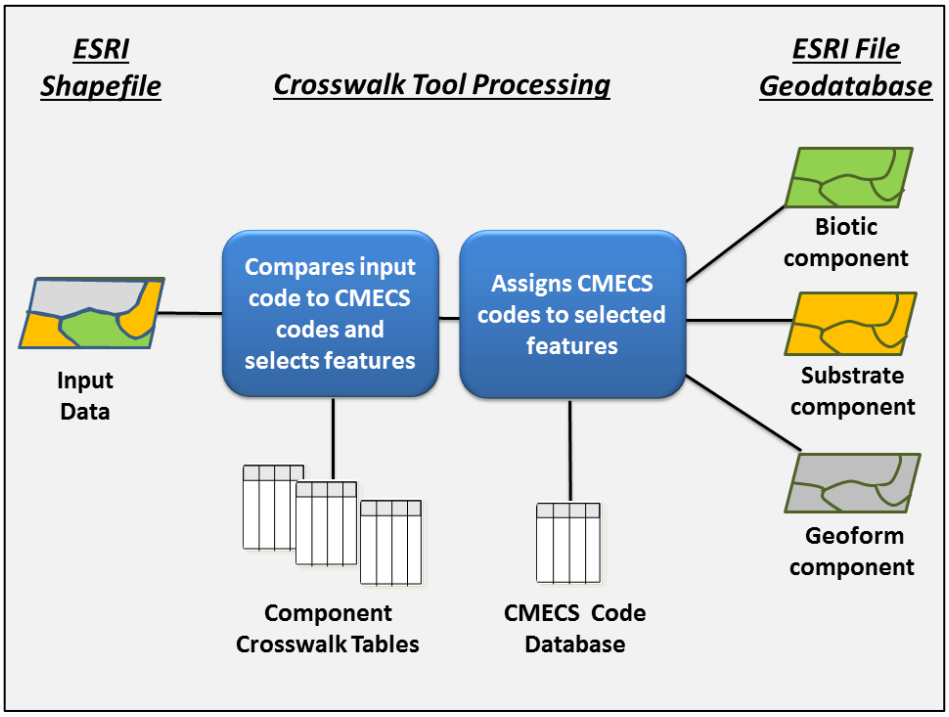

Figure 1: CMECS Crosswalk Tool model design

Figure 1 above illustrates the two-step process the tool follows to conduct the crosswalk and the multiple output data sets that result. The tool requires an Esri shapefile (*.shp) or feature class as the input data format. The attribute table for the input file must include a code field representing the alpha-numeric codes for the classification system. These codes are referenced in the look-up tables that facilitate the crosswalk process and allow the translation to CMECS. Additional fields in the source attribute table will not affect the tool's operation, but the alpha-numeric code is required. For data in this field to be compatible with the crosswalk tool, it must be copied into a new field called InputCode.

The tool creates a crosswalk between the existing classes and the CMECS classes. As mentioned in the previous section, the tool comes with look-up tables for three classification systems (SCHEME, NERRCS and Cowardin). These tables are essential for translating a benthic data set to CMECS. Each component look-up table (or LUT) contains two fields: 1) InputCode and 2) CMECS_Code. InputCode contains all the possible alpha-numeric codes for a particular classification system while CMECS_Code contains the CMECS equivalent for that particular component. These tables are provided in DBASE format (.dbf).

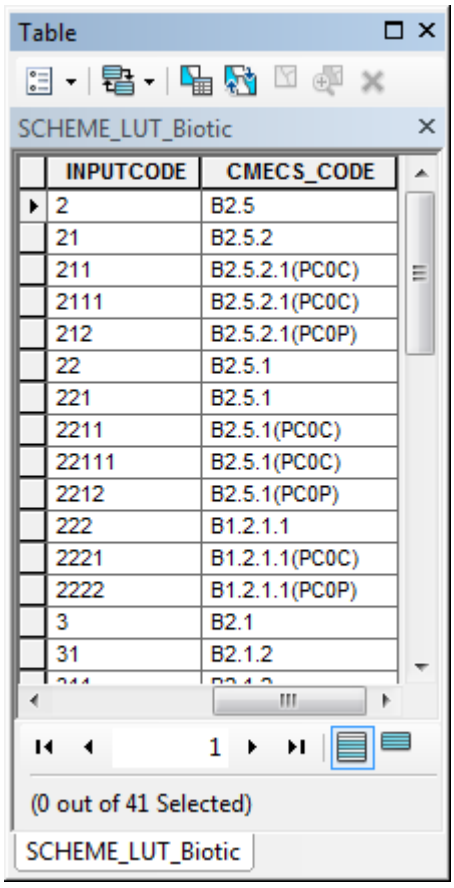

Figure 2: Example of SCHEME biotic look-up table, or LUT

Following the conceptual crosswalk recommended in the "Introduction" section, a review of the look-up tables is advisable. The tables can be edited in their native format in ArcMap or modified as Microsoft Excel files and then converted to *.dbf format using ArcMap. Figure 2 above illustrates the relationship between input source data codes and their CMECS equivalents.

One of the strengths of the CMECS system is that different aspects of the environment are organized into individual components (water column, substrate, biota, and geoform). The data are outputted into three component class shapefiles (substrate, biota, and geoform). The tool does not yet accommodate crosswalks using the water column component.

Instructions for Installing the Crosswalk Tool

The CMECS Crosswalk Tool is compatible with ArcGIS 10.2 with Service Pack 4 or higher and requires hardware similar to that recommended for proper operation of ArcGIS.

To install the tool, create a local directory called **C:\CMECS_CrosswalkTool** and extract the contents of the zip file to this folder.

