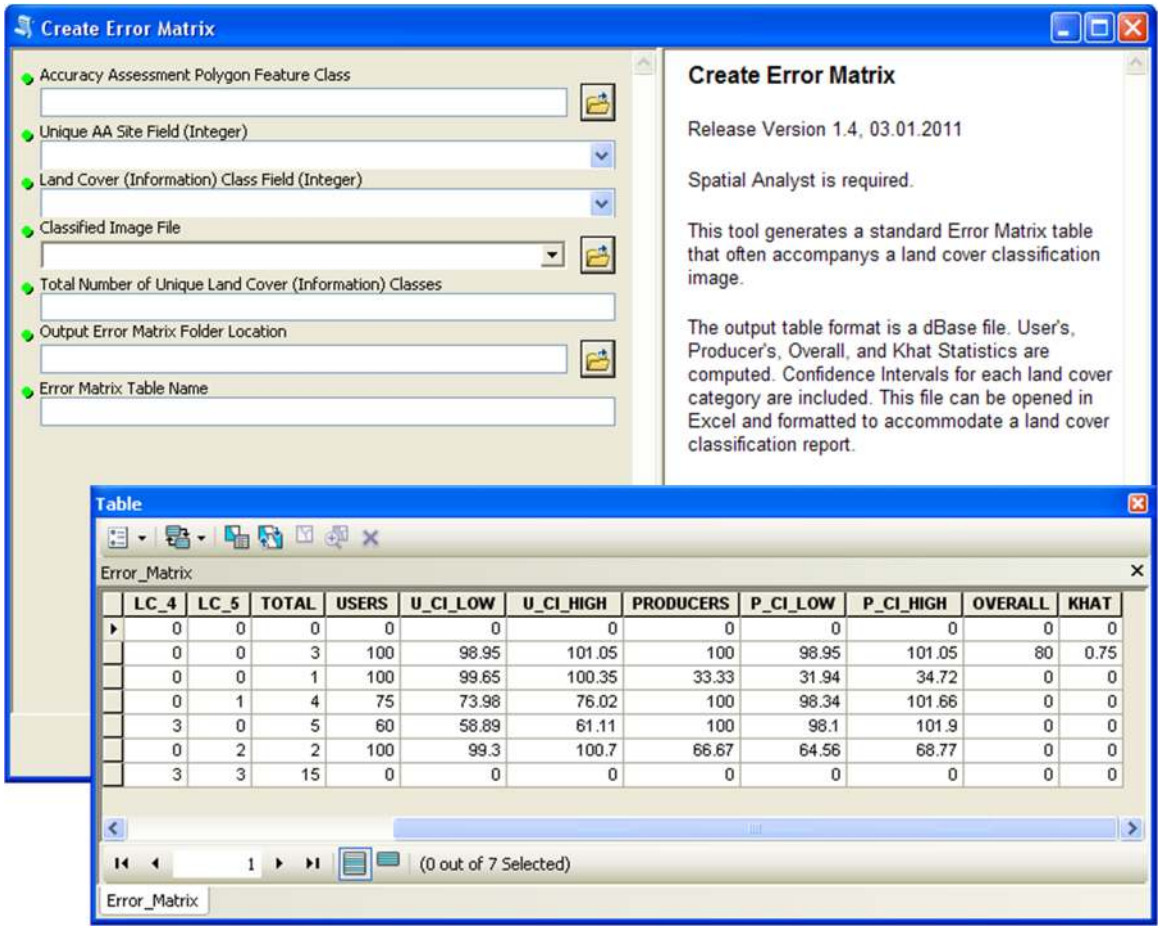
Accuracy Assessment Python Tool
for ArcGIS v 1.4



LC_4	LC_5	TOTAL	USERS	U_CI_LOW	U_CI_HIGH	PRODUCERS	P_CI_LOW	P_CI_HIGH	OVERALL	KHAT
0	0	0	0	0	0	0	0	0	0	0
0	0	3	100	98.95	101.05	100	98.95	101.05	80	0.75
0	0	1	100	99.65	100.35	33.33	31.94	34.72	0	0
0	1	4	75	73.98	76.02	100	98.34	101.66	0	0
3	0	5	60	58.89	61.11	100	98.1	101.9	0	0
0	2	2	100	99.3	100.7	66.67	64.56	68.77	0	0
3	3	15	0	0	0	0	0	0	0	0

