Wiscland-2 Quarterly Project Update 4

Prepared by Wiscland-2 Project Team
Contact: Howard Veregin
University of Wisconsin-Madison
howard.veregin@wisc.edu

December 31, 2015

Executive Summary
The objective of the Wiscland-2 project is to update the Wisconsin statewide land cover map. During this quarter (October-December 2015), the Project Team has been primarily involved in producing the Level 1 statewide map. In the upcoming quarter, planned tasks are largely focused on continuing the classification process by moving into the creation of higher level (2-4) maps.

Introduction

The purpose of this document is to update the Wiscland-2 Land Cover Guidance Team on the progress of the project in the previous quarter. Since issuing the previous quarterly report in September 2015, the tasks outlined in the "Current Tasks" section have been completed as follows:

- Determined features and variables for classification.
- Developed post-classification workflows for segmenting, mosaicking, and refining map classifications.
- Created Level 1 map classifications for each footprint.
- Evaluated the possibility of digitizing nursery/Christmas tree farms, and with the support of the Guidance Team, determined it is outside of the current scope of the project.

Results

Results related to <u>classification research and development:</u>

1) Assessed methods to improve to accuracy of Shrubland class.

The most difficult class to achieve consistent accuracies of >85% is the Shrubland class. The Project Team conducted research under several areas to eventually increase Shrubland accuracies from an average of 35% to an average of 85%.

First, the Project Team presented the issue to the Science Advisory Committee during this quarter, and received recommendations to investigate the utility of different texture features for discriminating between Shrubland and other similar classes. Texture metrics were calculated using three sources of imagery: Landsat 15m panchromatic (pan) band for four dates of imagery (winter, spring, summer, fall), 2013 NAIP pan, and 2010 WROC pan images. The best performing grey-level co-occurrence texture metrics (GLCM) were chosen for each image source by analyzing class separability for each metric.

Each metric was resampled to match the 30m resolution of the feature data, and tested as additional features in the classification. Unfortunately, none of these features were able to significantly improve the classification confusion between Shrubland and the Grassland, Forest, and Wetland classes. A review of the separability measurements of these texture features showed that, although several of the metrics have significant discriminatory power between Shrubland and other classes, they provided redundant information and ultimately did not improve the final map accuracy of any class.

Next, we briefly assessed whether incorporating the probability estimates from the classification results into a post-classification process would be feasible, but determined that the probability values were too variable across the state to determine a universal rule.

Lastly, we tested variations on the sampling technique to adjust the class distribution of training and testing data. Here, we were able to make sampling adjustments that maintained the shape of the class distribution, so that the number of samples is roughly proportional to its area on the ground, but increase the Shrubland points. Adjustments also included a more even sampling across the sample data sources in order to increase representation of different Shrubland subtypes. Doing so, Shrubland and overall accuracy increased statewide, with many areas surpassing the 85% benchmark for either or both precision and recall.

Results related to <u>Level 1 map classification:</u>

1) Developed and implemented workflow to classify areas outside of field collection area.

Although the vast majority of the state is covered by 11 Landsat footprints, there are two very small areas outside of these scenes: Washington and Rock Islands in Door County, several of the Apostle Islands and a portion of Bayfield County in northern Wisconsin. These areas were too small and remote to have sufficient field data collected, so we developed a workflow to use spectral signature extension for the classification. The signature extension utilizes sample data from the adjoining footprint along the same path to augment the sample data for the target footprint. The data is parsed to only include features (e.g. dates of imagery) and land cover classes that are shared between the two footprints. In doing so, fewer features are used but we increase the number of available sample points. Each of the areas was successfully classified using this approach.

2) Developed and implemented post-processing workflow.

After the pixel-based classification of each of the 13 Landsat footprints, several post-processing procedures were developed and implemented: 1) compositing areas of footprint overlap, 2) mosaicking footprints to statewide, and 3) segmenting and filtering the statewide map results.

Compositing and mosaicking

Because of orbital track of the Landsat satellite, each footprint, and consequently each footprint map, overlaps with up to 2 other footprints along each edge. Therefore, pixels in those overlapping areas may have two or three labels from different classifications and those labels may conflict. To determine the final label, we developed a compositing procedure to choose the best label according to the classifier.

As part of the classification process, the algorithm provides information on the probability that the pixel belongs to each of the available classes. During the compositing procedure, this probability value is retrieved for the mapped class and compared to the probability of the mapped class in the other footprint(s). The label with the highest probability for that pixel is chosen as the final map label.

After compositing each overlap area, the footprints are reprojected to Wisconsin Transverse Mercator (HARN), mosaicked into a statewide map, and clipped to the state boundary.

Segmentation and filtering

After producing the pixel-based classification results, we introduce an image segmentation process. Image segmentation partitions an image into segments based on the spectral (dis)similarity of adjacent pixels. Therefore, segmentation can be used to identify natural boundaries between landscape features and incorporate contextual information into the classification.

To create the segments, we used a multi-band image stack containing NAIP RGB-IR and WROC RGB imagery, both resampled to 10m resolution to decrease computing requirements. Segments were produced using an open-source mean-shift algorithm from the Orfeo Toolbox library and restricted to a minimum size of two acres to meet the minimum mapping unit.

The segments were integrated with the statewide map by applying a majority rule over each segment: the class that made up the majority of each segment was assigned to the entire segment. After this segmentation process, we used a sieve block to filter out spurious pixel and re-inforce the two-acre minimum mapping unit, creating the final statewide map.

3) Developed and implemented statewide accuracy assessment process.

Individual cross-validation results were created for each footprint as part of developing the classification process. However, the final map product requires an accuracy assessment be performed after the post-classification processing (mosaicking, segmentation, etc.) as well. During the classification process, a subset of the data for each class in each footprint was retained (outside of the data used in training the classifier) for this accuracy assessment. We augmented the workflow to compile these points, intersect the points with the statewide ‘final’ map, and finally create a confusion matrix and accuracy metrics for the statewide accuracy assessment. We implemented this workflow for the Level 1 interim product for the December 31st interim delivery.

Accuracy metrics for the Wiscland 2 Level 1 interim product		
Class	**Precision**	**Recall**
Urban	0.95749	0.89349

Grassland	0.88599	0.91565
Forest	0.87131	0.91773
Water	0.97108	0.93352
Wetlands	0.9121	0.89224
Barren	0.91169	0.86881
Shrubland	0.94807	0.70242
Overall Accuracy	0.90151	

The Project Team is currently having discussions regarding the accuracy assessment methodology and may be adjusting the technique slightly in the coming weeks. Updated statistics will be provided with any subsequent assessments.

Conclusions

During this quarter, the SCO has completed tasks related to creating the final Wiscland 2 land cover products. The overall accuracy of this Level 1 product is estimated at 90.2%, with class accuracies ranging from 87.1 to 97.1% Precision and 70.2 to 93.4% Recall. Shrubland accuracy continues to be the lowest performing class in general. Adding texture features to the classification failed to yield any significant increases in the accuracy although modifying the sampling schemes has improved the accuracy across the state. The Barren class also has lower accuracy, mostly due to spectral confusion between Barren and Urban areas. The Level 1 output informs the creation of the higher level maps, so a high-quality output at this level is important to the overall final product. In order to get a deeper review of the map results, this Level 1 initial classification was forwarded to the DNR for review while the Project Team continues to refine the workflows needed for the creation of a Level 2 map.

Current

The Project Team is currently focusing efforts on:

- Creating level 2, 3, and 4 map classifications for each footprint.
- Reviewing accuracy assessment methodology.
- Responding to DNR reviewer feedback on the interim products.

Upcoming

There are several priority tasks currently being worked on for the next quarter:

- Creating level 2, 3, and 4 classification maps for each footprint.
- Creating draft versions of the User's Guide and documentation that will be included in the final product delivery.
- Continuing outreach and coordination with other land cover mapping projects.