CLASSIFICATION ASSIGNMENT: PART III – ACCURACY ASSESSMENT

FELIPE CAMACHO HURTADO

SALZBURG UNIVERSITY

ADVANCED REMOTE SENSING

SALZBURG, AUSTRIA

JANUARY 7, 2023

CONTENTS

1. OBJECTIVES

1. Determine the importance of the accuracy assessment phase in an object-based image classification workflow.
2. Recognize different approaches to perform an accuracy assessment of a classification result.
3. Identify the different elements that can influence the results of an accuracy assessment.

2. USED SOFTWARE

1. ArcGIS Pro 3.0.0
2. eCognition Developer 10.1
3. eCognition Developer 10.2

3. USED DATA

1. Classification Data (Provided by the tutor)
2. Labelled Sample Points: result_NN_classification.shp (Reference Dataset)
3. Riparian Zones Polygons: Riparian_zone_LULC.shp (Reference Dataset)
4. Image: TTAmask_SPOT.tif and the associated *.csv file providing the class names.

4. INTRODUCTION

The initial idea of this exercise was to assess the accuracy of the classification performed in "PART I". However, as the classification was created using eCognition 9 – Trial version, it was not possible to import it into licensed eCognition 10 software. For that reason, instead of using the mentioned datasets, the classification provided by the tutor will be assessed.

In this document, several approaches on how to generate a confusion matrix will be presented, for both eCognition and ArcGIS Pro software, including the usage not only of control points but also of reference polygons.

5. eCOGNITION 10.1

For eCognition 10.1, the first step (after performing a classification) is to load the reference data (including the classes code .CSV) as Test and Training Area (TTA) mask through: Samples > Load TTA Mask. (Figure 1). This mask contains the training samples and information about classes for the map. Then, the linkage between classes (Figure 2) of the map and the classes of the TTA must be done through: Classification > Samples > Edit Conversion Table, to ensure working with the same classes.

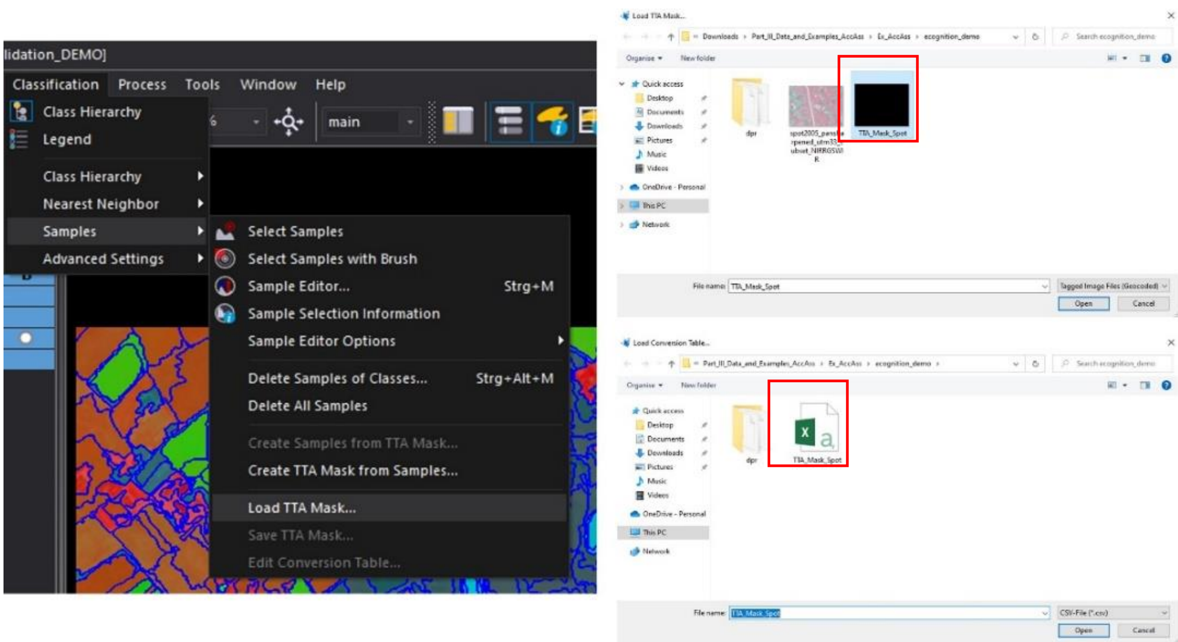


Figure 1. Adding TTA Mask and CSV files.

Once the reference data was added and the classes were linked correctly, the accuracy assessment can be carried out, using the “Calculate a confusion matrix based on the reference data” tool (Figure 3).

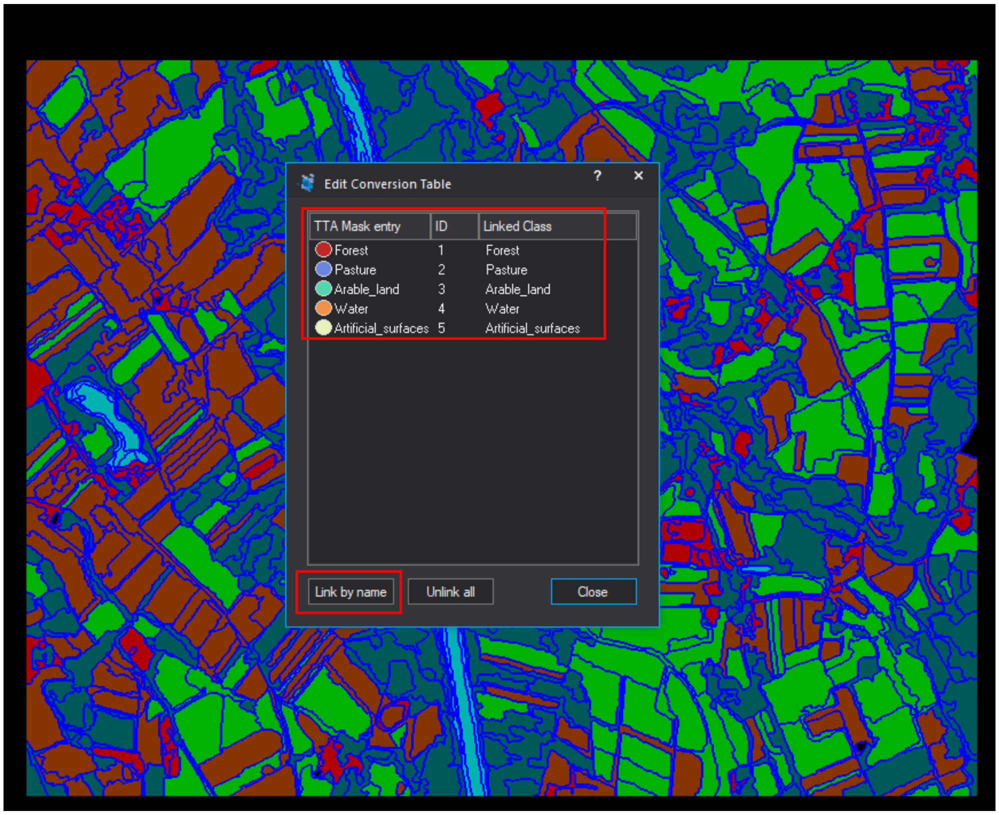


Figure 2. Linkage between TTA Mask classes and Map classes.

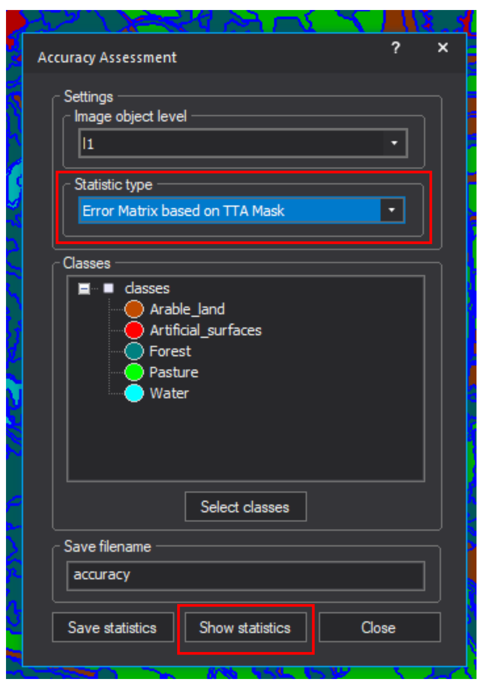
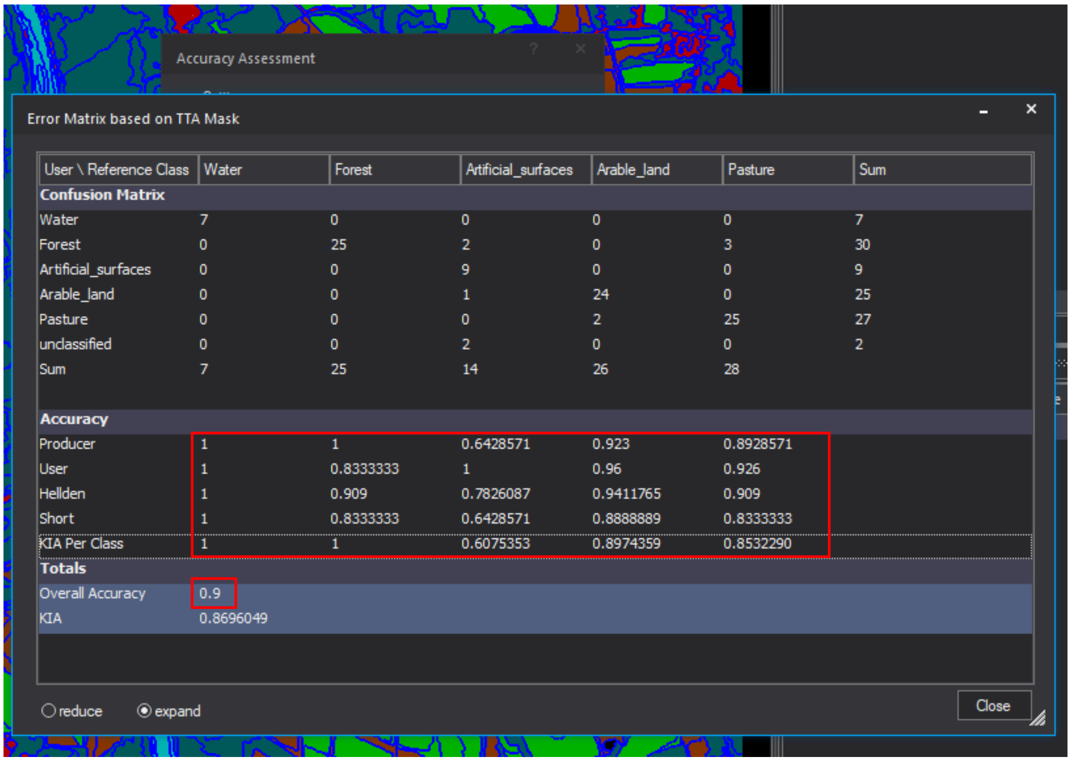


Figure 3. Creating Error Matrix based on TTA Mask.

After clicking on "Show Statistics", the "Error Matrix based on TTA Mask" will be generated (Figure 4). For the classification provided, the Error Matrix returned an overall Accuracy of

90%. The following were the User accuracy for all the classes: Water – 100%, Forest – 83.3%, Artificial Surfaces – 100%, Arable Land – 96%, and Pasture – 92.6%. In this case, the Artificial Surfaces and Water had the higher user accuracy, while the Forest class had the lowest.



User \ Reference Class	Water	Forest	Artificial_surfaces	Arable_land	Pasture	Sum
Confusion Matrix						
Water	7	0	0	0	0	7
Forest	0	25	2	0	3	30
Artificial_surfaces	0	0	9	0	0	9
Arable_land	0	0	1	24	0	25
Pasture	0	0	0	2	25	27
unclassified	0	0	2	0	0	2
Sum	7	25	14	26	28	
Accuracy						
Producer	1	1	0.6428571	0.923	0.8928571	
User	1	0.8333333	1	0.96	0.926	
Hellden	1	0.909	0.7826087	0.9411765	0.909	
Short	1	0.8333333	0.6428571	0.8888889	0.8333333	
KIA Per Class	1	1	0.6075353	0.8974359	0.8532290	
Totals						
Overall Accuracy	0.9					
KIA	0.8696049					

Figure 4. Resulting Error Matrix based on TTA Mask

6. eCOGNITION 10.2

For eCognition 10.2, two approaches how to perform Accuracy Assessment for the classification will be shown. The first is based on ground truth data as vector points, and the second one is based on reference data as a polygon vector dataset (considering an overlap threshold).

6.1 Based on Point Vector (Object Based)

After having added the vector reference points (Figure 5), you may have to delete any existing samples, by using the “delete samples” algorithm (Figure 6). This is done in case you still have samples that were used for the classification or other Accuracy assessment process, and do not correspond to ground truth data.

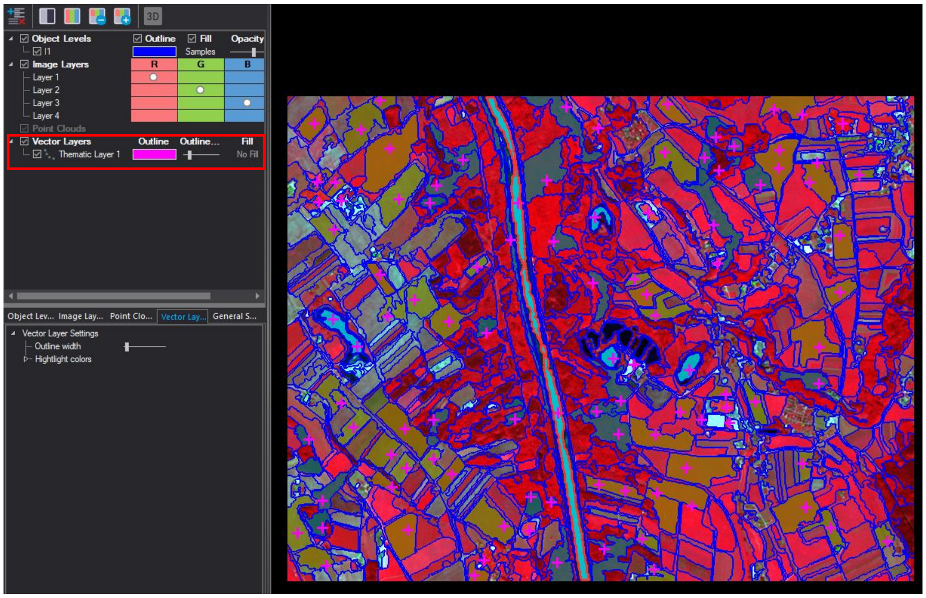


Figure 5. "Thematic Layer 1" – Reference Vector Points

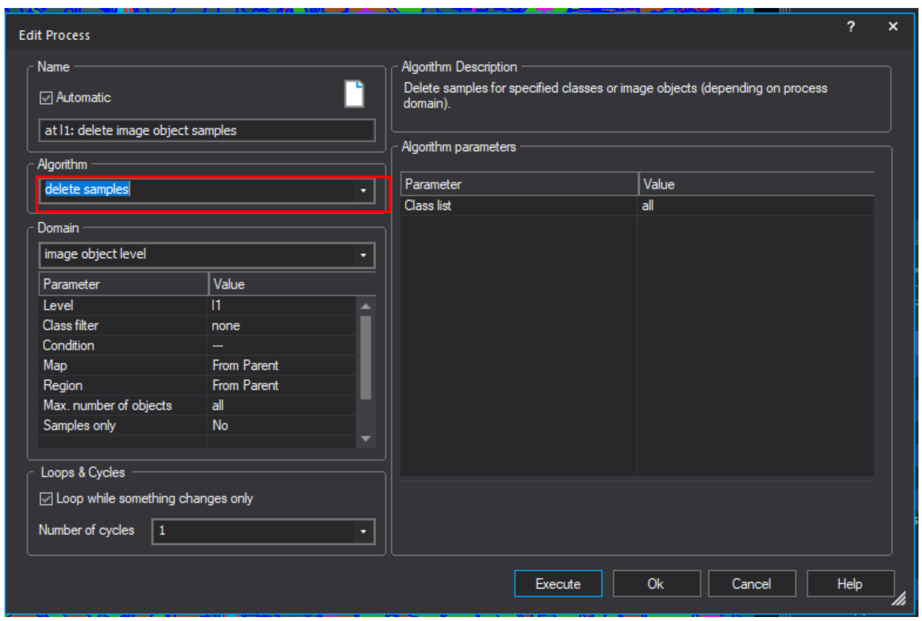


Figure 6. The "delete samples" process.

The next step is to use the "convert thematic objects to samples" process (Figure 7), after Remapping the Attribute Values (to match the classification convention of the reference data, with the classification convention of the map) of the reference data with the Mapped values, this process will assign the reference data class to the object that overlaps the vector points (Figure 8). Note: A new "String" field had to be created in ArcGIS Pro for the reference points layers, and be populated with the code classification convention, in order to use it as the input value for the Remap process.

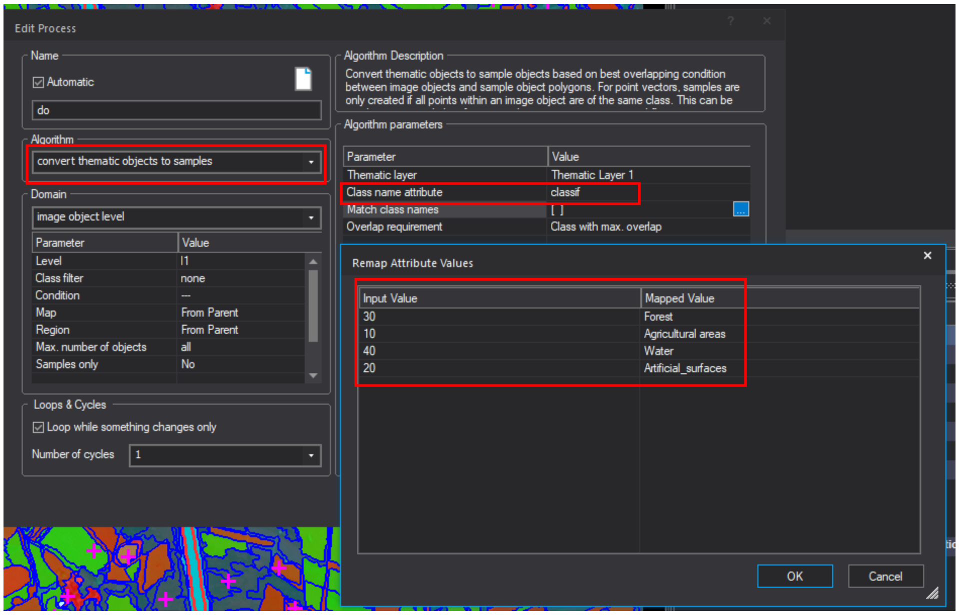


Figure 7. The "convert thematic objects to samples" process.

