Analysis and application of interactive supervised classification in Broome County, NY

Xingyu Yuan
Geography Department, Binghamton University, Binghamton, New York, 13905, U.S.
734106211@qq.com

ABSTRACT. Land cover classification is essential in urban sprawl. It helps the urban planners and the policy makers to understand their urban communities. Interactive supervised classification (ISC) function in ArcMap can produce land cover classification map. Based on the analysis and comparison of the accuracy of ISC map and land cover classification products, GLC, GLOBC, and MCD, ISC map has the highest user’s accuracy in identifying developed urban area, forest, water body and crop/pasture, and highest overall accuracy, which provides efficient technical support for urban planning and decision-making.

1. INTRODUCTION
Monitoring urban sprawl and identifying land cover type are concerns of urban planners and policy makers (Keranen and Kolvoord, 2014). There are many products that provide land cover classification which is used to monitor the urban changes. The classification products provide different results in the same location because of the potential errors, such as spectral data quality, humidity and resolution (Hansen, et al, 2000). The higher resolution of the image generally provides a more detailed map. In order to examine the accuracy of the classification products, accuracy assessment is applied to the results (Foody, 2002). Accuracy assessment provides producer’s accuracy, which occurs when we have omitted certain categories that actually exist on the ground, user’s accuracy, which occurs when we have identified categories that do not exist on the ground, and finally, overall accuracy, which is the total number of correct pixels (sum of the diagonal) divided by the total number of pixels in the error matrix.

Interactive supervised classification (ISC) is a tool provided by ArcGIS10.3.1 to generate land cover classification map based on given aerial image and sample choices (ESRI.com). In this paper, three classification products, GLC, GLOBC, and MCD, are examined and compared with the ISC map in Broome County to see the performance of each product. The results are examined based on the accuracy assessment, and the reference is the high resolution map of Google Earth in 2010.

2. STUDY AREA
Broome County, NY is chosen as the study area to examine the performance of ISC map and the three classification products, GLC, GLOBC, and MCD. Broome County is on the south line of New York. The south border abuts the border of Pennsylvania. Susquehanna River flows through the eastern part of the county. There are 22 cities and towns in this county, and most of them lie in the north and west part of the county. The classification of urban area mainly includes water body, developed urban area, crop/pasture, and forest. The study area, Broome County, includes the mentioned four types of land cover above.
3. METHODOLOGY

3.1 ISC MAP
The aerial image of Broome County is retrieved from Landsat 4-5 from USGS, and it was taken in early 2011. Interactive supervised classification function in ArcMap 10.3.1 is used on the image. Four general land cover type are determined, developed urban area (grey), water body (blue), forest (green), and crop/pasture (brown), and more than 50 samples of each classification are drawn to generate the ISC map (Figure 2).

3.2 100 RANDOM POINTS
Then, 100 random points are generated by using random points function in ArcMap 10.3.1 within Broome County as the points to complete the accuracy assessment (Figure 3). These points are compared with the high-resolution Google Earth map to identify the land cover type in 2010. In these 100 random points, 31 of them are developed (grey), 33 of them are forest (green), 30 of them are crop/pasture (brown), and 6 of them are water (blue).

### 3.3 CLASSIFICATION PRODUCTS
There are three classification products used to be examined, MCD, GLOBC, and GLC (Table 1). GLC and GLOBC are produced by European Commission Joint Research Center with resolutions of 988m and 309m respectively. MCD is produced by University of Wisconsin, Boston University with a resolution of 463m. There are 9 types of land cover classifications in MCD (Figure 4), 21 types of land cover classifications in FLOBC (Figure 5), and 4 types of land cover classifications in GLC (Figure 6). The three products have different classification type, but they can fit the four general land cover classification that is made for ISC map, developed, water, forest, and crop/pasture.

**Table 1: Information of classification products**

| Map                | Producer                                   | Resolution |
|--------------------|--------------------------------------------|------------|
| GLC                | Global Land Cover 2010                      | 988m       |
| GLOBC              | GlobCover 2010                              | 309m       |
| MCD                | MODIS Urban Land Cover 2010                 | 463m       |

![Figure 4: MCD in Broome County](image1)

![Figure 5: GLC in Broome County](image2)

![Figure 6: GLOBC in Broome County](image3)

These products provide land cover classification. Error matrix and accuracy assessment can be made from the comparison between the classifications and the 100 points. The classification products are expected to provide more accurate land cover classification with higher resolution. There are two products with close resolution, GLOBC with 309m resolution, and MCD with 463m resolution, and another product, GLC, with a large difference in resolution, 988m. The reason for this is to analyze if the products with higher resolution are more accurate.

### 3.4 ERROR MATRIX AND ACCURACY ASSESSMENT
Error matrix and accuracy assessment are made based on the comparison between the 100 points, the high resolution map of Google Earth in 2010, the classification products and ISC map (Table 2). The points that match the reality in the classified map are yellow.
Producer’s accuracy is the map accuracy from the point of view of the map maker (the producer). This is how real features correctly shown on the classified map, or the probability that a certain land cover is classified as such. User’s accuracy is the accuracy from the point of view of a map user, which tells how the class on the map will be shown on the ground, which is referred as the reliability. Overall accuracy tells what proportion of the sites is mapped correctly.