Content of CMECS Crosswalk Tool (CCT) Download Package

The following files are provided in the downloaded CCT.zip file:

Table 1: Files contained within CCT.zip

File	Description
CCT.tbx	CMECS Crosswalk Tool
SCHEME_LUT_Biotic.dbf	LUT for the SCHEME to CMECS Biotic Component
SCHEME_LUT_Substrate.dbf	LUT for the SCHEME to CMECS Substrate Component
SCHEME_LUT_Geoform.dbf	LUT for the SCHEME to CMECS Geoform Component
Cowardin_LUT_Biotic.dbf	LUT for the Cowardin to CMECS Biotic Component
Cowardin_LUT_Substrate.dbf	LUT for the Cowardin to CMECS Substrate Component
Cowardin_LUT_Geoform.dbf	LUT for the Cowardin to CMECS Geoform Component
NERR_LUT_Biotic.dbf	LUT for the NERR to CMECS Biotic Component
NERR_LUT_Substrate.dbf	LUT for the NERR to CMECS Substrate Component
NERR_LUT_Geoform.dbf	LUT for the NERR to CMECS Geoform Component
Biotic_CMECS_Codes.dbf	Alpha Numeric Codes for CMECS Biotic Component
Substrate_CMECS_Codes.dbf	Alpha Numeric Codes for CMECS Substrate Component
Geofrom_CMECS_Codes.dbf	Alpha Numeric Codes for CMECS Geoform Component
humboldt_bay_scheme_subset.mxd	Esri ArcMap Map Document
ArcataBay.shp	Sample SCHEME Input Data Set

1. After starting an ArcMap session, navigate to the **mxd** folder and open the Map Document called **ArcataBayTutorialSCHEME.mxd**.
2. Open **ArcToolbox** by selecting it from the **Geoprocessing** pull-down menu.
3. Right-click inside and select **Add Toolbox**.
4. Navigate to the **tool** folder and add **CCT.tbx**. This will add the **NOAA CMECS Crosswalk Tool** to your list of available tools.

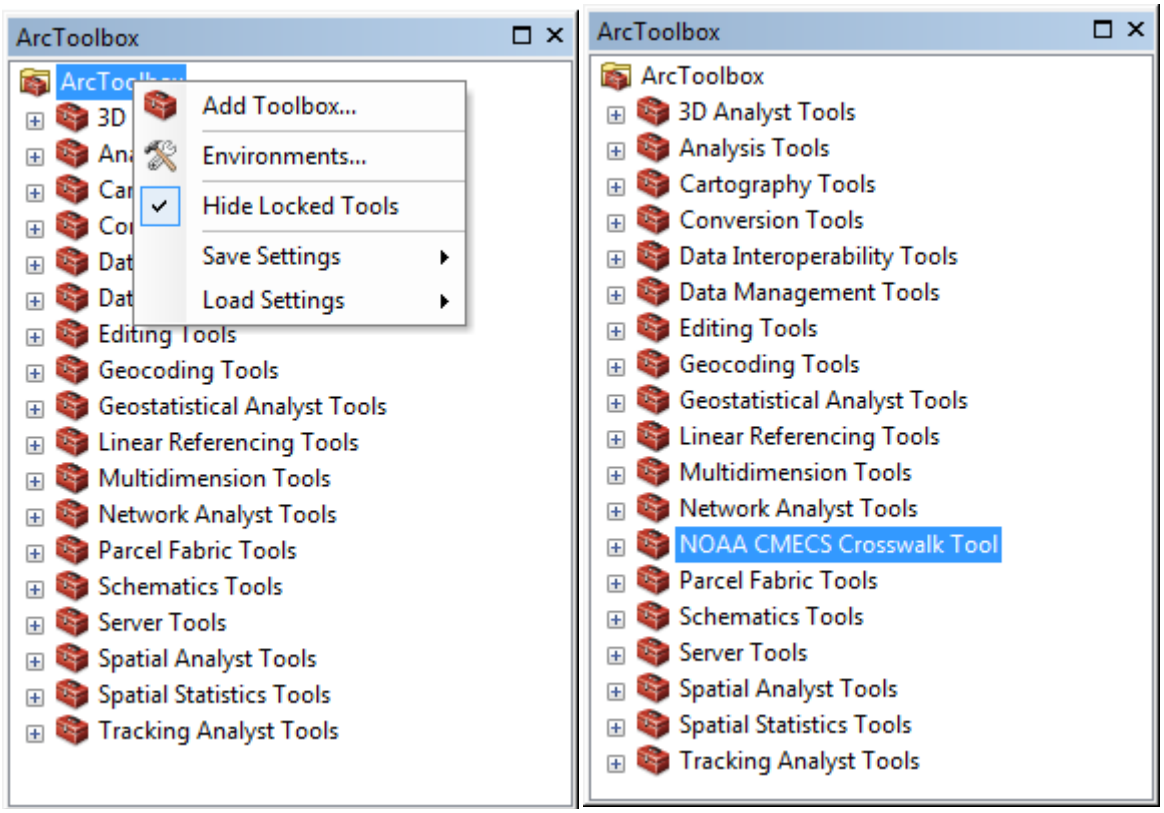

Figure 3: Adding the CMECS Crosswalk Tool to ArcToolbox

1. The Map Document will already contain a Benthic Habitat data set called **ArcataBay** classified to SCHEME (see figure 4).

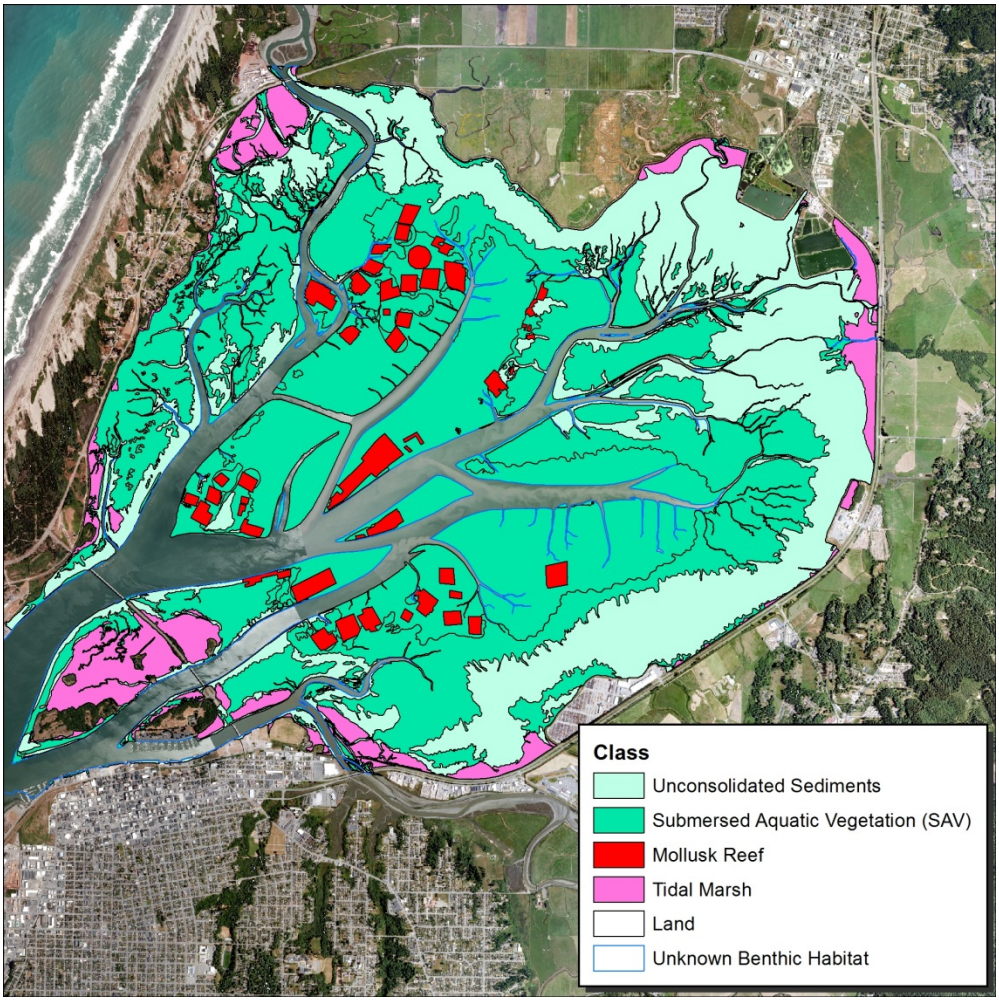

Figure 4: Input SCHEME shapefile and classes

2. Open the attribute table of the layer **ArcataBay** by right-clicking on it in the table of contents, and select **Open Attribute Table**.

3. Select **Add Field** from the Table pull-down menu.

4. Name the field **InputCode** and make the type **Text** with a length of **50**.

5. Right-click on the field **InputCode** and select **Field Calculator**.

6. Enter the Field name of your original alpha-numeric code field in the following expression **InputCode = [SCHEME_COD] (if using the VB Parser)** and click **OK**.

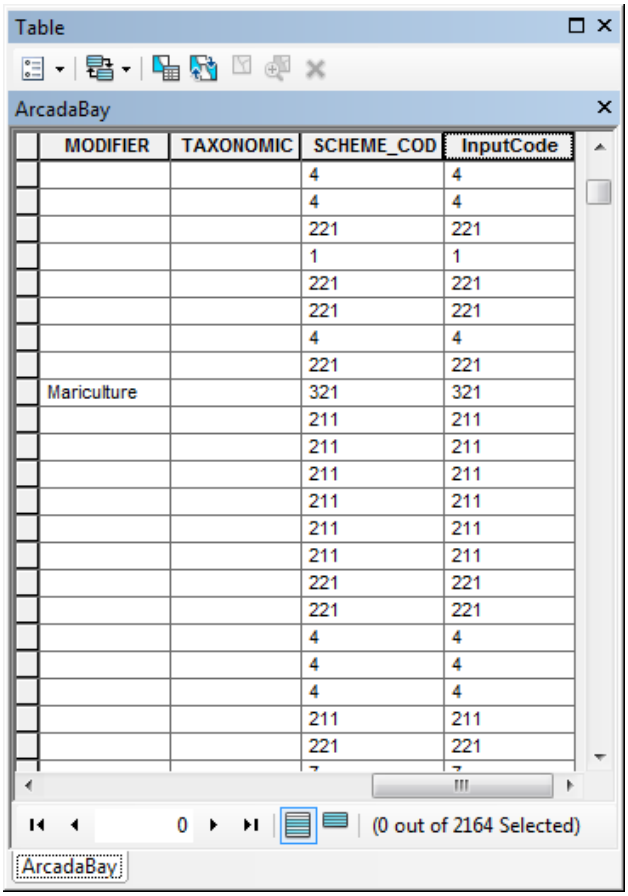

Figure 5: Attribute table for ArcataBay.shp after adding InputCode field

Applying the Crosswalk Tool

1. Open **ArcCatalog**.
2. Navigate to the directory you created when you extracted the CCT.zip file.
3. Right-click on the folder and select **New** –> **File Geodatabase**.
4. Name the File Geodatabase **CMECS_tutorial.gdb**.
5. Open the **NOAA CMECS Crosswalk Tool** from **ArcToolbox**.

