GST 105: Introduction to Remote Sensing Lab Series

Lab 6: Accuracy Assessment

Document Version: 2013-09-24 (Beta)

Organization: American River College
Author: Nathan Jennings

The development of this document is funded by the Department of Labor (DOL) Trade Adjustment Assistance Community College and Career Training (TAACCCT) Grant No. TC-22525-11-60-A-48; The National Information Security, Geospatial Technologies Consortium (NISGTC) is an entity of Collin College of Texas, Bellevue College of Washington, Bunker Hill Community College of Massachusetts, Del Mar College of Texas, Moraine Valley Community College of Illinois, Rio Salado College of Arizona, and Salt Lake Community College of Utah.

Contents

Introduction

A primary aspect of image classification is to quantitatively validate the resulting land cover dataset. A common method is to perform an accuracy assessment that uses an error matrix to compute a number of quantitative measures on the land cover dataset. Some image processing software packages provide this functionality, however, with ArcGIS, this functionality is not provided. For this lab, students will use a custom-built accuracy assessment routine that functions within ArcMap.

Objective: Perform an Accuracy Assessment

Students will use the custom Accuracy Assessment Tool created by the author to perform a simple accuracy assessment on a classified image. The results will be reviewed and interpreted to determine how well the image classification performed. The image classification has already been performed and the accuracy assessment sites required to compute the accuracy assessment is provided.

Lab Settings

Required Virtual Machines and Applications

Windows Machine User Account	Train
Windows Machine User Password	Train1ng$

All of the data for this lab can be found in the **Lab *6\Accuracy_Assessment_Python*** folder.

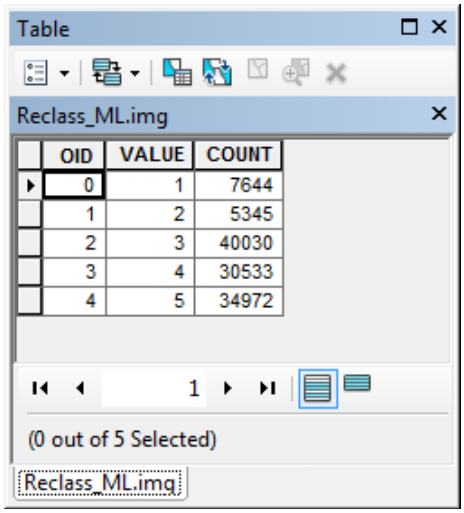
1 Open the Accuracy Assessment ArcMap Document

1. Log into the computer, using the information provided in the Lab Settings section.
2. Open the **Accuracy_Assessment.mxd** in ArcMap. This ArcMap document contains a small subset of a Landsat TM image (**tm_sacsub.img**) as well as a "classified image" (**ML_Class.img**), and a reclassed image named **Reclass_ML.img**.

   This image contains five land cover types:

   1) Agriculture
   2) Water
   3) Grassland
   4) Forest
   5) Urban

3. Right-click on the **Reclass_ML.img** and select **Open Attribute Table**. The following table will appear. The Value field contains the unique land cover values that correspond to the individual types of land cover (information classes).



4. Expand the ML_class.img layer in the Table of Contents (if not expanded). The **ML_class.img** file represents a typical output from a maximum likelihood image classification where every spectral class is categorized in the output.

- The **Reclass_ML.img** file is the recoded version of the **ML_class.img** where the spectral classes have been recoded to the five information classes shown above.

- The **Reclass_ML.img** file is the image that will be used in the accuracy assessment.

- In addition to the images, two polygon files are provided.
    - The **Spectral_Sigs** feature class represents the sites that were used to generate the training sites that were used to produce the **MLClass.img**.
    - The **AA_Sites** feature class contains the accuracy assessment polygons and will be used in the accuracy assessment.

2 Review the Accuracy Assessment Feature Class

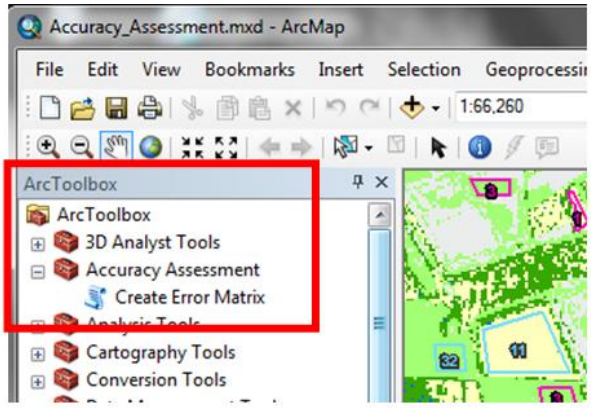
Open the attribute table for the AA Sites feature class. For each AA Site polygon, notice the **AA_Site_Num**, **AA_Sites_Name**, and the **AA_ID** columns. The **AA_Sites** polygon feature class is one of the required parameters in the Create Error Matrix tool used below.

AA_Site_Num	Contains the specific unique land cover classes (1-5).
AA_Site_Name	The name of the corresponding land cover class.
AA_ID	A unique id for each accuracy assessment polygon. Note, this is not the same as the ObjectID..

The **AA_ID** and the **AA_Site_Num** values will be used in the Accuracy Assessment program.

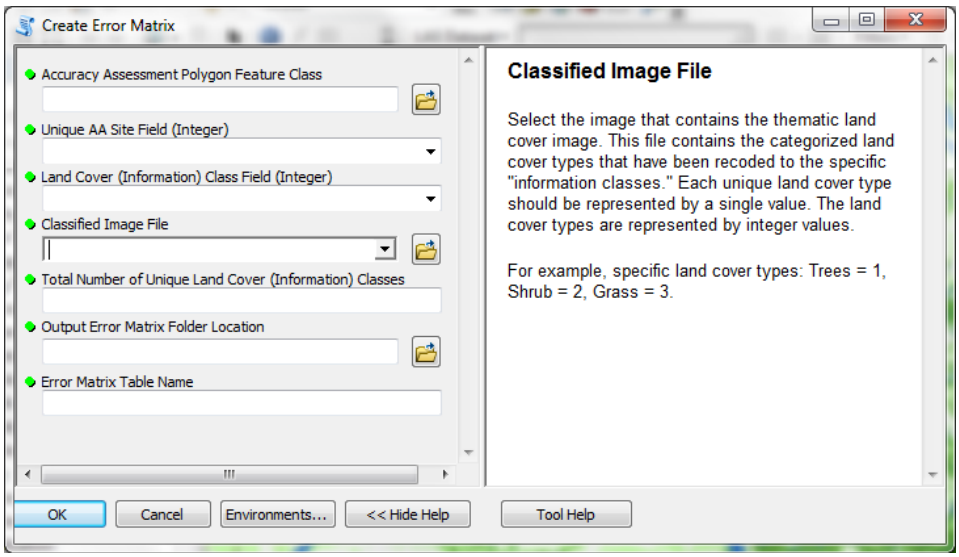
3 Run the Create Error Matrix Tool

1. Open the **ArcToolbox** tab and then open the **Create Error Matrix** tool within the **Accuracy Assessment** toolbox. Click on the Show Help button to see more information about each parameter.



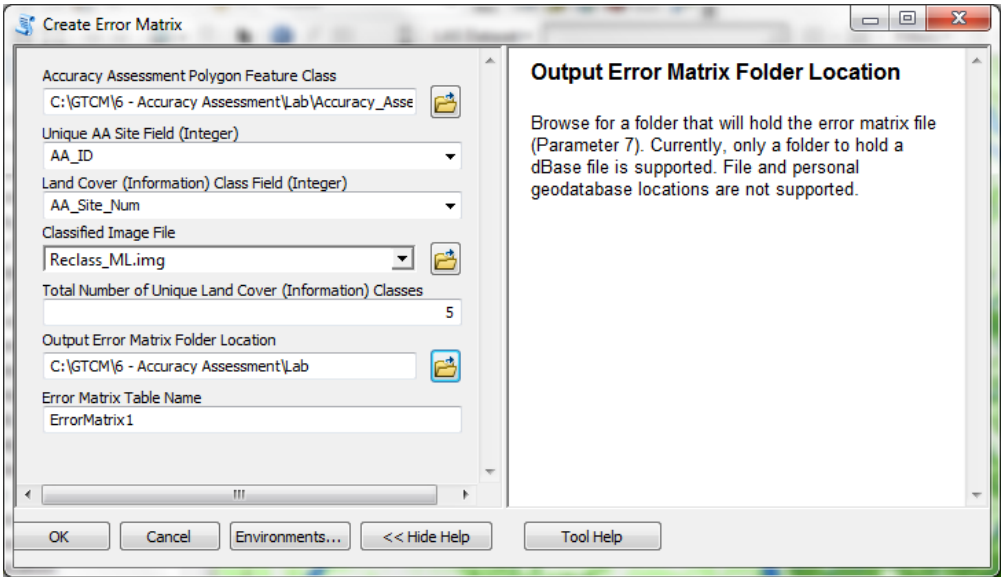
The **AA_Sites** feature class can be found in the **\Accuracy_Assessment_Python\ImageClassification.gdb** file geodatabase.