Created: 04.05.2010
Updated: 03.01.2011

Created by: Nathan Jennings
nate@jenningsplanet.com
www.jenningsplanet.com

Purpose

Currently, no accuracy assessment function exists in either ArcGIS or ERDAS Imagine that operate on an accuracy assessment polygon file. Computing an error matrix based on an "area" of land cover versus a single pixel is often desired to assess the quality of a digital land cover classification.

The Accuracy Assessment tool is a modified version of the original Accuracy Assessment Tool that was originally created using VBA and ArcObjects (ca. 2007) by the author. This version is supported by Python 2.5 and ArcGIS 9.3.1 and Python 2.6.5 and ArcGIS 10. A custom Accuracy Assessment Toolbox has been created to use a Python script.

The tool generates a standard Error Matrix for land cover classifications that compute user's and producer's categorical accuracy assessment values in addition to overall and Khat statistics. Confidence intervals are computed for each of the land use categories for both user's and producer's accuracies as well.

This application is provided to the end user as is. The author asks that the application not be sold, enhanced, or modified without the written consent. The author can be contacted at the above email address. Please submit all questions, issues, problems, etc. directly to the author.

Program Package

The package comes with sample data for the user to become familiar with the process and general data format. The data was derived for a remote sensing class the author of the program teaches. Although the actual data has not been rigorously validated, it does provide the context to create an error matrix. In practice, the accuracy assessment sites will be validated and the land cover data set will be thoroughly reviewed before assessing the accuracy.

The data package contains:

Accuracy_Assessment.mxd – ArcGIS 10 map document that contains the data and the toolbox

Accuracy Assessment.tbx – ArcGIS toolbox that contains the Create Error Matrix Tool

accuracy_assessment_v14ArcGISTool.py – Python script that creates and computes the error matrix

ImageClassification.gdb – file geodatabase that contains the following feature classes

a. **AA_Sites** – feature class of accuracy assessment polygons with unique accuracy assessment polygon ID (*AA_ID*) and land cover (information class) class number (*AA_Site_Num*). The feature class has 15 accuracy assessment polygons representing 5 unique land cover class values (1-5).

b. **Spectral_Sigs** – polygon feature class that represents the spectral signatures of the various land cover classes. These polygons were used to generate a spectral signature file that was used to create the land cover classification image. *SIG_NUM* – unique signature number (used in the image classification) and *SIG_NAME* – the name of the signature. This file is used for reference and is not used in the accuracy assessment tool.

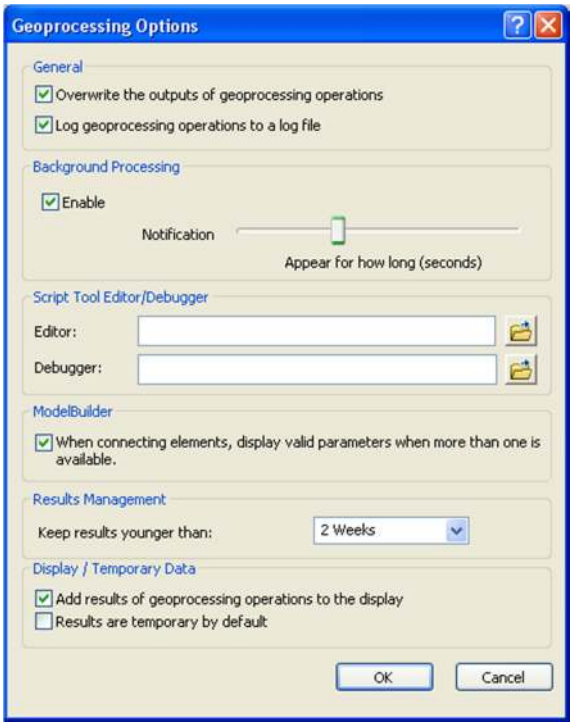
MLClass.img – Imagine image file representing the resulting supervised classification

Reclass_ML.img – Imagine image file that contains recoded thematic land cover types (i.e. information classes). The .RRD file represents the pyramid file that often accompanies image files to assist with faster display.
tm_sacsub.img – Image file subset of Landsat TM satellite image.

Error_Matrix.dbf – sample output dbase file. Users can choose their own name. This file represents the output from the accuracy assessment tool.

maj_table.dbf – sample zonal statistics table. This table is a temporary table and will be overwritten every time the program runs. The important statistic in this table that is used in the error matrix is the majority attribute which contains the class number that represents the majority of pixels found in the accuracy assessment polygon.

NOTE: The Python script contains an overwrite parameter that should take care of overwriting any existing data (i.e. maj_table.dbf or the error matrix with the same file name). The user can also check the Geoprocessing Options to make sure the Overwrite Output is checked.


Geoprocessing Options
General
Overwrite the outputs of geoprocessing operations
Log geoprocessing operations to a log file
Background Processing
Enable
Notification
Appear for how long (seconds)
Script Tool Editor/Debugger
Editor:
Debugger:
ModelBuilder
When connecting elements, display valid parameters when more than one is available.
Results Management
Keep results younger than:
2 Weeks
Display / Temporary Data
Add results of geoprocessing operations to the display
Results are temporary by default
OK
Cancel

Requirements

The tool has added the capability of generating the majority land cover class from each accuracy assessment polygon using the Spatial Analyst's Zonal Statistics as Table tool. Users must have the Spatial Analyst license available to run the Zonal Statistics as Table tool and to run the Python script. The script performs a check to see if the user has the spatial analyst license and if not, provides a warning message.

The current version of this tool only supports a dBase (.DBF) table. File and personal geodatabase tables are not supported. Future versions may include this functionality.

Analysts can use the sample data provided to format their own accuracy assessment polygon feature class.

Data Requirements

1. Accuracy Assessment polygon feature class that contains at a minimum a unique ID representing each accuracy assessment polygon and a field that represents the specific "true" land cover type. Both fields must contain integer numbers. Currently, the feature class format can be shapefile, personal, or file based geodatabase. The accuracy assessment tool contains a parameter that asks for a folder for the output error matrix table (currently in .DBF format).

2. Classified thematic land cover image. This file must contain the unique land cover categories (i.e. information classes). Normally, this value is stored as an integer within an image file. The file format can be any commonly used format accepted by ArcGIS. The tool has been tested with the ERDAS Imagine format (.img file).

Using the Error Matrix Tool

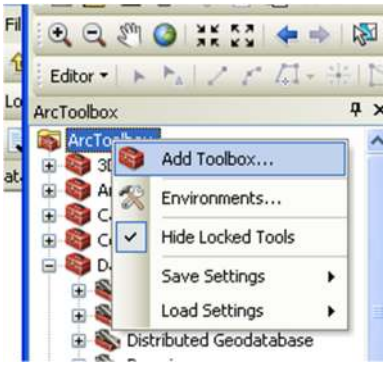
The program package contains an Accuracy Assessment Toolbox that holds the Create Error Matrix script tool. End users of the tool will need to add this toolbox and likely need to re-path the Create Error Matrix Tool Python script (see below).