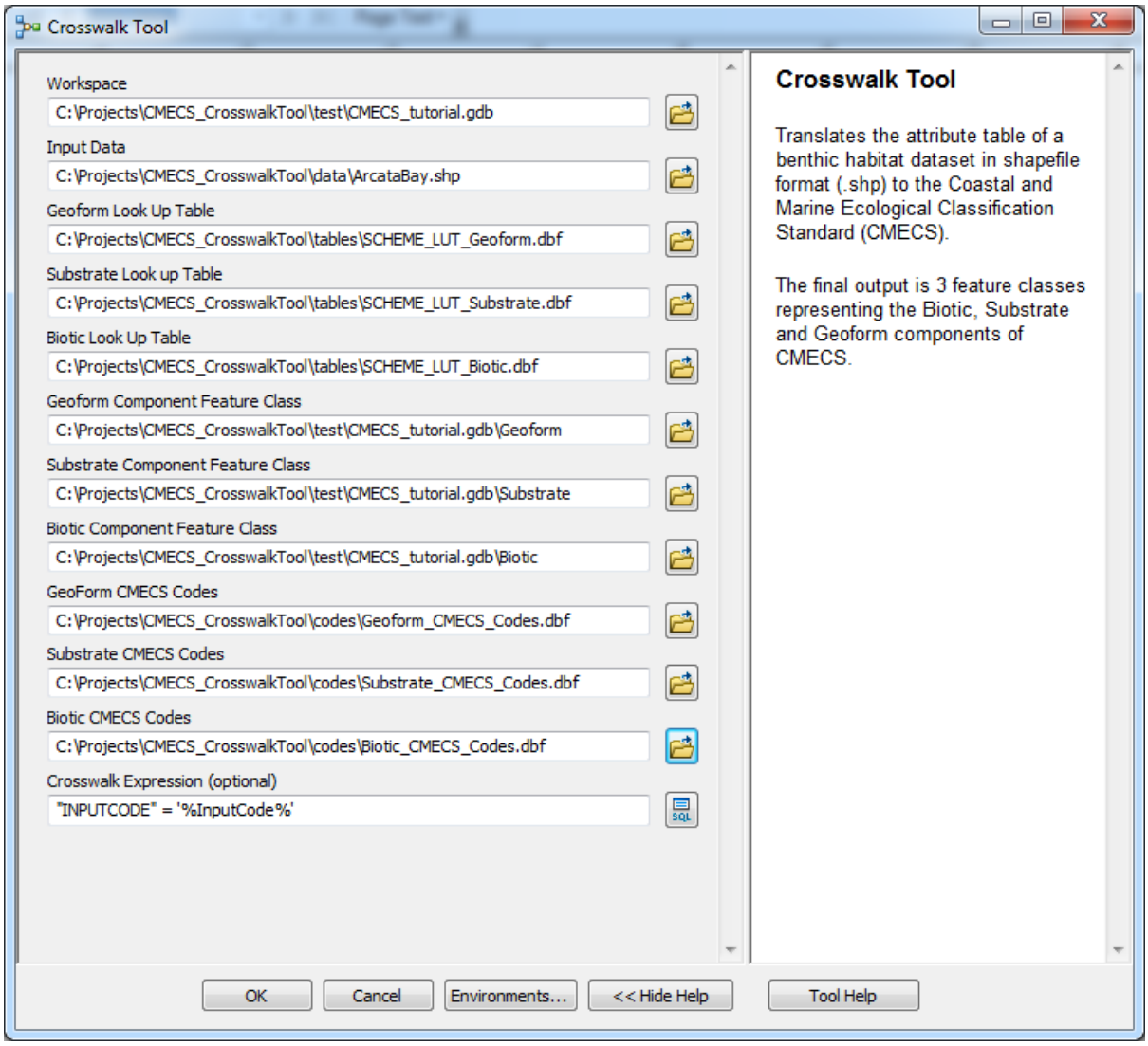

Figure 6: CMECS Crosswalk Tool user interface

6. Enter the previously created File Geodatabase **CMECS_tutorial.gdb** in the **Workspace** dialog.
7. **Input Data** is where the source data shapefile or feature class is indicated. For this tutorial, enter **ArcataBay.shp**.
8. Enter the appropriate .dbf files in the **Geoform**, **Substrate**, and **Biotic Look Up Tables**. Be sure they all apply to the same classification system. For this example, use the look-up tables associated with SCHEME, which can be found in the **Tables** folder of the CCT directory.
9. Name the feature classes for each of the three components to be output in your File Geodatabase. *Care should be taken when naming output feature layers, since there may be cases where more than one layer may be needed for substrate or biota.*
10. Enter the CMECS code .dbf files for the **Geoform**, **Substrate**, and **Biotic Components**. These can be found in the **Codes** folder of the CCT directory.
11. Copy and Paste **"INPUTCODE" = '%InputCode%'** into the **Crosswalk Expression** input. *To accommodate one-to-one and one-to-many relationships, the tool provides two different*

expressions for crosswalking data. The expression "INPUTCODE" = '%InputCode%' should be used when the codes in your data set should be an exact match with the codes in the look-up table. If your input data set contains additional modifiers that are not captured in the look-up table, the expression "INPUTCODE" LIKE '%InputCode%%' should be used.

12. Press **OK** to activate the tool.
13. When the tool is finished, the output feature layers are automatically added to the table of contents of the ArcMap session for visual inspection. Review the attributes to ensure the crosswalk was correctly applied.

Understanding the Outputs from the CMECS Crosswalk Tool

Figures 7-9 below illustrate the component outputs from applying the crosswalk tool to the SCHEME data for Arcata Bay. Transparent polygons in the output feature classes have NULL CMECS values either because the tool couldn't determine which CMECS unit should label the polygon or because there is no information available in the source data set for that polygon (such as an open water area).

For example, in figure 7 the SCHEME data contained unconsolidated sediments in the high intertidal zone. No biotic information was available for these areas in the source data, and therefore they are shown as NULL (empty) polygons below. Should biotic information become available for these areas those attributes could be added later.

Both the substrate and geoform outputs have a majority of NULL polygons. This is due primarily to the biotic focus of the SCHEME system and the fact that geomorphology is not generally within the conceptual domain of that system. In the case of the geoform output, it is likely that the user may wish to re-delineate the existing polygon boundaries along more geomorphologic lines.

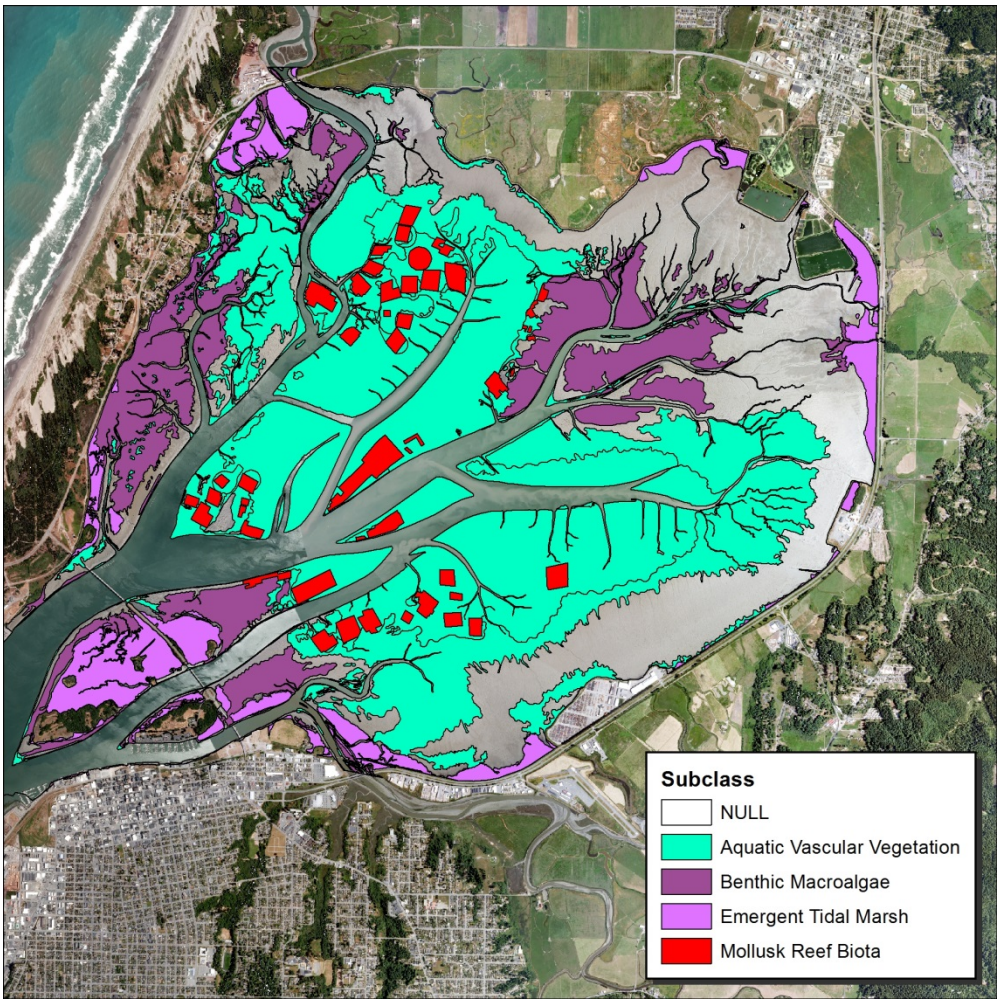

Figure 7: CMECS Biotic Component output feature class

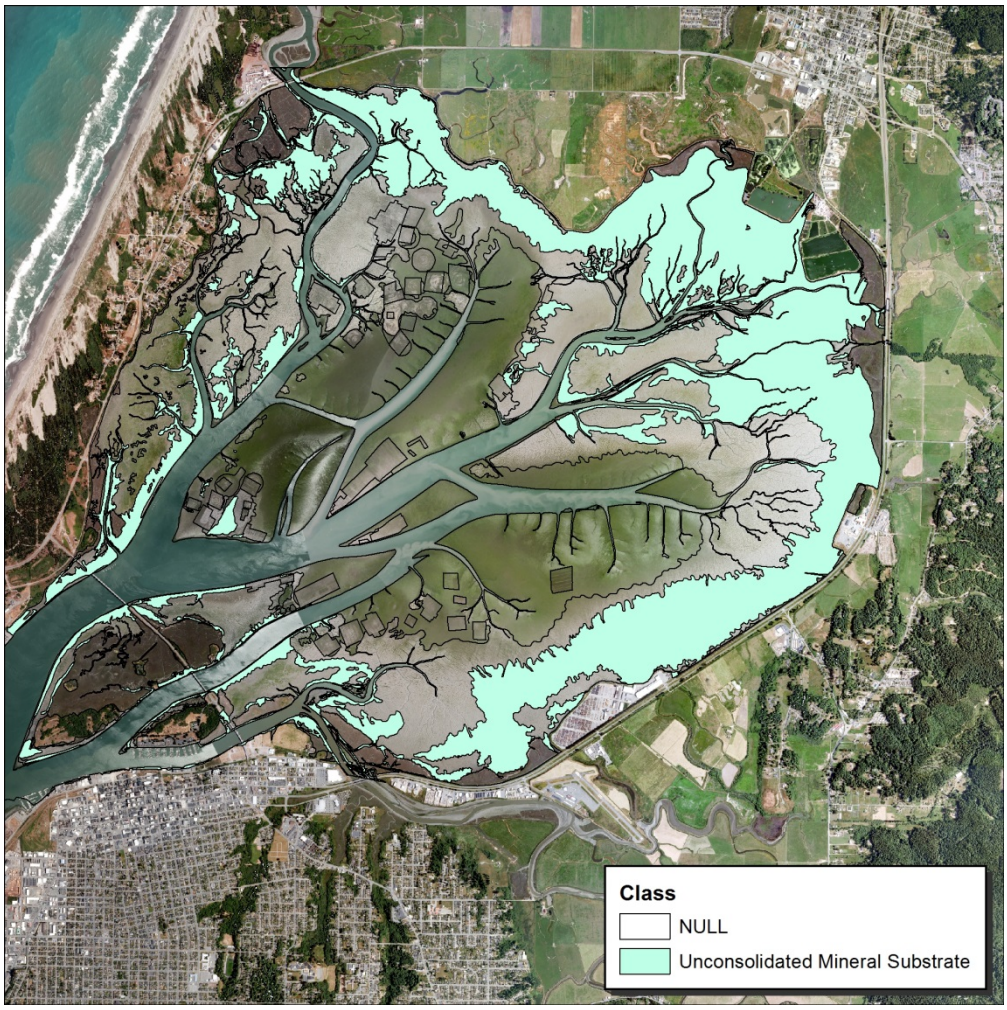

Figure 8: CMECS Substrate Component output feature class

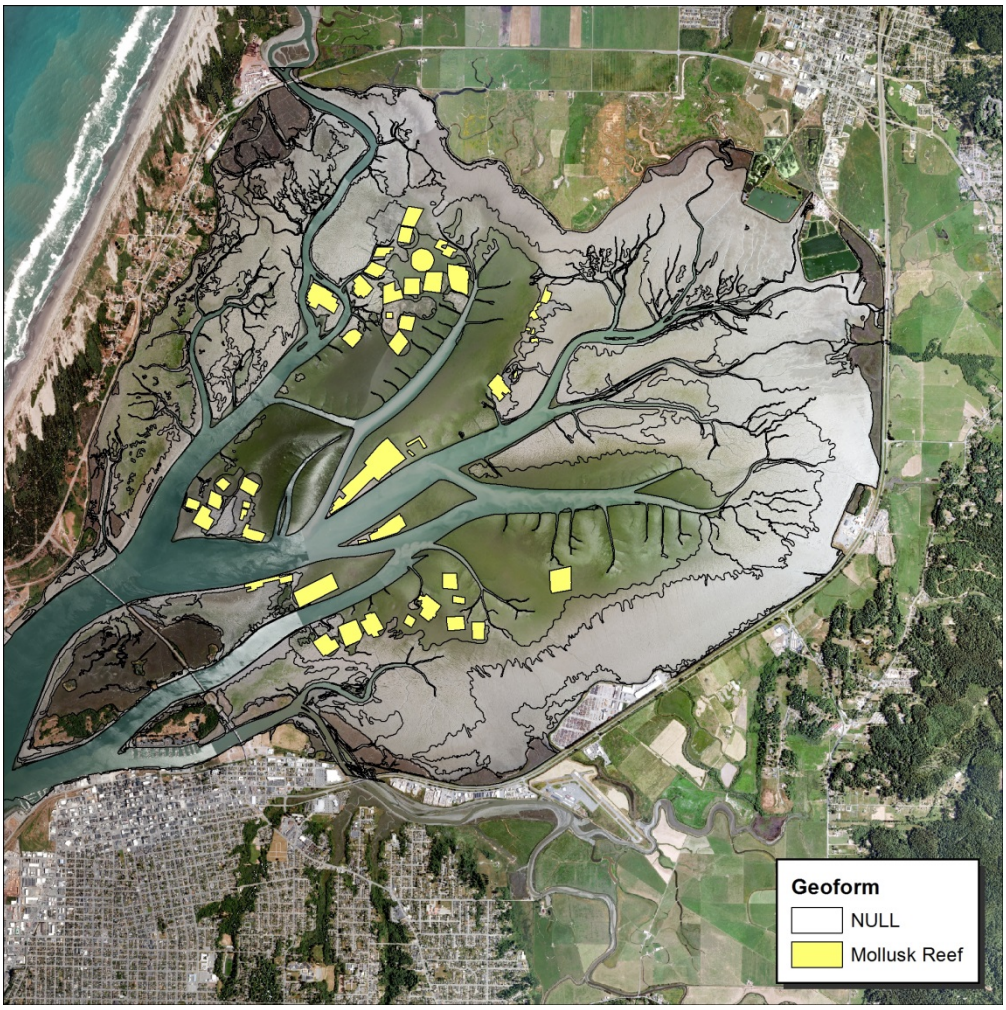

Figure 9: CMECS Geoform Component output feature class

Modifying the CCT Look-Up Tables to Crosswalk Additional Classification Systems

The look-up tables (LUT) at the heart of the CMECS crosswalk tool are designed so that they can be modified to import classification systems other than those included with the tool (SCHEME, NERRSCS, NWI). As previously mentioned, users are strongly encouraged to conduct a conceptual crosswalk prior to automating this process through the CCT tool. This requires the user to consider any information that might be generalized or lost during a crosswalk, clarify any assumptions about source data attributes and characteristics, and identify key features that might be important to highlight in the output product.

