Mapping Impervious Surfaces Globally at 30m Resolution Using Landsat Global Land Survey Data

Impervious surfaces, mainly artificial structures and roads, cover less than 1% of the world's land surface (1.3% over USA). Regardless of the relatively small coverage, impervious surfaces have a significant impact on the environment. They are the main source of the urban heat island effect, and affect not only the energy balance, but also hydrology and carbon cycling, and both land and aquatic ecosystem services. In the last several decades, the pace of converting natural land surface to impervious surfaces has increased. Quantitatively monitoring the growth of impervious surface expansion and associated urbanization has become a priority topic across both the physical and social sciences.

The recent availability of consistent, global scale data sets at 30m resolution such as the Global Land Survey from the Landsat satellites provides an unprecedented opportunity to map global impervious cover and urbanization at this resolution for the first time, with unprecedented detail and accuracy. Moreover, the spatial resolution of Landsat is absolutely essential to accurately resolve urban targets such a buildings, roads and parking lots. With long term GLS data now available for the 1975, 1990, 2000, 2005 and 2010 time periods, the land cover/use changes due to urbanization can now be quantified at this spatial scale as well.

In the Global Land Survey – Imperviousness Mapping Project (GLS-IMP), we are producing the first global 30 m spatial resolution impervious cover data set. We have processed the GLS 2010 data set to surface reflectance (8500+ TM and ETM+ scenes) and are using a supervised classification method using a regression tree to produce continental scale impervious cover data sets. A very large set of accurate training samples is the key to the supervised classifications and is being derived through the interpretation of high spatial resolution (~2 m or less) commercial satellite data (Quickbird and Worldview2) available to us through the unclassified archive of the National Geospatial Intelligence Agency (NGA). For each continental area several million training pixels are derived by analysts using image segmentation algorithms and tools and then aggregated to the 30m resolution of Landsat. Here we will discuss the production/testing of this massive data set for Europe, North and South America and Africa, including assessments of the 2010 surface reflectance data. This type of analysis is only possible because of the availability of long term 30m data sets from GLS and shows much promise for integration of Landsat 8 data in the future.

Mapping Impervious Surfaces Globally at 30m Resolution Using Global Land Survey Data

B41D-0442

Eric Brown de Colstoun[1], Chengquan Huang[2], Robert Wolfe[1], James Tilton[1], Bin Tan[1,4], Sarah Smith[1,3], Jacqueline Phillips[1,3], Panshi Wang[2], Pui-Yu Ling[2], James Zhan[2], Sike Li[2], and Michael Taylor[1,5]

1 NASA Goddard Space Flight Center, Greenbelt MD
2 Department of Geographical Sciences, University of Maryland
3 Goddard Earth Science Technology and Research, USRA
4 Earth Resources Technology, Inc.
5 Sigma Space Corp.

Introduction

- Since about 2008, the U.N. estimates more people live in cities than rural areas. Higher urban growth rates are expected in the developing world in the next 30 years.
- Cities still represent a relatively small ‘footprint’ globally (~3% of land area).
- However, the process of urbanization is most often irreversible, modifying carbon, water, energy cycles at various spatial scales.
- New global scale data sets from Landsat (GLS:Global Land Survey) provide great opportunities to map and monitor urbanization at the APPROPRIATE spatial scale, with future integration of Landsat 8 data to the present.

Project Objectives:

1. Produce Global, 30m resolution surface reflectance data from Landsat for 2010 using Global Land Survey Data.

2. Produce the first Global scale, 30m resolution percent impervious cover data sets for 2000 and 2010 and assess areas of significant urbanization in the 2000-2010 period.

Methods

Our overall approach is essentially a supervised ‘classification’ approach using a continental/global archive of very high resolution commercial satellite data for training, Hierarchical Image Segmentation analysis tools for training generation, and a regression tree algorithm for continuous value output at continental scales.

Obtain Imagery from the National Geospatial-Intelligence Agency (NGA)

The NGA provides access to federal U.S. government users to a global archive of unclassified high-resolution commercial satellite imagery. These data make it possible to derive training at continental scales and to extrapolate to the entire world using the GLS data. We look for scenes that are cloud free, between the years 2009 and 2011, represent periods of leaf-on vegetation, with view angles close to nadir, and high sun angles. As much as possible, the images are spatially distributed to represent the geographical variations present in the landscape.

Orthorectification

The NGA data are not orthorectified and need be matched up with the GLS data. The high resolution images are orthorectified with the Rational Polynomial Coefficients (RPC) model [1]. This model is an alternative solution for the rigorous physical sensor model. The RPC model performs transformations between image and object spaces through a ratio of two cubic polynomials. All required coefficients are embedded in the meta data of the high resolution data. See Figure 1 below.

Figure 1. Training site for the mountainous city of Nuestra Senora de La Paz, Bolivia with original image overlaying orthorectified image at 60% transparency. Note the difference between sections 1 and 2 of the training site. The orthorectification process made significant adjustments to section 2 which lies lower in the mountain valley.

Subset in ENVI

Each newly re-projected GeoTIFF is subset into one to three 2500 x 2500 pixel sub-scenes for interpretation into high quality impervious/non-impervious training.

Hierarchical Image Segmentation Functions/Tools

All subsets are then batch processed using the *HSeg* Software which is a form of region growing segmentation that directly forms a segmentation hierarchy based on user interaction. We have developed an in-house tool called *HSeglearn* that allows spatially disjoined region classes to be merged and results are continuously updated as the analyst submits binary selections of either impervious or non-impervious features. *HSegViewer* is used to fine-tune classifications by allowing the user to manipulate the labeling of segmentation hierarchies and as a final QA tool. Errors commonly encountered include misclassified roads, open pit mines, bare fields, etc.

