

# Inventory & Monitoring

# Remote Sensing Tips



United States Department of Agriculture  
Forest Service—Engineering  
Remote Sensing Applications Center



San Dimas Technology & Development Center

November 2002 RSAC-4002-TIP1

## Analysis of Canopy Cover and Impervious Surface Cover of Zone 41

**Bonnie Ruefenacht, Alex Hoppus, Jule Caylor**  
*Remote Sensing Applications Center, Salt Lake City, UT*

**David Nowak, Jeff Walton**  
*Northeast Research Station, Syracuse, NY*

**Limin Yang**  
*US Geological Survey, EROS Data Center, Sioux Falls, SD*

**Greg Koeln**  
*Earth Satellite Corporation, Rockville, MD*

**Prepared for: Inventory & Monitoring Steering Committee, Bob Simonson**  
*San Dimas Technology & Development Center, San Dimas, CA*

### Introduction

Satellite imagery from sources such as Landsat Enhanced Thematic Mapper (ETM+) is typically classified through unsupervised and supervised classification techniques (Lillesand and Kiefer, 2000). While these techniques can be highly accurate, they are also time consuming and expensive for large areas. Other classification techniques such as regression-tree analysis have been proven effective for large areas (Lawrence and Wright, 2001). Using regression-tree analysis with a variety of imagery to classify large areas produces highly accurate results in a relatively short time and is inexpensive (Lawrence and Wright, 2001). For this project, using regression-tree analysis to classify images involved the following four steps: 1) classifying high-resolution imagery through standard techniques (i.e., unsupervised/supervised classifications), 2) combining the classified high-resolution imagery with lower resolution imagery and ancillary data,

3) using regression-tree software to build prediction models, and 4) using the prediction models to produce a classified layer. Using these four steps, percent canopy-cover and impervious-surface-cover layers were derived for Zone 41 located in Minnesota and Wisconsin (figure 1).

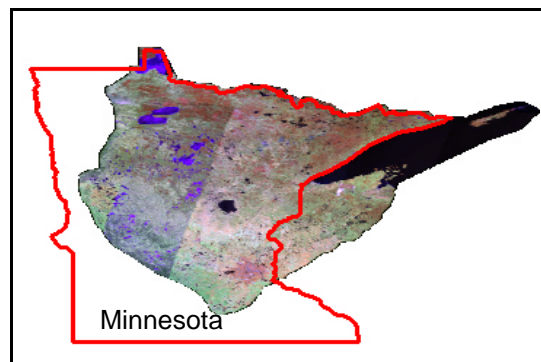


Figure 1—Zone 41 and Minnesota.

**For additional information**, contact Jule Caylor, Remote Sensing Applications Center, 2222 West 2300 South, Salt Lake City, UT 84119; phone: 801-975-3750; e-mail: [jcaylor@fs.fed.us](mailto:jcaylor@fs.fed.us). This publication can be downloaded from our Web site: <http://fsweb.rsac.fs.fed.us>. **For information about the San Dimas Technology & Development Center Program**, visit the SDTDC Web site: <http://fsweb.sdtc.wo.fs.fed.us>

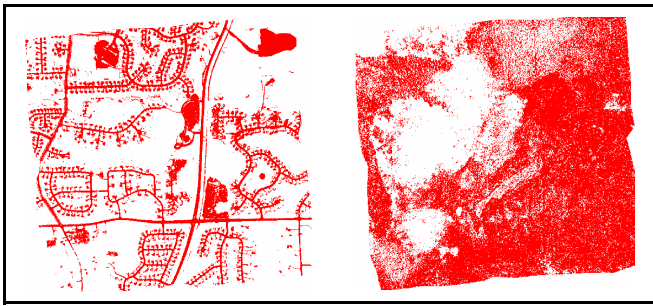


Figure 2—Impervious-surface-cover and canopy-cover classifications used to analyze Zone 41.

