

ANALYSIS OF CANOPY COVER AND IMPERVIOUS SURFACE COVER FOR USGS ZONE 41 USING REGRESSION-TREES AND ERDAS IMAGINE

Classification of satellite imagery, such as Landsat Enhanced Thematic Mapper (ETM+), is typically achieved through unsupervised and supervised classification techniques. While these techniques can be highly accurate, they can be time consuming and expensive to perform over large areas. Classifying large areas sion tree analysis with a variety of imagery yields highly accurate results in a short time and is inexpensive to perform. The Remote Sensing Applications Center (RSAC) used Cubist (www.rulequest.com) and a suite of Leica-Geosystems ERDAS Imagine programs developed at RSAC to classify the US Geological Survey (USGS) Zone 41 into percent canopy-cover and percent impervious-surface-cover.



Step 1: Eliminate Clouds



RSAC predicted the digital number values of a cloud-free image and compared the results with the original digital number values.





	Band 1	Band 2	Band 3	Band 4	Band 5	Band /
Max Difference	21	25	25	31	60	32
Min Difference	-14	-23	-23	-53	-77	-55
Mode	0	0	0	0	-1	0
Median	0	0	0	0	1	Ö
Mean Difference	0.176	0.038	0.037	-0.118	0.758	0.435
St. Dev.	1.915	2.787	2.778	4.678	6.658	4.820

Step 2: Select High-resolution imagery



This suite of programs can aide in the selection of the high-resolution imagery to use for the training data in the regression-tree analyses. For the analysis of Zone 41, the USGS chose 28 digital aerial photographs to use for the training data. These programs were not used for the selection of these images.



Create Water Mask



Select Block



Convert Image to Thematic



Subset High-resolution Images









The objective of this suite of programs is to convert the high-resolution classified imagery into percent cover layers and combine these layers with the subsetted low-resolution imagery. The "Recode Classified Image to Binary" model recodes the high-resolution images to images containing ones and zeros; the ones represent the class of interest. The "Percent Cover" program calculates how many pixels with a value of one lie within a pixel of the lowresolution imagery and converts this to a percentage. The "Percent Cover" output images have the same spatial resolution of the high-resolution imagery. The spatial resolution must equal the spatial resolution of the low-resolution imagery. The "Change Spatial Resolution" program changes the spatial resolution of the percent cover images. The "Layer Stack" program stacks the percent cover imagery with the subsetted low-resolution imagery.

Step 6: Select Training/Testing Blocks



This program randomly selects blocks of imagery to use as training and testing in the regression-tree analyses. RSAC did not use this program for the analysis of Zone 41 because of the limited training data that were available.



Training Data Blocks

RSAC classified 28 digital aerial photographs using ISODATA in ERDAS Imagine. The images were classified into three categories: canopy-cover, impervious-surface-





cover classification

Canopy-cover classification



Select the training image (*.img)

Step 7: Prepare Data For Cubist

Step 8: Run Cubist (<u>www.rulequest.com</u>)

Output Cubist Data File (*.data)

Cubist, the regression-tree software used for the analysis of Zone 41, cannot read remote sensing data directly. This program converts imagery into commadelimited ASCII formats.

	The rout theit
used if the high-	
elected using the "Select" suite of programs. If	
s used to select imagery,	
ution and high-	
t their boundaries line	
7 (1	

Step 5: Calculate Percent Cover





This program uses the CubistSam code available at <u>www.rulequest.com</u> to apply the Cubist model to an image. It is strongly recommended that Cubist is run using the "Rules alone" option. The other two options can take a long time to implement. This program was used to generate the output shown on the right.

Accuracy Assessment of the

Percent Canopy-Cover and Percent Impervious-Surface-Cover Analysis of Zone 41

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 coefficients are averaged. These averages are nearly unbiased predictors of the ture 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.







Percent Canopy-Cover For Zone 41



Percent Impervious-Surface-Cover For Zone 41











Percent Canopy-Cover For Zone



Percent Impervious-Surface-Cover For Zone 41



Percent Impervious-Surface-Cover For Zone 41

Acknowledgements

This report was initiated and funded by the Inventory and Monitoring (I & M) Steering Committee of the USDA Forest Service. Oversight for the project was provided by the San Dimas Technology and Development Center. Additional support was provided by the Remote Sensing Applications Center. Inventory Analysis and Accuracy Assessment Program. We would also like to thank the USDA Forest Service Northeastern Research Station, the US Geological survey, and the Earth Satellite Corporation for their support and assistance.

Prepared by: Bonnie Ruefenacht February 2003 (bruefenacht@fs.fed.us)