Accuracy Assessment Toolbox and Create Error Matrix Tool Set Up

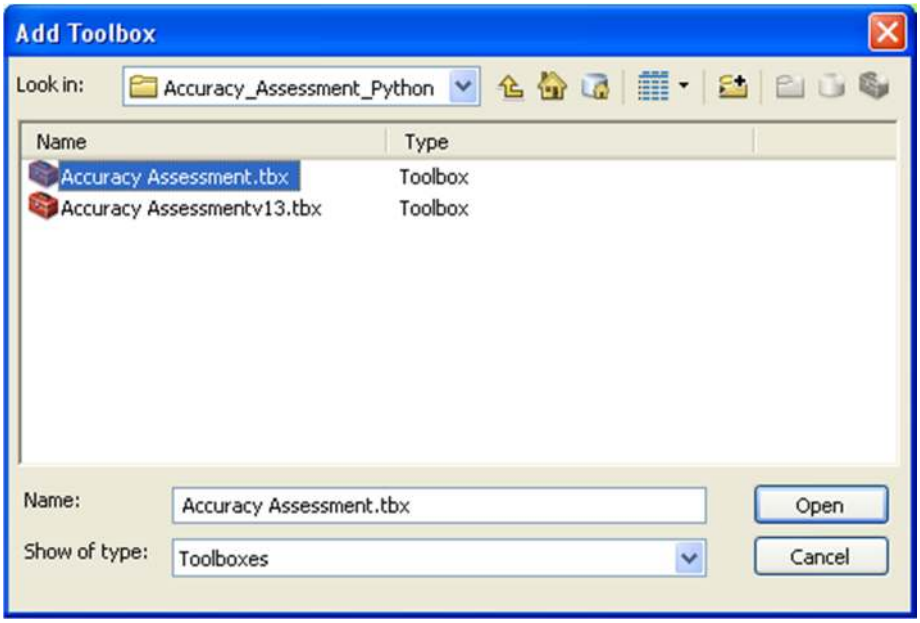
The user can copy the toolbox to the same folder as the ArcMap document for use within that document. The demo data already includes a folder with the Accuracy Assessment Toolbox and an ArcMap document (ArcGIS 10) that contains the toolbox in the ArcToolbox list.

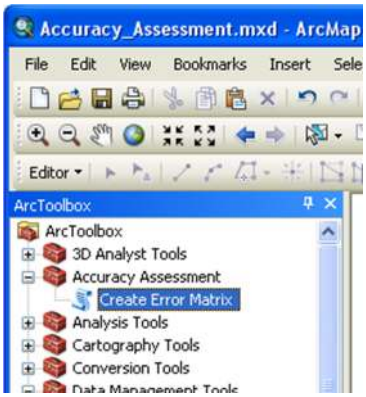
The user can add the toolbox on their own for use in other ArcMap documents by doing the following:

1. Open up ArcMap and Add the Accuracy Assessment Toolbox to the ArcToolbox. Right click on the ArcToolbox name.

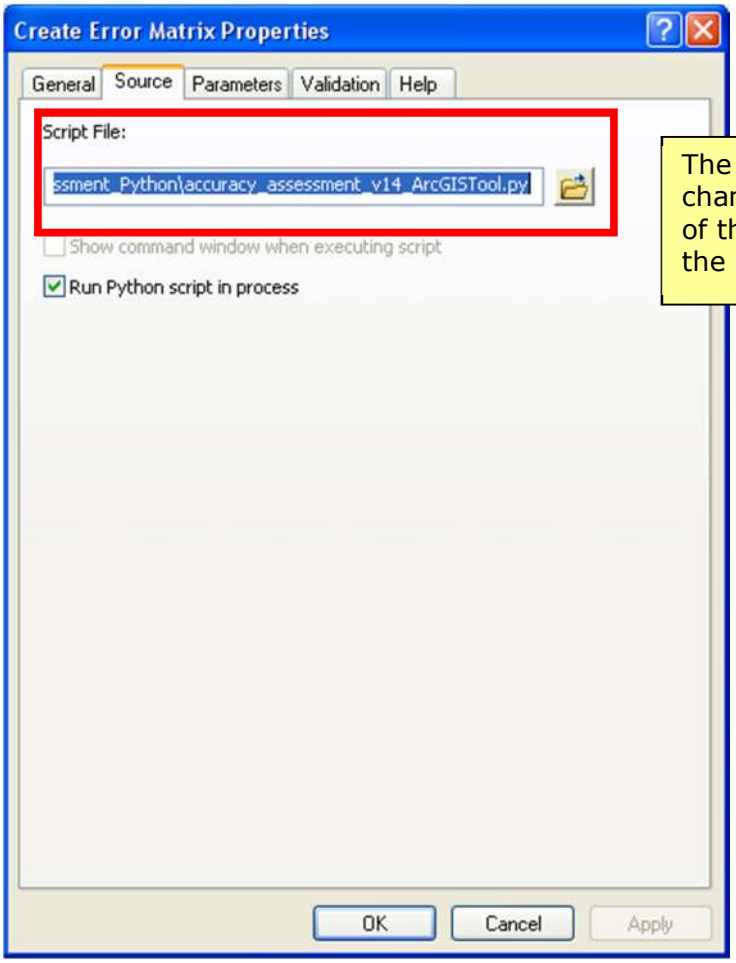


2. Locate the Accuracy Assessment tool. Browse to the path that contains the Accuracy Assessment toolbox file.





3. Right click on the Create Error Matrix and Select Properties. Select the Source tab and change the path to the Python script to the location on the user's computer. The user must do this before running the tool the first time, even with the demo data. Currently, **accuracy_assessment_v14_ArcGISTool.py**.

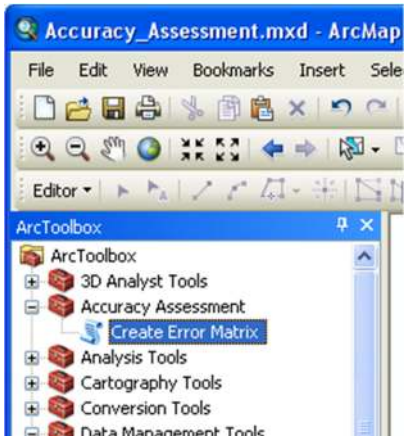


The path will need to be changed to the location of the Python script on the user's computer.