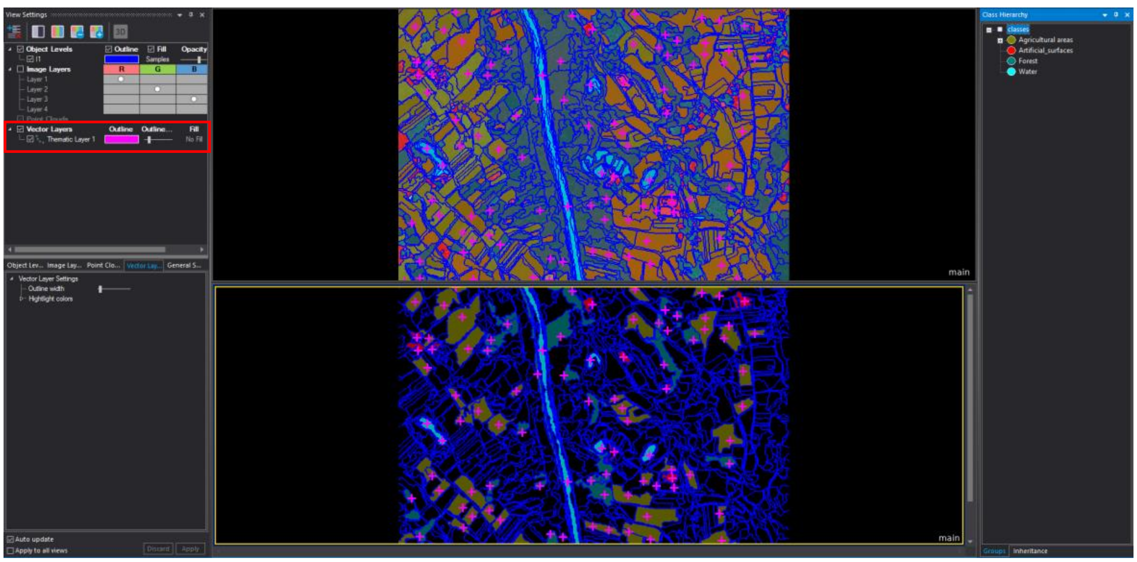
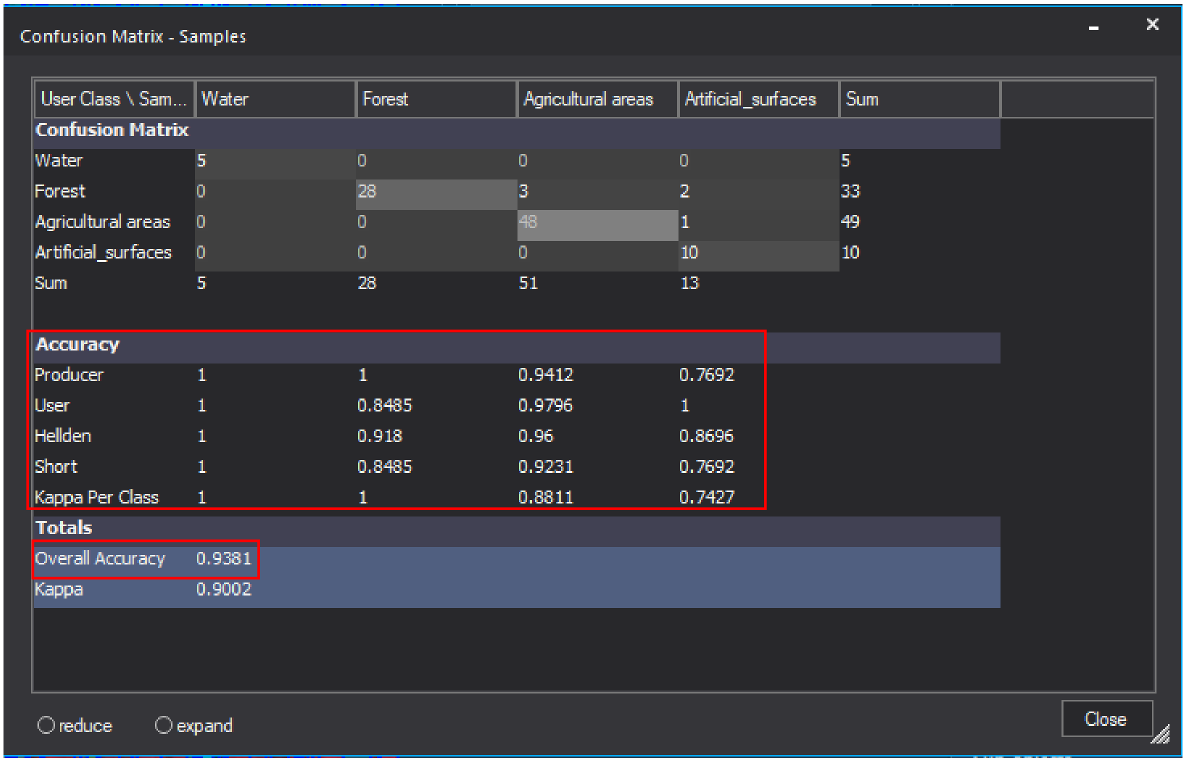


Figure 8. Resulting samples (bottom image). The Upper layer image is the user classification.

Then, the accuracy assessment can be carried out, using the "Calculate a confusion matrix based on the reference data" tool as with eCognition 10.1. The resulting matrix (Figure 9) shows an Overall Accuracy of 93.81%. The following were the User accuracy for all the classes: Water – 100%, Forest – 84.85%, Agricultural Areas – 97.96%, and Artificial Surfaces – 100%. As in the accuracy assessment carried out in eCognition 10.1, the Artificial Surfaces and Water classes had the highest user accuracy, while the Forest class had the lowest.



Confusion Matrix - Samples

User Class \ Sam...	Water	Forest	Agricultural areas	Artificial_surfaces	Sum
Confusion Matrix					
Water	5	0	0	0	5
Forest	0	28	3	2	33
Agricultural areas	0	0	48	1	49
Artificial_surfaces	0	0	0	10	10
Sum	5	28	51	13	
Accuracy					
Producer	1	1	0.9412	0.7692	
User	1	0.8485	0.9796	1	
Hellden	1	0.918	0.96	0.8696	
Short	1	0.8485	0.9231	0.7692	
Kappa Per Class	1	1	0.8811	0.7427	
Totals					
Overall Accuracy	0.9381				
Kappa	0.9002				

reduce expand Close

Figure 9. Resulting Error Matrix based Vector Points.

6.2 Based on Polygon Vector (Object Based)

The first step will be to add the Polygon reference layer (Figure 10).