A Data Connection may be required to locate the data for this lab. Connect to the ***Accuracy_Assessment_Python*** folder and then locate it in the **Look in:** dropdown.



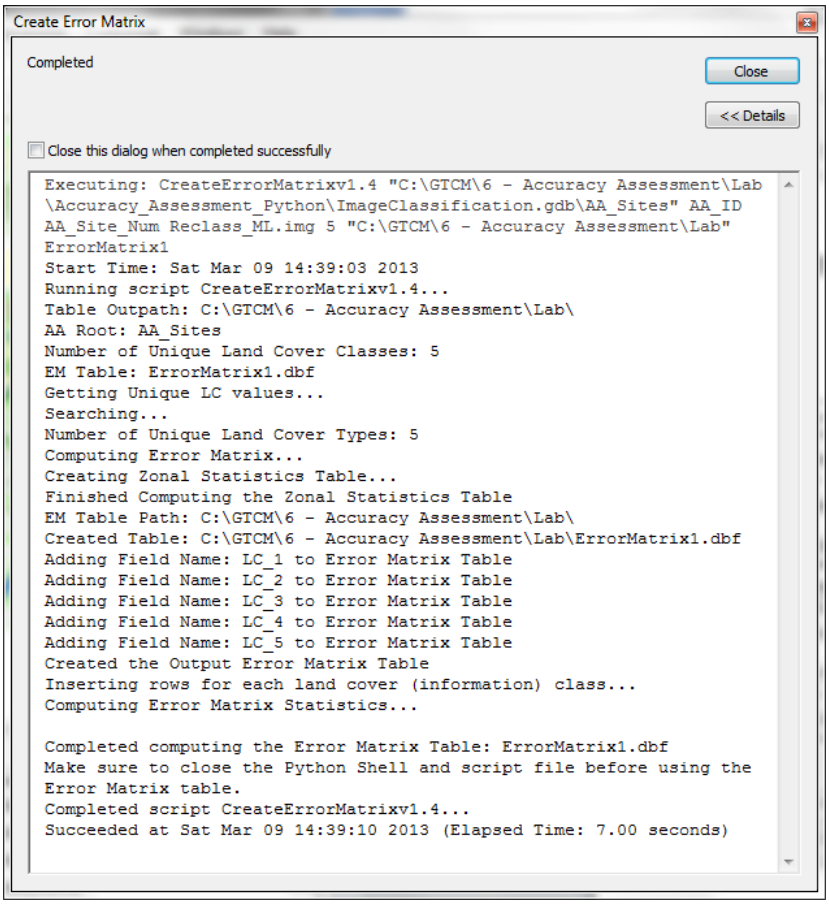
2. Fill in the parameters using the values provided in the table below.

Parameter	**Value**
Accuracy Assessment Polygon Feature Class	**AA_Sites**
Unique AA Site Field	The **AA_ID** Attribute from the AA_Sites feature class
Land Cover (Information) Class Field (Integer)	The **AA_Site_Num** field from the AA_Sites feature class
Classified Image File	The **ReclassML.img** file that is found within the MXD
Total Number of unique Land Cover (Information) Classes	Enter **5**. This number must be the same as the total number of unique land cover values (see the attribute table for the ReclassML.img file).
Output Error Matrix Folder Location	The folder on disk that will contain the resulting error matrix file. This parameter must be a folder, not a geodatabase workspace (e.g. c:\temp). Note, you may need to make a data connection to access a specific folder of your choice.
Error Matrix Table Name	Enter a name without spaces, dashes, or other special characters and do not put a file extension on it (e.g. *ErrorMatrix1*). The output table will be a .DBF (dBase) file.



3. Click **OK**.

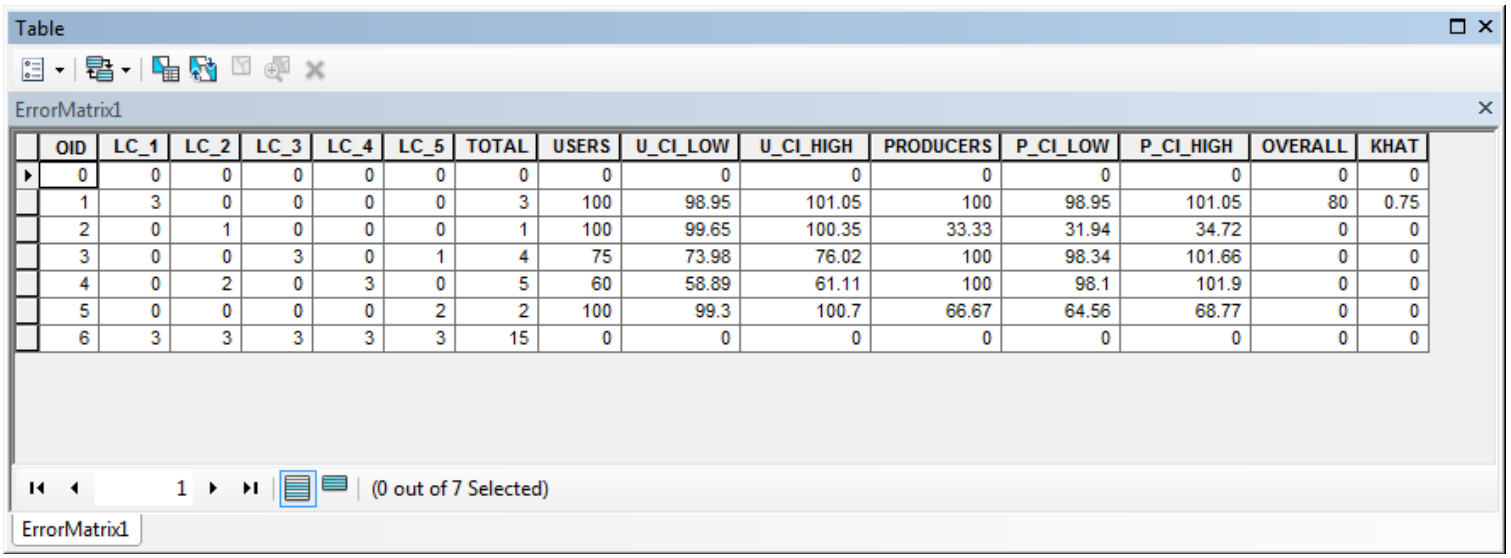
The program takes a few seconds to run. Progress messages are displayed on the screen.



The Error Matrix has been computed.

4 Review and Interpret the Error Matrix

Add the error matrix dBase table to the Table of Contents and open it to review the data,



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

The left side of the table contains the matrix. The right side contains the various statistics computed for the error matrix.

LC_x columns represent the specific land cover classes. Note the first row is all zeros. This is just a "spacer" row. Row OID 6 contains the column totals from rows with OID (1-5). The Total column contains the accuracy assessment (in this case 15).

The right side of the table contains the user's and producer's accuracies in addition to the low and high confidence intervals. The last two columns contain a single value for the overall accuracy and the Khat statistic.

The error matrix needs to be evaluated to determine which classes are confused and to note which categories of the lowest user's and producer's accuracies.

Conclusion

This completes the quantitative analysis of an image classification dataset. Students performed a simple image classification and were able to observe and interpret the user's, producer's, overall, and KHAT measures, in addition to the confidence intervals for each land cover category. Students should now be familiar with the specific methods and measures that are often required to perform, document, and explain an accuracy assessment.

Review Exercises

Please refer to the error matrix created in section 4 and answer the following questions.

Exercise A: *Which class has the most confusion for a given column (only those with LC_x)? How do you know?*

Exercise B: *Which row (1-5) has the most confusion? How do you know?*

Exercise C: *Is there anything suspect about this error matrix? Describe and explain.*

Exercise D: *Which class has the lowest user's accuracy?*

Exercise E: *Which class has the lowest producer's accuracy?*

Exercise F: *Given that there are 15 accuracy assessment sites for five different land cover classes and the overall accuracy is 80%, would you think the accuracy assessment is valid? Why or why not? Justify your answer.*