### Classifying High-Resolution Imagery

The Remote Sensing Applications Center (RSAC) obtained 28 scanned, natural-color, aerial photographs from the Natural Resources Conservation Service (NRCS). Each image covered 289 hectares and had a scanned spatial resolution of 0.3 meters. Only the effective area of each photograph was used in this analysis due to geometric distortions at the edges. Occasionally, other areas were excluded due to poor scan quality. The 28 digital aerial photographs were classified using unsupervised classification techniques into three categories: impervious surface cover, canopy cover, and background (figure 2).

### Combining Classifications with Lower Resolution Imagery and Ancillary Data

After the digital aerial photographs were classified, they were converted to percent cover layers (figure 3) with a spatial resolution of 30 meters, which matched the spatial resolution of the Landsat ETM+ imagery used for the analysis. The percent cover layers were combined with Landsat ETM+ imagery, state soil geographic (STATSGO) data, Tasseled Cap transformations, elevation, slope, and aspect information. All of these data layers, which had a spatial resolution of 30 meters were obtained from the U.S. Geological Survey, EROS Data Center (EDC). The Landsat ETM+ data consisted of three composite images

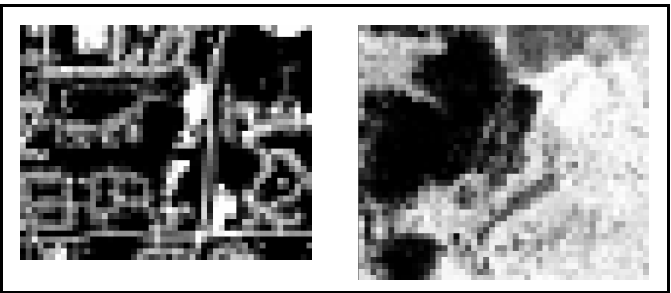


Figure 3—Percent impervious-surface-cover and canopy-cover layers at 30 meter spatial resolution.

from spring (March–May), leaf-on (July–August), and leaf-off (October–November) time periods (figure 4). The STATSGO imagery consisted of three layers: soil carbon, soil quality, and soil available-water capacity.

### Removing Clouds from the Landsat ETM+ Imagery

The Landsat ETM+ scenes selected for the composite images were the most cloud-free ones available. However, some clouds occurred in the images. RSAC has developed an automated process that eliminates clouds by predicting what the digital-number (DN) values would be if there were no clouds (figure 5). This process successfully removed all the clouds from the Landsat ETM+ scenes. RSAC tested the accuracy of this method by predicting DN values for nonclouded imagery and comparing these values with the original ones. The results are shown in table 1. On average, the predicted DN values did not differ by more than one for all bands.

### Regression-Tree Classification

The aerial photography and Landsat ETM+ images were combined to build prediction models using Cubist ([www.rulequest.com](http://www.rulequest.com)). Cubist does not read remotely sensed imagery directly. Thus, the imagery must first be converted to comma-delimited, ASCII text.