- The data are then aggregated to 30m resolution and matched to the GLS-2010 surface reflectance data for training.
- The training pixels for the entire continent (~2.63M for Europe, ~5.08M for North America, 1.93M for South America and 3.54M for Africa) are used within the Cubist Regression Tree algorithm to create percent impervious cover product for each GLS-2010 scene (see Fig. 4).

Results

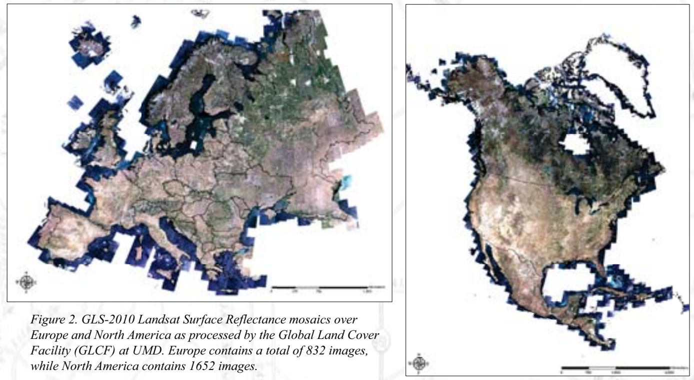

Figure 2. GLS-2010 Landsat Surface Reflectance mosaics over Europe and North America as processed by the Global Land Cover Facility (GLCF) at UMD. Europe contains a total of 832 images, while North America contains 1652 images.

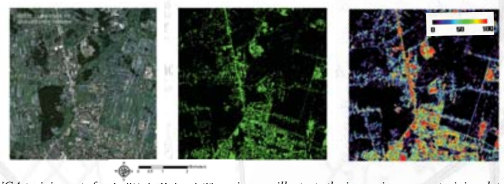

Figure 3. NGA training sets for Jedlinsk, Poland. These images illustrate the impervious cover training data generation process. Very high spatial resolution Quickbird or Worldview-2 NGA imagery is classified and interpreted into impervious (green) and non-impervious (black) classes using Hierarchical Image Segmentation. The interpreted data are then aggregated to 30m resolution to derive percent impervious cover training pixels.

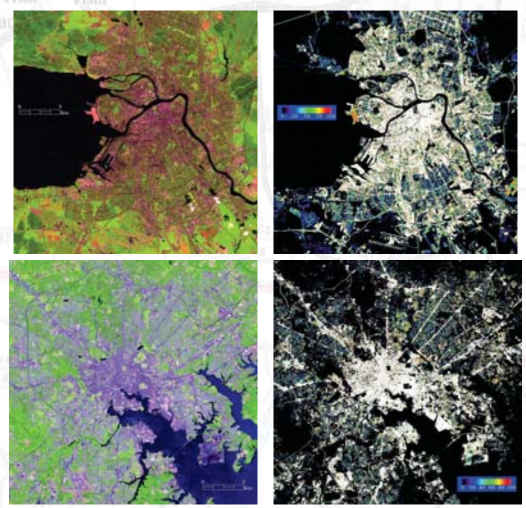

Figure 4. GLS-2010 Surface Reflectance subset over St. Petersburg, Russia (top left) and Baltimore, MD (bot. left). Right panels show derived percent impervious cover results using ~1.5M training pixels from NGA data over Europe and ~4M for North America and Cubist regression tree. Results over dense urban centers are quite good but errors of commission over bare fields and crops are still a challenge due to phenology variations in the GLS-2010 imagery. Landsat bands 5, 4, and 2 are coded as R, G, and B, respectively.

Progress & Status

- Processing of 8555 Landsat TM and ETM+ data to surface reflectance for GLS-2010 (Figure 2)
- Developed tools to significantly facilitate the work of analysts during production of training data using *HSeg*
- Enhanced the process and quality of training significantly compared to previous methods used by the PIs
- Ordered and reviewed a global archive of over 1,800 high-res images
- Completed training for Europe, North America, Africa, and South America (See Fig. 5)
- Exploring the uses of GIS tools, data, and services as an effective tool for QA and also for identification and removal of false positives
- Testing methods for improving training data quality using regression and decision tree classifiers, outlier identification and filtering, etc. (Our training data is dominated by low impervious cover samples which can bias regression trees.)
- Developed and executed QA tools and metrics for training data.
- Automated image orthorectification for better co-registration of training and GLS-2010 data.

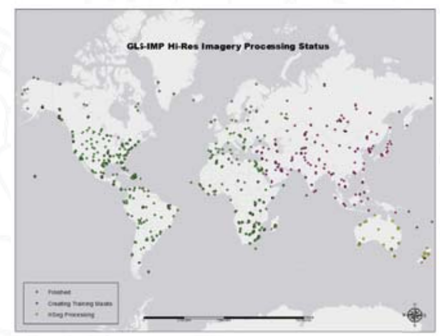

Figure 5. Status of training data selection and processing for the GLS-IMP Project. Over 1800 Quickbird and Worldview-2 scenes have been acquired to date.

Next Steps

- Apply continental masks of water, shadow, cloud, and snow.
- Automate training data outlier removal.
- Reduce commission errors through ancillary data.
- Complete and assess continental scale results for Europe/North America.
- Create training data for Asia and process Australia in Hseg.

This project is funded under NASA’s Land Cover Land Use Change Program

References

[1] Kaichang Di, Ruijin Ma, and Rong Xing Li. Rational functions and potential for rigorouos sensor model recovery. Photogrammetric Engineering & Remote Sensing, Vol 69 pp33-41, 2003.