4. ANALYSIS

ISC map has an overall accuracy of 0.83. The producer’s accuracies of developed urban area, forest, crop/pasture, and water are 0.94, 0.73, 0.83, and 0.83 respectively. The user’s accuracies of developed urban area, forest, crop/pasture, and water are 0.94, 0.83, 0.74, and 0.83 respectively.

The overall accuracies for MCD, GLC, and GLOBC are 0.56, 0.39, and 0.48 respectively. The range of user’s accuracy in identifying developed urban area, forest, crop/pasture, and water for MCD, GLC, and GLOBC is between 0 and 0.64. The range of producer’s accuracy in developed urban area, forest, crop/pasture, and water for MCD, GLC, and GLOBC is between 0 and 0.9.

The overall accuracy of ISC map is the highest in the four classification methods. The ISC map can also identify the land cover more reliable than the other products. Although there are a few classifications of the products have higher producer’s accuracy than the ISC map, for example, the

| Table 2: Accuracy assessment of ISC, GLC, GLOBC, and MCD |
|---------------------------------|------------------|------------------|------------------|------------------|------------------|
| ISC Google Earth                | developed        | forest           | crop/pasture     | water            | total            | user’s accuracy  |
| developed                       | 29               | 1                | 0                | 1                | 31               | 0.94             |
| forest                          | 0                | 24               | 5                | 0                | 29               | 0.83             |
| crop/pasture                    | 2                | 7                | 25               | 0                | 34               | 0.74             |
| water                           | 0                | 0                | 5                | 0                | 6                | 0.83             |
| total                           | 31               | 33               | 30               | 6                | 100              |                  |
| producer’s accuracy             | 0.94             | 0.73             | 0.83             | 0.83             |                  |                  |
| overall accuracy                | 0.83             |                  |                  |                  |                  |                  |

| MCD Google Earth                | developed        | forest           | crop/pasture     | water            | total            | user’s accuracy  |
| developed                       | 9                | 8                | 0                | 2                | 19               | 0.47             |
| forest                          | 10               | 20               | 1                | 2                | 33               | 0.61             |
| crop/pasture                    | 10               | 3                | 27               | 2                | 42               | 0.64             |
| water                           | 2                | 2                | 2                | 0                | 6                | 0.00             |
| total                           | 31               | 33               | 30               | 6                | 100              |                  |
| producer’s accuracy             | 0.29             | 0.61             | 0.90             | 0                |                  | 0.56             |
| overall accuracy                |                  |                  |                  |                  |                  |                  |

| GLC Google Earth                | developed        | forest           | crop/pasture     | water            | total            | user’s accuracy  |
| developed                       | 6                | 5                | 7                | 2                | 20               | 0.30             |
| forest                          | 13               | 20               | 10               | 2                | 45               | 0.44             |
| crop/pasture                    | 8                | 6                | 13               | 2                | 29               | 0.45             |
| water                           | 4                | 2                | 0                | 0                | 6                | 0.00             |
| total                           | 31               | 33               | 30               | 6                | 100              |                  |
| producer’s accuracy             | 0.19             | 0.61             | 0.43             | 0                |                  | 0.39             |
| overall accuracy                |                  |                  |                  |                  |                  |                  |

| GLOBC Google Earth              | developed        | forest           | crop/pasture     | water            | total            | user’s accuracy  |
| developed                       | 6                | 2                | 9                | 0                | 17               | 0.35             |

| 4 |
The accuracy of the products is expected to be more accurate with a higher resolution. In the three products, GLOBC has a higher resolution and GLC has the lower resolution, however, the MCD has a highest overall accuracy, and higher user’s accuracy, which shows more reliability than others. Therefore, the accuracy of the classification product might not be heavily affected by the resolution within the study area.

5. CONCLUSION

Urban land use analysis is an important part of urban planning and management, especially the application of remote sensing technology is an important symbol of urban modernization and scientific management. ISC map has the highest overall accuracy and high-level user’s accuracy; therefore, it is a more efficient and more reliable method for the urban planners and policy-makers to use to monitor urban sprawl. MCD, GLC, and GLOBC have the lowest accuracy in identifying water bodies, thus, the applicability of their resolution in the study area requires more scientific research.

REFERENCES

[1] “Esri.” GIS Mapping Software, Location Intelligence & Spatial Analytics, www.esri.com/en-us/home.
[2] Foody, Giles M. “Status of Land Cover Classification Accuracy Assessment.” Remote Sensing of Environment, vol. 80, no. 1, 2002, pp. 185–201., doi:10.1016/s0034-4257(01)00295-4.
[3] Hansen, M. C., et al. “Global Land Cover Classification at 1 Km Spatial Resolution Using a Classification Tree Approach.” International Journal of Remote Sensing, vol. 21, no. 6-7, 2000, pp. 1331–1364., doi:10.1080/014311600210209.
[4] Joint Research Centre. (2020, April 30). Retrieved from https://ec.europa.eu/info/departments/joint-research-centre_en
[5] Keranen, K., & Kolvoord, R. (2014). Making spatial decisions using Gis and remote sensing: a workbook. Redlands, CA: ESRI Press.
[6] MODIS Web. (n.d.). Retrieved from https://modis.gsfc.nasa.gov/data/dataprod/mod12.php
[7] Potere, D., Schneider, A., Angel, S., & Civco, D. L. (2009). Mapping urban areas on a global scale: which of the eight maps now available is more accurate?. International Journal of Remote Sensing, 30(24), 6531-6558.
[8] Usgs. Home. Retrieved from https://earthexplorer.usgs.gov/