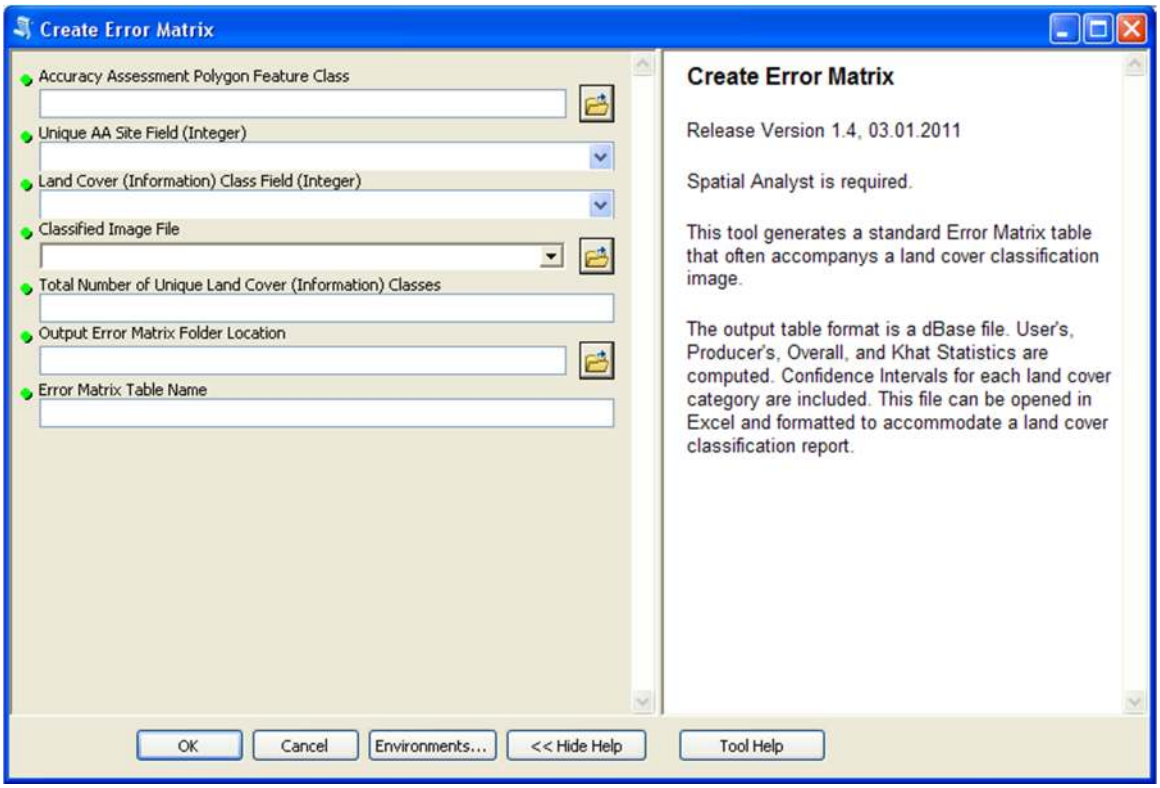
Operate the Create Error Matrix Tool

Once the Create Error Matrix Tool is set up and the required data is formatted properly, the user can fill in the tool parameters and click OK.

1. Locate the Accuracy Assessment toolbox and select the Create Error Matrix Tool



2. Fill in the parameters with the required information. All parameters are required.



Accuracy Assessment Polygon Feature Class

Locate the Accuracy Assessment polygon feature class. This can be a shapefile, file or personal geodatabase polygon file.

Unique AA Site Field (Integer)

Select from the list the specific field that represents the unique accuracy assessment site ID. This must be an integer field. (*AA_ID*)

Land Cover (Information) Class Field (Integer)

Select from the list of specific fields that represent the unique land cover classes. This attribute will contain specific values for land cover (information) class types. (*AA_Site_Num*)

Classified Image File

Locate the classified image file that contains the thematic land cover image (*Reclass_ML.img*). This file will contain the reclassified (or recoded) values and represent information classes.

Total Number of Unique Land Cover Classes

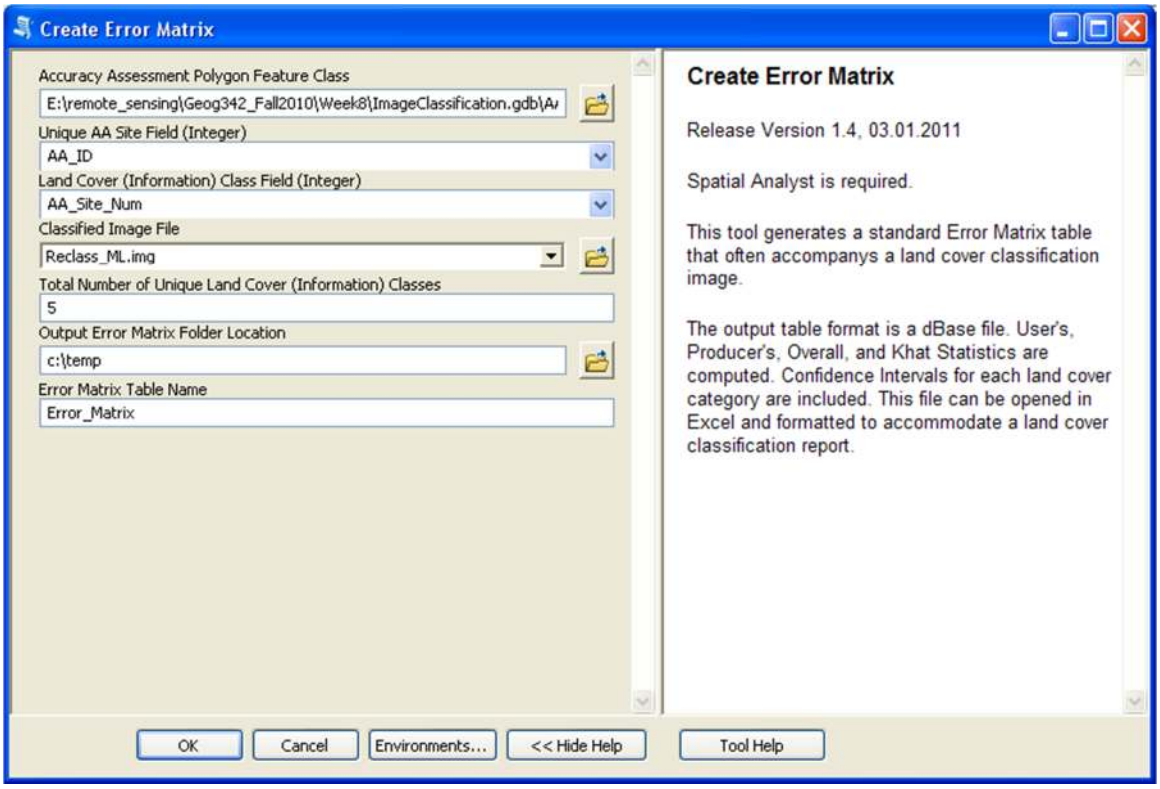
Enter a number that represents the unique land cover types. This is used to build the error matrix and to ensure that a field and row exist to hold the accuracy assessment values. For the sample data, this will be 5.