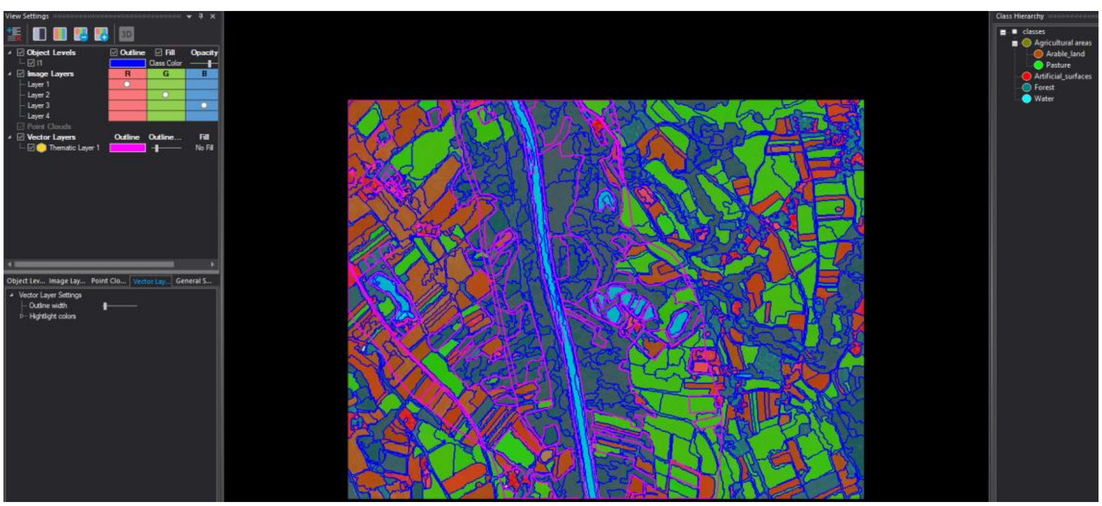


Figure 10. Polygon Reference layer added.

Then, as with the previous accuracy assessment workflows, add the "convert thematic objects to samples" process. However, as we are working with polygon surfaces, we must consider the "Overlap threshold" parameter. This parameter will limit the selection of samples to those mapped objects that overlap the reference polygon in at least the defined threshold. In our case, we defined an overlap threshold of 60%, which means that just the mapped objects whose surfaces overlap at least 60% of the reference polygon, will be assigned as a determined class (based on the "Remap Attribute Values" configuration). (Figure 11)

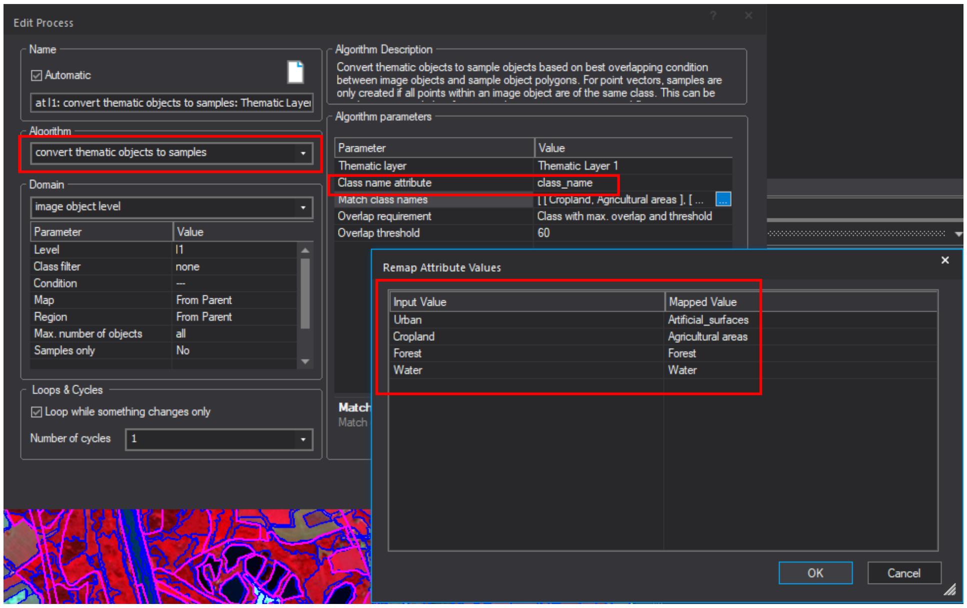


Figure 11. Polygon Reference layer added.

As result, the samples will be displayed the image layer (Figure 12).

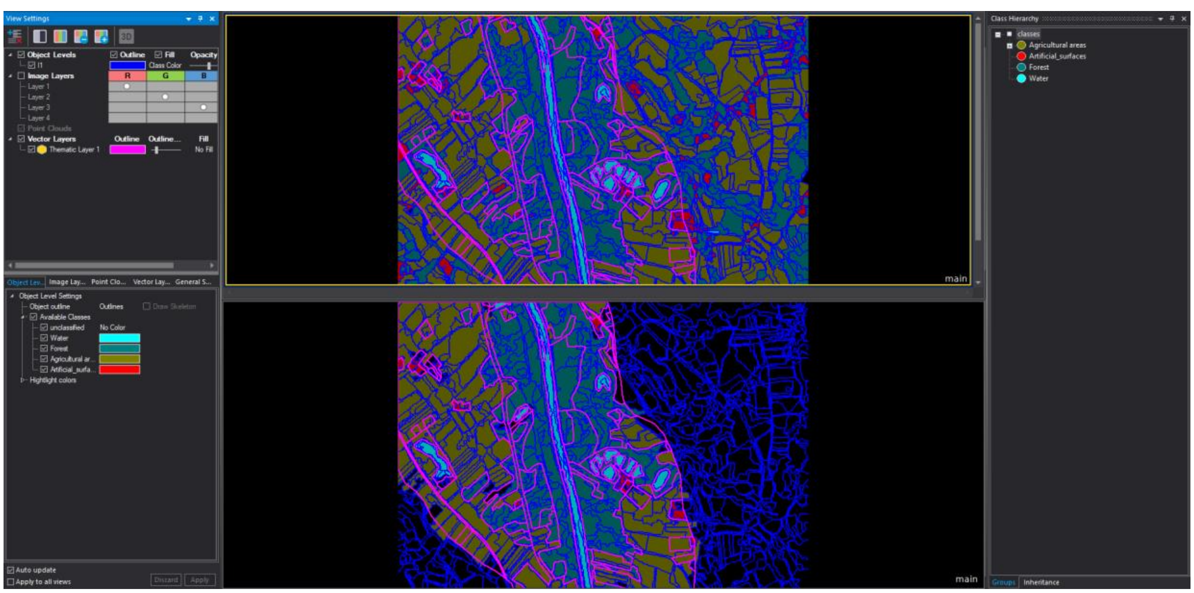
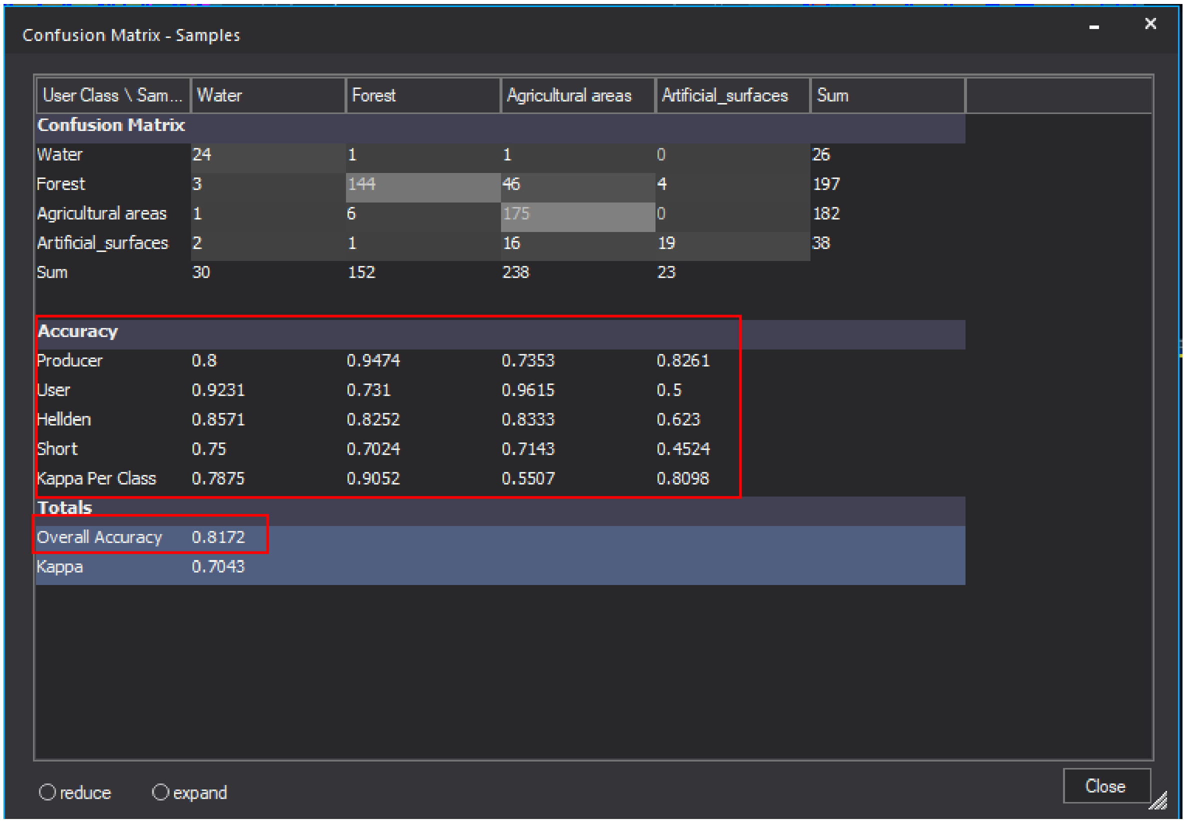


Figure 12. Resulting samples (bottom layer image). The upper layer image is the user classification.

Using the “Calculate a confusion matrix based on the reference data” tool, the Confusion Matrix was generated (Figure 13). It shows an Overall Accuracy of 81.72%. The following were the User accuracy for all the classes: Water – 92.31%, Forest – 73.1%, Agricultural Areas – 96.15%, and Artificial Surfaces – 50%. For this accuracy assessment, the water class had the highest user accuracy, while the Artificial Surfaces class had the lowest.



Confusion Matrix - Samples

User Class \ Sam...	Water	Forest	Agricultural areas	Artificial_surfaces	Sum
Confusion Matrix					
Water	24	1	1	0	26
Forest	3	144	46	4	197
Agricultural areas	1	6	175	0	182
Artificial_surfaces	2	1	16	19	38
Sum	30	152	238	23	
Accuracy					
Producer	0.8	0.9474	0.7353	0.8261	
User	0.9231	0.731	0.9615	0.5	
Hellden	0.8571	0.8252	0.8333	0.623	
Short	0.75	0.7024	0.7143	0.4524	
Kappa Per Class	0.7875	0.9052	0.5507	0.8098	
Totals					
Overall Accuracy	0.8172				
Kappa	0.7043				

reduce expand Close

Figure 13. Resulting Error Matrix based Vector Polygons.

7. ARCGIS PRO

As for eCognition, ArcGIS Pro provides different geoprocessing tools to carry out Accuracy Assessment. In the following sections, the tools “Create Accuracy Assessment Points”, “Update Accuracy Assessment Point”, and “Compute Confusion Matrix” will be introduced for usage with reference points and polygons.

7.1 Based on Polygon Vector (Object Based)

In this section, the accuracy assessment will be carried out based on the provided riparian polygons. The first step is to randomly generate 500 sample points along the reference polygons area (Figure 14) and populate them with the reference Ground truth values.

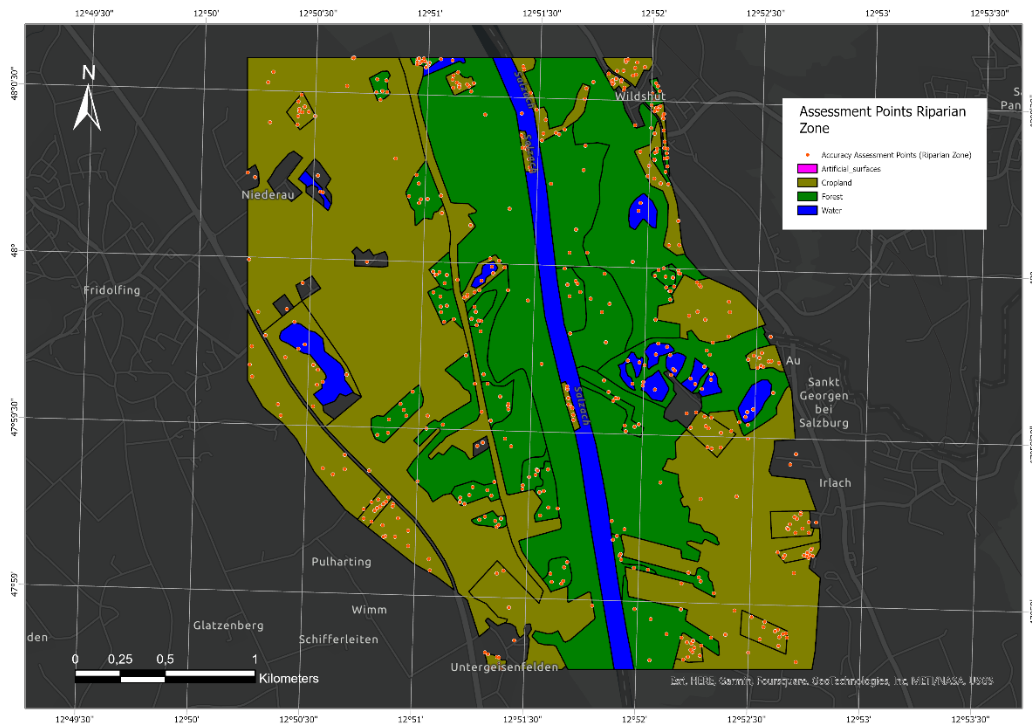


Figure 14. Generated Samples and Riparian polygons.

For that purpose, the "Create Accuracy Assessment Points" was used, with the Stratified random parameter as the "Sample Strategy", this parameter determines that the number of samples for each class will be proportional to its relative area (Figure 15).

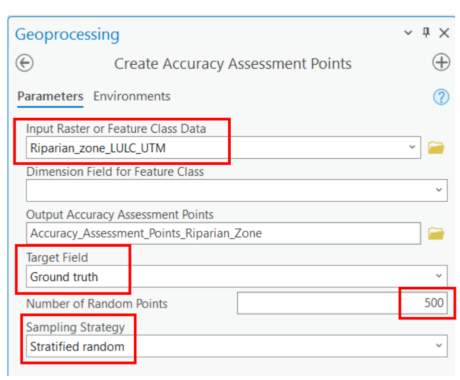
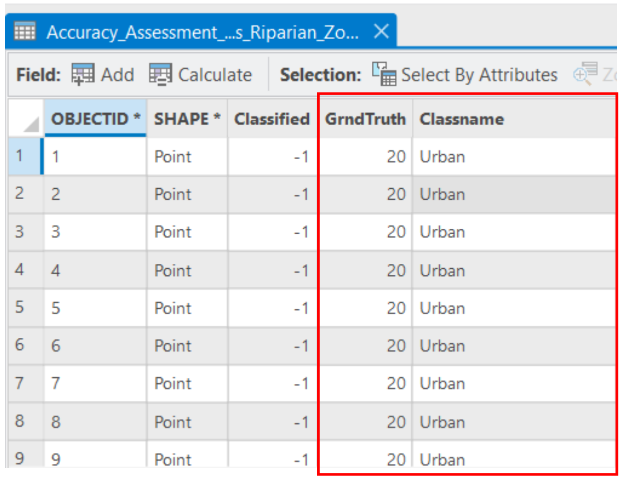


Figure 15. The "Create Accuracy Assessment Points" tool configuration.

As result, 500 samples were generated. Each of the points contains the value of the associated riparian polygon class (Figure 16).



	OBJECTID *	SHAPE *	Classified	GrndTruth	Classname
1	1	Point	-1	20	Urban
2	2	Point	-1	20	Urban
3	3	Point	-1	20	Urban
4	4	Point	-1	20	Urban
5	5	Point	-1	20	Urban
6	6	Point	-1	20	Urban
7	7	Point	-1	20	Urban
8	8	Point	-1	20	Urban
9	9	Point	-1	20	Urban

Figure 16. Output for the "Create Accuracy Assessment Points" tool.

Then, using the "Update Accuracy Assessment points" tool (Figure 17)., the class value (resulting from the user's classification) of each object, will be added (if that's the case) as an attribute to the Assessment points layer (Figure 18).

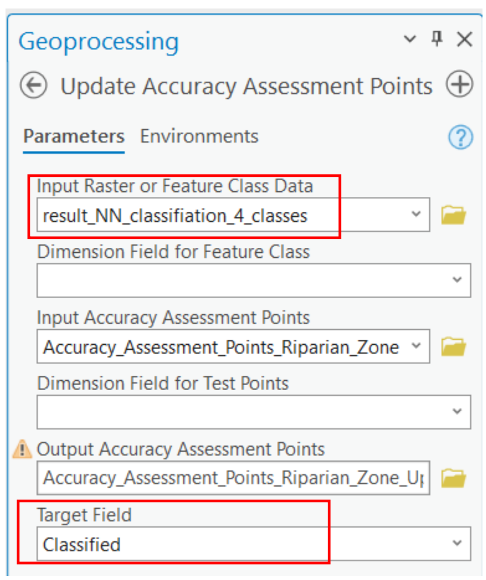


Figure 17. The "Update Accuracy Assessment Points" tool configuration.

As result, for each of the sample points we have not only the ground truth value as an attribute but also the assigned user-classified value. The final step will be to run the "Compute Confusion Matrix" tool, which generates the confusion Matrix.

Accracy_Assessment_...an_Zone_Upd...

Field: Add Calculate Selection: Select By Attributes

	OBJECTID *	SHAPE *	Classified	GrndTruth	Classname
1	1	Point	30	20	Urban
2	2	Point	10	20	Urban
3	3	Point	20	20	Urban
4	4	Point	30	20	Urban
5	5	Point	20	20	Urban
6	6	Point	30	20	Urban
7	7	Point	20	20	Urban
8	8	Point	30	20	Urban
9	9	Point	30	20	Urban

Figure 18. Assessment Points with both User classification and Ground truth values

The resulting matrix (Figure 19) shows an Overall Accuracy of 74.74%. The following were the User accuracy for all the classes: Water – 57.14%, Forest – 69.14%, Agricultural Areas (Cropland) – 90.64%, and Artificial Surfaces (Urban) – 47.05%. For this Accuracy Assessment, the Cropland class had the highest user accuracy, while the Urban class had the lowest.