Appendix H of the CMECS standard document,[4] available at the CMECS website (*www.coast.noaa.gov/digitalcoast/publications/cmecs*), contains guidance and best practices for conducting

the conceptual crosswalk. Once the conceptual crosswalk has been accomplished, the user can create look-up tables to translate new source attributes to CMECS.

The crosswalk tool employs a two-step process to convert data, relying on two *.dbf files, but only the LUT file needs to be edited to convert a new classification system. The input source data requirements associated with the three prepackaged crosswalk tables will still apply with new systems, namely

- The data must be in vector (point, line, or polygon) ESRI shapefile format
- The data must have a coding system
- The data must have an attribute named **InputCode** (added as part of the tool prep process)

To create a custom look-up table for a new classification,

1. Open a new Microsoft Excel spreadsheet and label column A as “ID”, column B “InputCode” and column C as “CMECS_CODE”.
2. The ID column should be a numeric sequence starting at 1.
3. Enter the source code values that will be crosswalked to CMECS for that component into the InputCode column.
4. Enter the equivalent CMECS codes in the CMECS_CODE column.
5. Save and name the file in a way that indicates which CMECS component the table represents (geoform, substrate, or biotic). Repeat the process with separate Excel files for the remaining two components. NOTE: A table for each of the three components is required for the tool to run. If you only want to crosswalk biotic values you will still need tables for geoform and substrate. In that case. Create a spreadsheet with header columns but no records.

Figure 10 below illustrates examples of both populated and empty crosswalk spreadsheets.

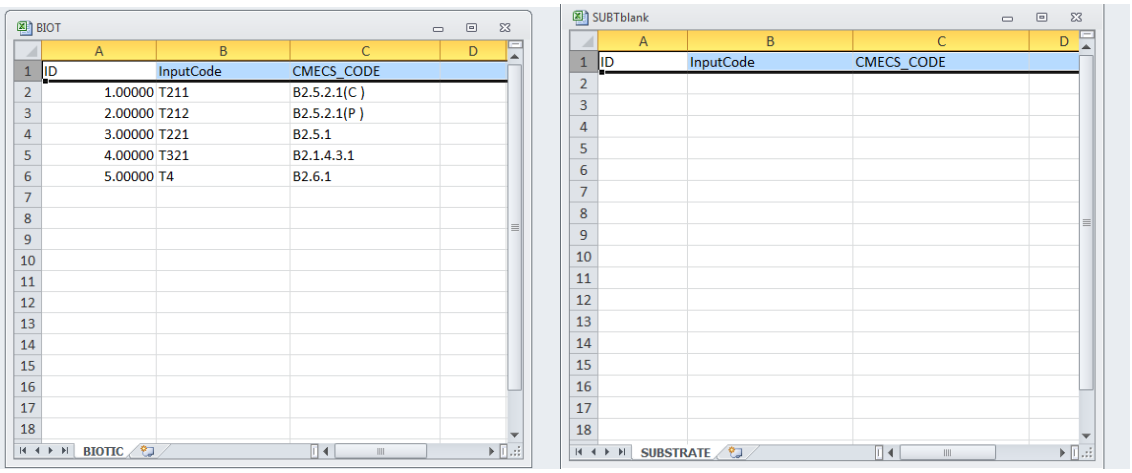

Figure 10: Example crosswalk spreadsheets in *.xlsx format for the biotic crosswalk at left and an “empty” table for the substrate component, for which no crosswalk will be conducted.

6. Once spreadsheets have been completed for the three components, bring the new LUT file into an ArcMap session using the **Add Data** button.

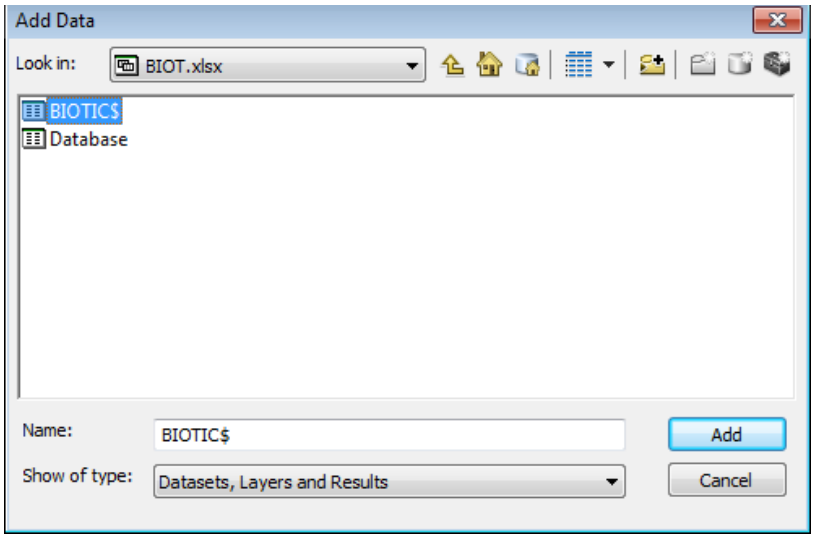

Figure 11: Exporting the Excel spreadsheet to *.dbf format

Then export the Excel (*.xlsx) files to a standard *.dbf format by selecting the table and using a right-mouse button. Be sure to name these by component. As the export completes, ArcMap will prompt you to add the *.dbf file to the view. It is recommended that you allow this and take the opportunity to review the attributes to make sure your codes are accurate. These *.dbf files (LUTs) will then be entered in the CCT tool dialog box.