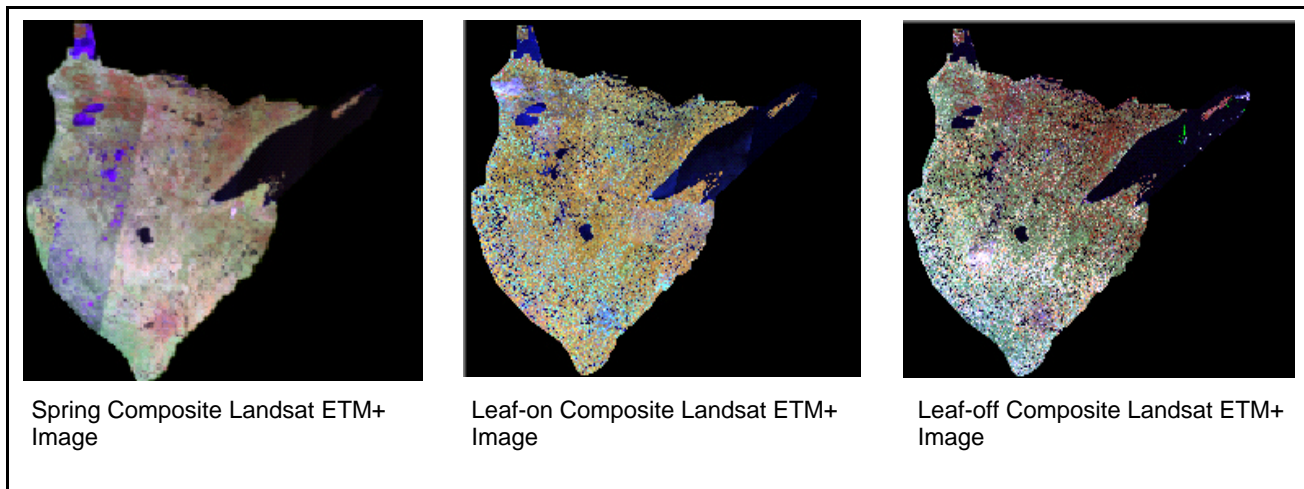


Figure 4—Landsat EMT+ composites used to derive percent impervious cover and canopy-cover layers for Zone 41.



Figure 5—Above left, the original Landsat EMT+ image with clouds; above right, the EMT+ image with the clouds removed by using the eliminate clouds procedure developed at RSAC.

Cubist can be run three different ways: “Rules alone,” “Instances and rules,” and “Let Cubist decide.” Models were created using all three of these options. The “Instances and rules” and “Let Cubist decide” options produced models with higher correlation coefficients and lower average errors than the “Rules alone” approach. However, due to the complexity of the models generated, it would have taken more than 250 days to apply these models to the entire data set. Therefore, models generated using the “Rules alone” option were applied to the entire data set, and only took 18 hours to implement.

Several Cubist trials using different image layers were run to derive the best models. For predicting impervious surface cover, the model that included bands 1–5 and 7 of the Landsat ETM+ composite images, elevation, slope, and aspect had the lowest amount of error and highest correlation. For predicting canopy cover, the model that included bands 1–5 and 7 of the Landsat ETM+ composite images, elevation, and slope had the lowest amount of error and highest correlation. However, when

the model was applied to Zone 41, noticeable seam lines appeared in the imagery. The lines occurred where the different dates of Landsat ETM+ imagery joined each other in the composites. To make the results more accurate, the final model chosen to depict canopy cover included bands 1–5 and 7 of the composite Landsat ETM+ imagery, elevation, and the thematic Landsat ETM+ composite date bands. The final predicted canopy-cover image appears in figure 6. The final predicted impervious-surface-cover image is shown in figure 7.

### Accuracy Assessment

To assess the accuracy of the models, RSAC did a 10-fold cross validation using Cubist. A 10-fold cross validation technique divides the data into 10 equal groups. With each iteration, nine groups are used to build models and the remaining group is used to test the models. At the end of each iteration, error rates and correlation coefficients are calculated. At the end of the 10-fold cross validation, the error rates and correlation

Table 1—Summary of the Differences Between Predicted and Original DN Values for Landsat ETM+ Images

	Band 1	Band 2	Band 3	Band 4	Band 5	Band 7
<b>Maximum Difference</b>	21	25	25	31	60	32
<b>Minimum Difference</b>	-14	-23	-23	-53	-77	-55
<b>Mode</b>	0	0	0	0	-1	0
<b>Median</b>	0	0	0	0	1	0
<b>Mean Difference</b>	0.176	0.038	0.037	-0.118	0.758	0.435
<b>Standard Deviation</b>	1.915	2.787	2.778	4.678	6.658	4.820

coefficients are averaged. These averages are nearly unbiased predictors of the true error rates and true correlation coefficients of the model built with all the data. The 10-fold cross validation for the percent canopy-cover model had an average error of 14.5, relative error of 0.55, and correlation coefficient of 0.77. The 10-fold cross validation for the percent impervious-surface-cover analysis had an average error of 4.6, relative error of 0.36, and correlation coefficient of 0.82.

## Conclusion

Compared to other classification techniques that could have been used to derive these percent cover layers, these accuracy levels are considered high. Additionally, these methods are much faster and less expensive than traditional approaches (Lawrence and Wright, 2001). These procedures provide geospatial data users an improved method to derive percent cover layers.

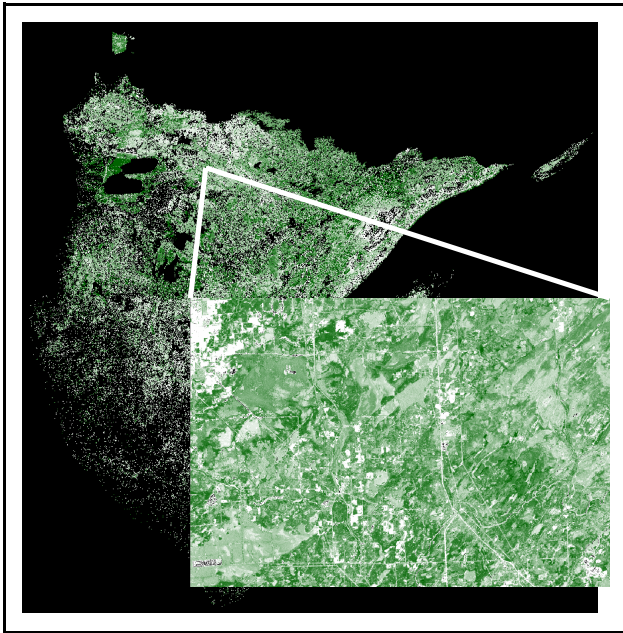


Figure 6—Predicted canopy cover for Zone 41.

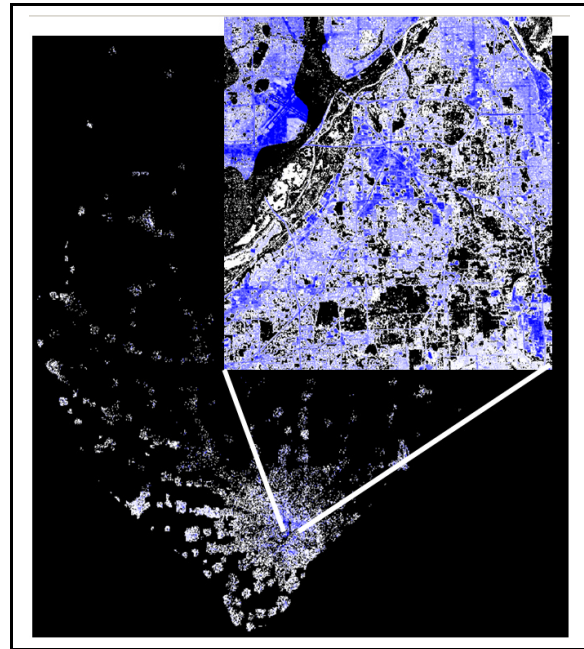


Figure 7—Predicted impervious surface cover for Zone 41.

## Suggested Reading

Lillesand, T.M., and R.W. Kiefer. 2000. "Remote Sensing and Image Interpretation," John Wiley & Sons, Inc., New York, NY.

Lawrence, R.L., and A. Wright. 2001. "Rule-Based Classification Systems Using Classification and Regression Tree (CART) Analysis," *Photogrammetric Engineering and Remote Sensing*, 67:1137–1142.

The Forest Service, United States Department of Agriculture (USDA), has developed this information for the guidance of its employees, its contractors, and its cooperating Federal and State agencies, and is not responsible for the interpretation or use of this information by anyone except its own employees. The use of trade, firm, or corporation names in this document is for the information and convenience of the reader. Such use does not constitute an official evaluation, conclusion, recommendation, endorsement, or approval by the department of any product or service to the exclusion of others that may be suitable.

The U.S. Department of Agriculture (USDA) prohibits discrimination in all its programs and activities on the basis of race, color, national origin, sex, religion, age, disability, political beliefs, sexual orientation, or marital or family status. (Not all prohibited bases apply to all programs). Persons with disabilities who require alternative means for communication of program information (Braille, large print, audiotape, etc.) should contact USDA's TARGET Center at 202-720-2600 (voice and TDD).

To file a complaint of discrimination, write USDA, Director, Office of Civil Rights, Room 326-W, Whitten Building, 1400 Independence Avenue, SW, Washington, DC 20250-9410 or call 202-720-5964 (Voice and TDD). USDA is an equal opportunity provider and employer.