Confusion_Matrix_Riparian_Zone

Field: Selection: Rows:

	OBJECTID *	ClassValue	C_10	C_20	C_30	C_40	Total	U_Accuracy	Kappa
1	1	C_10 - Cropland	155	1	13	2	171	0,906433	0
2	2	C_20 - Urban	6	8	0	3	17	0,470588	0
3	3	C_30 - Forest	71	9	186	3	269	0,69145	0
4	4	C_40 - Water	18	0	0	24	42	0,571429	0
5	5	Total	250	18	199	32	499	0	0
6	6	P_Accuracy	0,62	0,444444	0,934673	0,75	0	0,747495	0
7	7	Kappa	0	0	0	0	0	0	0,583809

Figure 19. Resulting Error Matrix based Vector Polygons.

7.2 Based on Point Vector (Object Based)

Finally, using the Accuracy Assessment points (Figure 20) already provided and prepared by the tutor (with ground truth and classified values) as the input for the" Compute Confusion Matrix", the confusion matrix was generated.

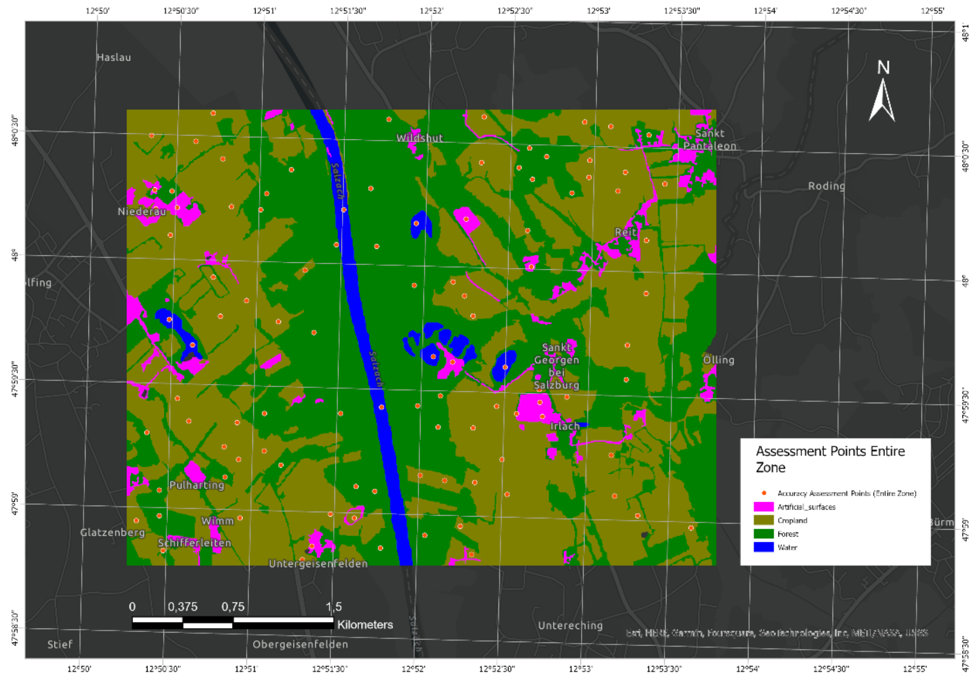


Figure 20. User classification and provided accuracy assessment points.

The resulting matrix (Figure 21) shows an Overall Accuracy of 94.17%. The following were the User accuracy for all the classes: Water – 100%, Forest – 84.84%, Agricultural Areas (Cropland) – 98.11%, and Artificial Surfaces (Urban) – 100%. For this Accuracy Assessment, the Water class had the highest user accuracy, while the Forest class had the lowest.

Confusion_Matrix_Entire_Zone

Field: Selection: Rows:

	OBJECTID *	ClassValue	C_10	C_20	C_30	C_40	Total	U_Accuracy	Kappa
1	1	C_10	52	1	0	0	53	0,981132	0
2	2	C_20	0	10	0	0	10	1	0
3	3	C_30	3	2	28	0	33	0,848485	0
4	4	C_40	0	0	0	7	7	1	0
5	5	Total	55	13	28	7	103	0	0
6	6	P_Accuracy	0,945455	0,769231	1	1	0	0,941748	0
7	7	Kappa	0	0	0	0	0	0	0,906236

Figure 21. Resulting Error Matrix based Vector Points.

8. CONCLUSIONS

In conclusion, even though the accuracy assessment (ACC) is the last step in the presented object-based image classification workflow, it may be the most important one, as it can determine whether the performed classification is good enough or in the worst cases, decide when a project has to start all over again because of the low accuracy.

Regarding the ACC carried out through eCognition and ArcGIS software, they provided us with a confusion matrix that quantitatively informed us on how well our samples and classification performed in comparison to reality. Each software was user-friendly and highly responsive to the demanded processes.

Finally, based on the different “Overall Accuracy” and User Accuracy percentages obtained from the different ACC, we can evidence that the accuracy which is mainly determined by the user classification can be influenced by the geometry of the reference data (points or polygons), and the methods used to generate the samples and to carry out the AC. This means that prior to performing the ACC, we must consider which reference data we have and what’s the best workflow to be applied.

9. REFERENCES

Accuracy Assessment Tools. Documentation eCognition Suite (Developer User Guide) (n.d.). Retrieved January 7, 2023, from https://docs.eCognition.com/v9.5.0/eCognition_documentation/User%20Guide%20Developer/11%20Accuracy%20Assessment.htm#:~:text=Error%20Matrices,type%20used%20for%20accuracy%20assessment.&text=Test%20areas%20are%20used%20as,satellite%20imaging)%20based%20on%20pixels

Compute Confusion Matrix (image analyst). Compute Confusion Matrix (Image Analyst)-ArcGIS Pro | Documentation. (n.d.). Retrieved January 7, 2023, from https://pro.arcgis.com/en/pro-app/latest/tool-reference/image-analyst/compute-confusion-matrix.htm

Create Accuracy Assessment Points (image analyst). Create Accuracy Assessment Points (Image Analyst)-ArcGIS Pro | Documentation. (n.d.). Retrieved January 7, 2023, from https://pro.arcgis.com/en/pro-app/latest/tool-reference/image-analyst/create-accuracy-assessment-points.htm

Introduction to the Accuracy Assessment workflow. (2021). *YouTube*. Retrieved January 7, 2023, from https://www.youtube.com/watch?v=7jkORYYoXEg.

Update Accuracy Assessment Points (image analyst). Update Accuracy Assessment Points (Image Analyst)-ArcGIS Pro | Documentation. (n.d.). Retrieved January 7, 2023, from https://pro.arcgis.com/en/pro-app/latest/tool-reference/image-analyst/update-accuracy-assessment-points.htm