Folder Location

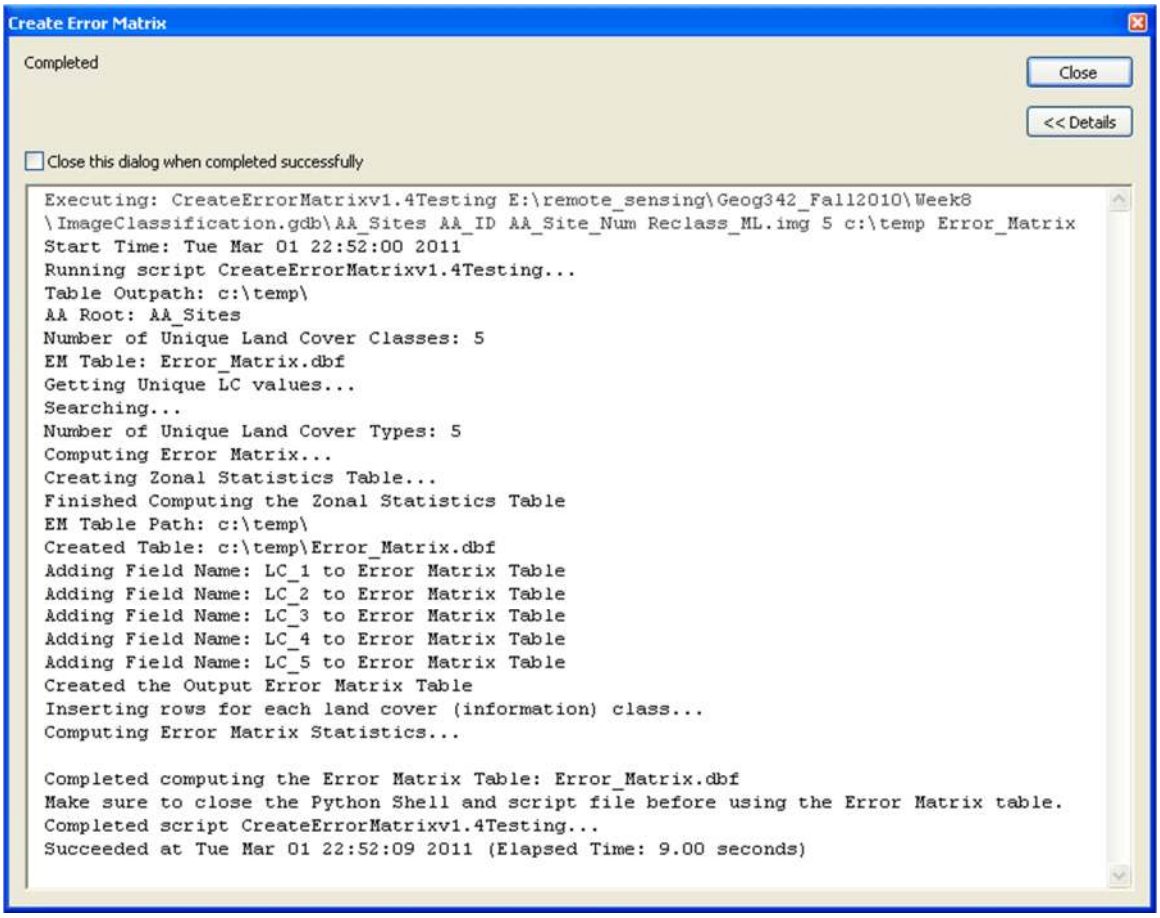
Input a folder where the new Error Matrix will be created. Ex. C:\temp.

Error Matrix Table Name

Enter a name for the error matrix table. If a .DBF extension is not added, it will be through the table creation process. Ex. Error_Matrix.

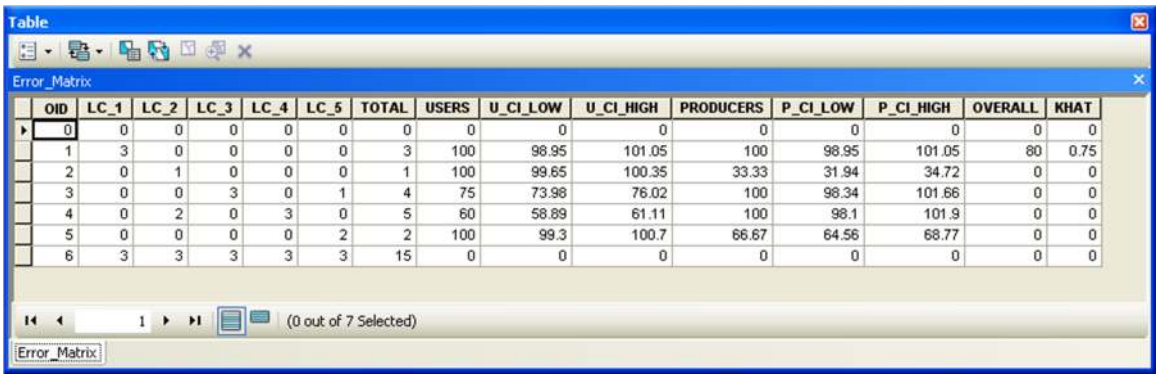


The following process dialog box appears showing progress.



Error Matrix Table

The following table is produced from the process. This can be loaded into ArcMap for viewing and using or can be brought into Excel for formatting for a document.



OID	LC_1	LC_2	LC_3	LC_4	LC_5	TOTAL	USERS	U_CI_LOW	U_CI_HIGH	PRODUCERS	P_CI_LOW	P_CI_HIGH	OVERALL	KHAT
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	3	0	0	0	0	3	100	98.95	101.05	100	98.95	101.05	80	0.75
2	0	1	0	0	0	1	100	99.65	100.35	33.33	31.94	34.72	0	0
3	0	0	3	0	1	4	75	73.98	76.02	100	98.34	101.66	0	0
4	0	2	0	3	0	5	60	58.89	61.11	100	98.1	101.9	0	0
5	0	0	0	0	2	2	100	99.3	100.7	66.67	64.56	68.77	0	0
6	3	3	3	3	3	15	0	0	0	0	0	0	0	0

Table

Error_Matrix

OID	LC_1	LC_2	LC_3	LC_4	LC_5	TOTAL	USERS	U_CI_LOW
0	0	0	0	0	0	0	0	0
1	3	0	0	0	0	3	100	98.95
2	0	1	0	0	0	1	100	99.65
3	0	0	3	0	1	4	75	73.98
4	0	2	0	3	0	5	60	58.89
5	0	0	0	0	2	2	100	99.3
6	3	3	3	3	3	15	0	0

1 (0 out of 7 Selected)

Error_Matrix

Table

Error_Matrix

LC_4	LC_5	TOTAL	USERS	U_CI_LOW	U_CI_HIGH	PRODUCERS	P_CI_LOW	P_CI_HIGH	OVERALL	KHAT
0	0	0	0	0	0	0	0	0	0	0
0	0	3	100	98.95	101.05	100	98.95	101.05	80	0.75
0	0	1	100	99.65	100.35	33.33	31.94	34.72	0	0
0	1	4	75	73.98	76.02	100	98.34	101.66	0	0
3	0	5	60	58.89	61.11	100	98.1	101.9	0	0
0	2	2	100	99.3	100.7	66.67	64.56	68.77	0	0
3	3	15	0	0	0	0	0	0	0	0

1 (0 out of 7 Selected)

Error_Matrix

Please send all requests, questions, comments, issues, etc to the email above.