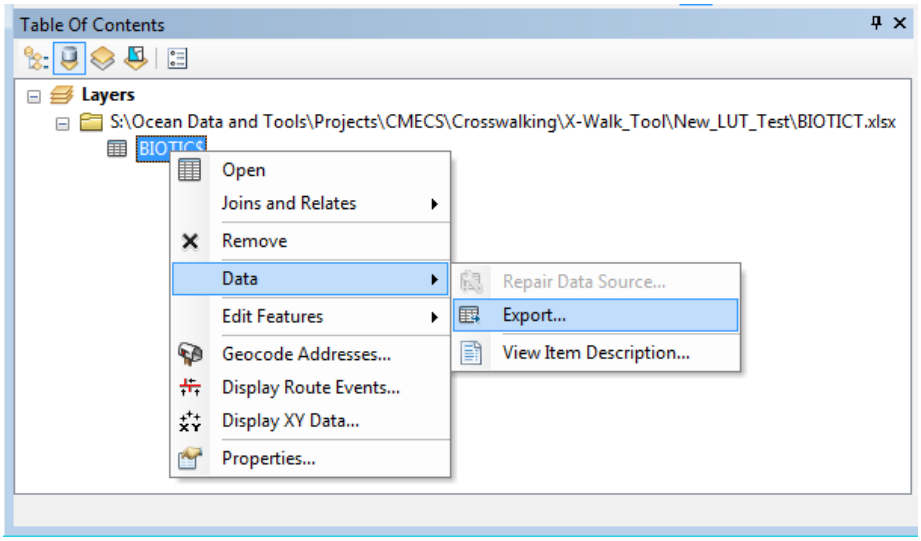

Figure 12: Exporting the Excel spreadsheet to *.dbf format

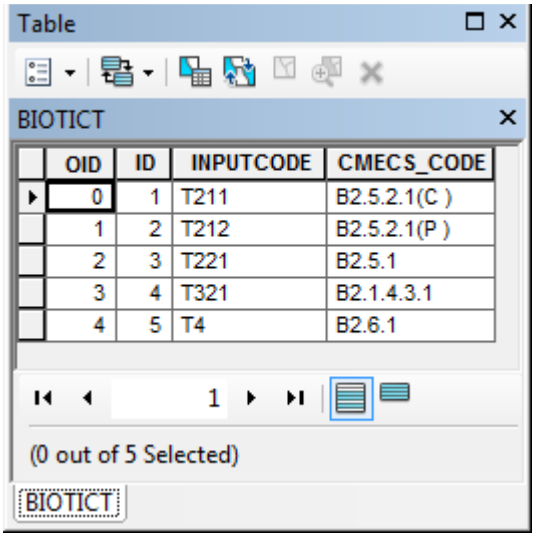

Figure 13: Attributes of Exported Biotic Look Up Table (*.dbf format). Ready for use in CCT.

At this point you can open the CCT tool and go through the process of selecting the input data set, workspace, and LUTs to be used for the crosswalk. The CMECS codes files will be the same ones as provided with the tool and found in the codes folder.

NOTE: if only crosswalking one of the components, you will still need to assign output feature classes for the non-used components in order for the tool to run. These will contain all the original vector geometry of the input file but will have null values for all the CMECS attributes. If unneeded, these can be deleted after running the tool.

After running the tool, you should see your new crosswalk results as the three feature layers in your target workspace (geodatabase). Review the output results to ensure that input classes were translated to the correct CMECS units and that the data captures your features of interest.

If working with a system other than the ones provided with the crosswalk tool, users are advised to first conduct a conceptual crosswalk before applying the tool. This will avoid an incorrect crosswalk where unit relationships are ambiguous or emphasize different ecological elements.

Users are encouraged to share their experience as well as any LUTs they are willing to share with the broader CMECS user community. The CMECS development team is available to consult on the use of this tool and the CMECS standard itself. For more information, please contact the CMECS Implementation Group at *nos.csc.cmecs_ig@noaa.gov.*

References

1. Madley, K., A. B. Sargent, and F. J. Sargent. 2002. *Development of a System for Classification of Habitats in Estuarine and Marine Environments (SCHEME) for Florida*. Report to the U.S. Environmental Protection Agency Gulf of Mexico Program (Grant Assistance Agreement MX- 97408100). St. Petersburg, FL: Florida Marine Research Institute, Florida Fish and Wildlife Conservation Commission.

2. FGDC (Federal Geographic Data Committee). 1996. *Classification of Wetlands and Deepwater Habitats of the United States*. FGDC- STD-004. Reston, VA: Federal Geographic Data Committee.

3. Kutcher, T. E. 2008. *Habitat and Land Cover Classification Scheme for the National Estuarine Research Reserve System.* Wells, ME: Wells Bay National Estuarine Research Reserve.

4. FGDC (Federal Geographic Data Committee), Marine and Coastal Data Subcommittee. 2012. *Coastal and Marine Ecological Classification Standard*. FGDC-STD-018-2012. Washington, D.C.: Federal Geographic Data Committee.

Users are encouraged to share their experience in using this tool, and to consult the CMECS webpage for additional crosswalk tables. The CMECS development team is available to consult on the use of this tool and the CMECS standard itself. For more information, please contact the CMECS Implementation Group at *nos.csc.cmecs_ig@noaa.gov.*