Integrating Python and ArcGIS to create a suite of tools for land cover classification and accuracy assessment on the Madre de Dios region, Peru

Authors: Andrea Nicolau[1,2] (andrea.puzzinicolau@nasa.gov), Africa Flores[1,2], Kel Markert[1,2], Kelsey Herndon[1,2], Robert Griffin[1] | April, 2018
[1]University of Alabama in Huntsville - Earth System Science Center, [2]NASA SERVIR

Introduction: In this project, a suite of tools was created using ArcPy and ModelBuilder for a rapid land cover classification based on spectral indices for a pre- and post- scenes in the Madre de Dios region where illegal mining sites have been increasing over the years (Swenson et al., 2011; Elmes et al., 2014; Asner et al., 2017).

Methods: The toolbox created consists in three python scripts written in IDLE along with a model built in ModelBuilder. With Landsat scenes as inputs, the toolbox first calculates the Modified Normalized Difference Water Index (NDWI), and the Bare Soil Index (BSI). Secondly, it creates land cover rasters with 3 classes (Vegetation, Bare Soil, and Water), then it does the sampling and an unsupervised classification based on a reference raster, and finally outputs csv files with the overall accuracy of classification. The user can define the Landsat scenes, the thresholds of each index for the classification, the study area, the number of random points for each class in the sampling method, and the reference data to be used as the truth.

Results: In this case, one Landsat 5 TM scene from August 2008 and one Landsat 8 OLI scene from August 2014 were used, 30 random points were generated for each of the 3 classes, the reference raster was downloaded and reclassified from the GEOBOSQUES website and the overall accuracy for the pre- and post- scenes were 81% and 73%, respectively. The results indicated an increase in bare soil extent from 21,000 ha in 2008 to 33,000 ha in 2014, and a decrease in vegetation extent from 540,000 ha in 2008 to 523,000 ha in 2014, approximately.

Conclusion: Although this suite of tools demonstrated to be an easy-to-use way to generate land cover maps and the use of spectral indices demonstrated to be satisfactory method for classification, the lack of a reference raster as the truth may compromise the applicability of the toolbox.

Future Work: A comparison between other land cover classification methods, e.g.: Random Forests and Spectral Mixture Analysis.

References: Asner, P.; Tupayachi, R. (2017), "Accelerated losses of protected forests from gold mining in the Peruvian Amazon", Environmental Research Letters, 12. https://doi.org/10.1088/1748-9326/aa7dab; GEOBOSQUES (2018), "Bosque y no bosque 2000 - Pérdida de bosque 2001 - 2016 (raster)". http://geobosques.minam.gob.pe/geobosque/view/descargas.php#download; USGS, EarthExplorer; Swenson, J. J., et al. (2011), "Gold Mining in the Peruvian Amazon: Global Prices, Deforestation, and Mercury imports", PLoS ONE 6(4). doi:10.1371/journal.pone.0018875. Elmes, A., et al. (2014), "Mapping licit and illicit mining activity in the Madre de Dios region of Peru", Remote Sensing Letters 5:10, 882-890. http://dx.doi.org/10.1080/2150704X.2014.973080

Funding for this research provided by NASA SERVIR through NASA Cooperative Agreement NNM11AA01A

Workflow: