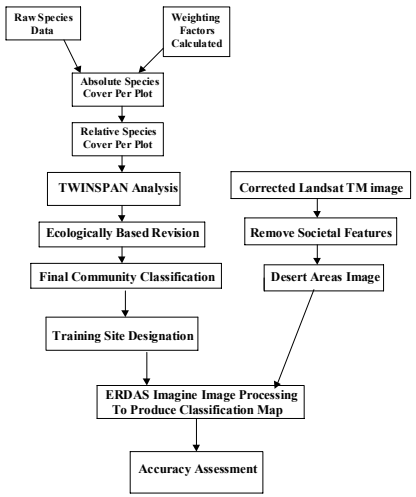
Preliminary classification of desert plant communities using Landsat Thematic Mapper (TM) data

Arthur Stiles[1] and Samuel Scheiner[2]

1. Department of Plant Biology, Arizona State University
2. Division of Environmental Biology, National Science Foundation

Abstract

Remote sensing is a valuable means by which to collect and subsequently generate spatially explicit information. Landsat Thematic Mapper (TM) is a versatile and well-utilized platform from which to gather spatial data. In concert with image processing software, the user is able to analyze larger expanses than would be possible with ground sampling alone. This project is concerned with the creation of a map depicting Sonoran desert community distribution, based on woody species (shrubs, trees, and cacti), across the CAP-LTER study area centered on the Phoenix metropolitan area, Arizona. Using species data from desert areas collected in the 200-sites survey of Spring 2000 and Stiles' data in remnant patches, a community classification was created using TWINSPAN. A supervised classification of TM imagery was performed using the community classification as a guide for the delineation of training groups. It was found that conventional image classification techniques were inadequate for the creation of a sufficiently accurate community distribution. Some gross-scale patterns were faithful to known distributions on the ground, but this was not true for all large-scale patterns and per-pixel accuracy fell below acceptable levels. Better results are possible using some modifications to the strategy. Proposed solutions are discussed.



Community Classification

I. Larrea-Dominated Lowland (42 sites)

II. Transitional
- A. Upper bajada (51 sites)
- B. Larrea / Ambrosia codominant (20 sites)
- C. Wash dominated (18 sites)
- D. Larrea / Encelia codominant (10 sites)
- E. Ambrosia dominated (9 sites)

III. Arizona Upland (20 sites)

IV. Alternatives
- A. Atriplex dominated (3 sites)
- B. Riparian area (2 sites)
- C. Encelia dominated (2 sites)

Figure 1: This community classification was generated using 177 ground-based samples. Here, the organization is centered around community structuring based on water limitation. Larrea tridentata-dominated stands are lower elevational locations receiving the least precipitation; Uplands are at the highest elevation and receive the most precipitation. "Transitional" types are generally intermediate with regard to water availability and elevation. "Alternative" types deviate from this continuum by the intercession of one or more environmental factors (A: saline soils, B: perennial water source, C: fire recovery & possibly soil factors).

Methodology

I. The two data sets used, 200-sites and Stiles-sites, were sampled using different designs. 200-sites data were collected within a single quadrat, 30 x 30 m^2; Stiles-sites consisted of five circular 100 m^2 quadrats arrayed along a transect. Since the sample grains are not equivalent, the data were transformed to allow comparison (described below).

II. For the 200-sites survey, measurements were taken on selected individuals of each species encountered in each plot. Length and width values were averaged, then squared and multiplied by pi (3.14156); all individuals were averaged together to yield a mean coverage value per species. This value was used as a weighting factor to compensate for differences in plant size.

III. Plant counts per sample were multiplied by the weighting factors to yield absolute coverage per species. Each species value was divided by the total sample cover to yield relative coverages.

IV. TWINSPAN from PCOrd was used to classify the samples into groupings. This function ordinates the samples based on Reciprocal Averaging and then splits the grouping into two parts depending on the presence or absence of one or more indicator species. This operation continues until all groups contain less than 5 samples, regardless of the ecological realism. This output was revised based on site and ecological information to produce the final community classification.

V. An atmospherically corrected Landsat TM image from August 1999 was obtained. This date was chosen for two reasons: 1. Appreciable precipitation amounts (promotes active metabolism in woody species, expected to increase spectral reflectance), and 2. Summer herbaceous coverage is generally lower than spring cover (decreases herbaceous impact on woody reflectance). All societal features (e.g. roads, buildings) were extracted from the image to yield a view of solely undeveloped desert areas.

VI. Woody plant community designations were used to determine training regions, which define the reference spectra for each community type. Image pixels are compared to the reference spectra and assigned a community category based on signal proximity using ERDAS Imagine's Image Classifier. This generates a map of woody plant community distribution.

VII. For more common community types, accuracy assessment was performed. Prior to the training process (VI.), each class group was split so that at least 10 samples were retained for reference signature generation. The remainder was used after the image processing steps to assess the accuracy of each category.

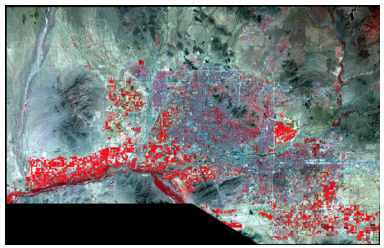


Figure 2: This is the Landsat TM image used in this study. Bands 4,3, and 2 are projected through the Red, Green, and Blue color guns, respectively. City grids (geometrically patterned around center), major roads, and agricultural fields (bright red polygons mostly in the lower half of image) were extracted. White Tank mountains are due west of Phoenix, South Mountain is due south of Phoenix, and the Fountain Hills / Verde River region is in the northeast corner.

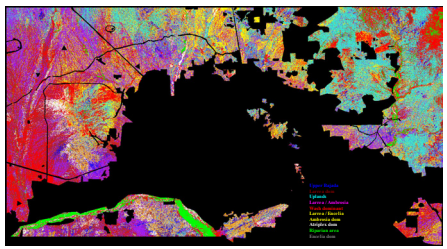


Figure 3: Trial One: This classification was created with all community designations listed in figure 1. Examining gross-scale patterns, riparian areas and the Arizona Uplands were positioned accurately for the most part. Atriplex-dominated stands were overemphasized and placed wrongly in the White Tank and South Mountains. Wash dominated stands were overemphasized in the flats west and southwest of the White Tanks. Per-pixel accuracies are shown below.

Classified Data

Ground Based Data (Reference)	Wash Inf	Larrea Dom	Uplands	Lar/Amb Cod	Wash Dom	Total
Upper Bajada	29	9	0	8	0	46
Larrea Dom	3	39	2	5	1	42
Uplands	22	2	26	11	8	69
Lar/Amb Cod	12	15	10	12	0	49
Wash Dom	8	10	3	3	11	35
Lar/Enc Cod	45	23	0	13	1	83
Ambrosia Dom	13	28	10	6	2	51
Atriplex Dom	14	3	3	4	2	26
Encelia Dom	36	5	0	4	0	45
Riparian	1	2	0	0	9	12
Total	187	116	54	66	35	458

Producers Accuracy: totals from Reference data

	Correct	Total	Accuracy
Upper Bajada	29	46	63.0%
Larrea Dom	39	42	71.4%
Uplands	26	69	36.2%
Lar/Amb Cod	12	49	24.5%
Wash Dom	11	35	31.4%

User's Accuracy: totals from Classified Data

	Correct	Total	Accuracy
Upper Bajada	29	187	15.5%
Larrea Dom	39	116	33.6%
Uplands	26	54	48.1%
Lar/Amb Cod	12	66	18.2%
Wash Dom	11	35	31.4%

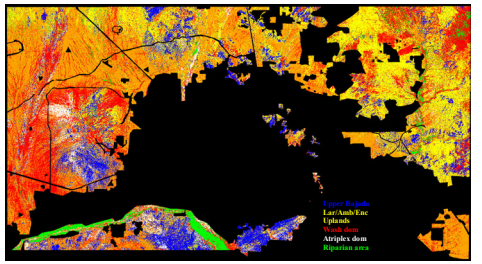


Figure 4: Trial Two: Given the unsatisfactory performance of the previous classification, the analysis was repeated after all types dominated by Larrea, Ambrosia, and Encelia were combined into a single group. It was thought that of the communities, these differ from each other less than other groupings. Using fewer divisions would decrease the "reference load" and may reduce confusion between classes. Gross-scale discrimination of flatlands above the White Tanks improved dramatically as wash courses are now perceptible from the "LAE" area. Other patterns are roughly similar, except that the Upper Bajada now dominates South Mountain. Per-pixel accuracies were improved over Trial One.

Classified Data

Ground Based Data (Reference)	LAE combo	Wash Inf	Uplands	Wash Dom	Total
LAE combo	68	11	15	1	95
Upper Baj	114	107	1	0	222
Uplands	36	39	29	9	115
Wash Dom	23	12	5	12	52
Atriplex Dom	4	17	4	2	27
Riparian	3	1	0	9	13
Total	250	187	54	33	524

Producers Accuracy: totals from Reference data

	Correct	Total	Accuracy
LAE combo	68	95	71.6%
Upper Baj	107	222	48.2%
Uplands	29	115	25.2%
Wash Dom	12	52	23.1%

User's Accuracy: totals from Classified Data

	Correct	Total	Accuracy
LAE combo	68	250	27.2%
Upper Baj	107	187	57.2%
Uplands	29	54	53.7%
Wash Dom	12	33	36.4%

Future Directions

This work represents the preliminary effort in the mapping of community types within undeveloped desert locales based on woody plant species. Further work is needed before a sufficiently accurate distribution map can be generated. Landsat TM data includes the combined per-pixel signal of all cover types (plants, rocks, soil, asphalt, etc.); especially given the inherently sparse nature of desert vegetation, it is important to account for soil and geological features. This would likely involve a procedure by which plant samples are stratified by soil / geology and analyzed separately. Designation of training and accuracy assessment sites would be improved if larger, continuous ground samples are used. Unsampled proximal vegetation may have obscured the "purity" of training signatures. This will entail additional ground-sampling efforts in which large polygons (at least 4 pixels) are recorded around a single community type, devoid of others, using GPS. Additionally, the "expert system approach" will be used. This means that remote sensing analyses are carried out using other data sources in concert with the image data. GIS layers (e.g. Digital Elevation Models) may be used to constrain the decision space; in other words, assignment of types can be restricted to only those areas in which those communities may be found. For example, Atriplex-dominated communities could only be located on flat ground at lower elevations.

ASU Arizona State University

CAP LTER
Central Arizona-Phoenix
Long-Term Ecological Research