MAPPING ARID VEGETATION SPECIES DISTRIBUTIONS IN THE WHITE MOUNTAINS, EASTERN CALIFORNIA, USING AVIRIS, TOPOGRAPHY, AND GEOLOGY

C. Van de Ven* and S. B. Weiss**[1]

1. INTRODUCTION

1.1 Problem Statement

Our challenge is to model plant species distributions in complex montane environments using disparate sources of data, including topography, geology, and hyperspectral data. From an ecologist's point of view, species distributions are determined by local environment and disturbance history, while spectral data are "ancillary." However, a remote sensor's perspective says that spectral data provide picture of what vegetation is there, topographic and geologic data are "ancillary." In order to bridge the gap, all available data should be used to get the best possible prediction of species distributions using complex multivariate techniques implemented on a GIS.

Vegetation reflects local climatic and nutrient conditions, both of which can be modeled, allowing predictive mapping of vegetation distributions. Geologic substrate strongly affects chemical, thermal, and physical properties of soils, while climatic conditions are determined by local topography. As elevation increases, precipitation increases and temperature decreases. Aspect, slope, and surrounding topography determine potential insolation, so that south-facing slopes are warmer and north-facing slopes cooler at a given elevation. Topographic position (ridge, slope, canyon, or meadow) and slope angle affect sediment accumulation and soil depth. These factors combine as complex environmental gradients, and underlie many features of plant distributions. AVIRIS data, digital elevation models, digitized geologic maps, and 378 ground control points were used to predictively map species distributions in the central and southern White Mountains, along the western boundary of the Basin and Range province. Minimum Noise Fraction (MNF) bands were calculated from the visible and near-infrared AVIRIS bands, and combined with digitized geologic maps and topographic variables using Canonical Correspondence Analysis (CCA). CCA allows for modeling species "envelopes" in multidimensional environmental space, which can then be projected across entire landscapes.

1.2 Setting

The location for this study is the White Mountains, east of, and within the rain shadow of, the Sierra Nevada (see Figure 1). Elevation ranges from 1220 m in adjacent Owens Valley to 4330 m at White Mountain Peak. Geologic substrates include granitoids, metavolcanics, quartzites, argillites, phyllites, limestones, and dolomites. Vegetation includes desert shrubs, pinyon-juniper woodland, sagebrush meadows, subalpine bristlecone-limber pine forests, and alpine fell fields (Mooney 1973; Hall 1993).

---

[1] *Dept. of Geological and Environmental Sciences, Stanford University, Palo Alto, California; vandeven@pangea.stanford.edu
** Creekside Center for Earth Observations, Menlo Park, California 94025; StuBWeiss@netscape.net

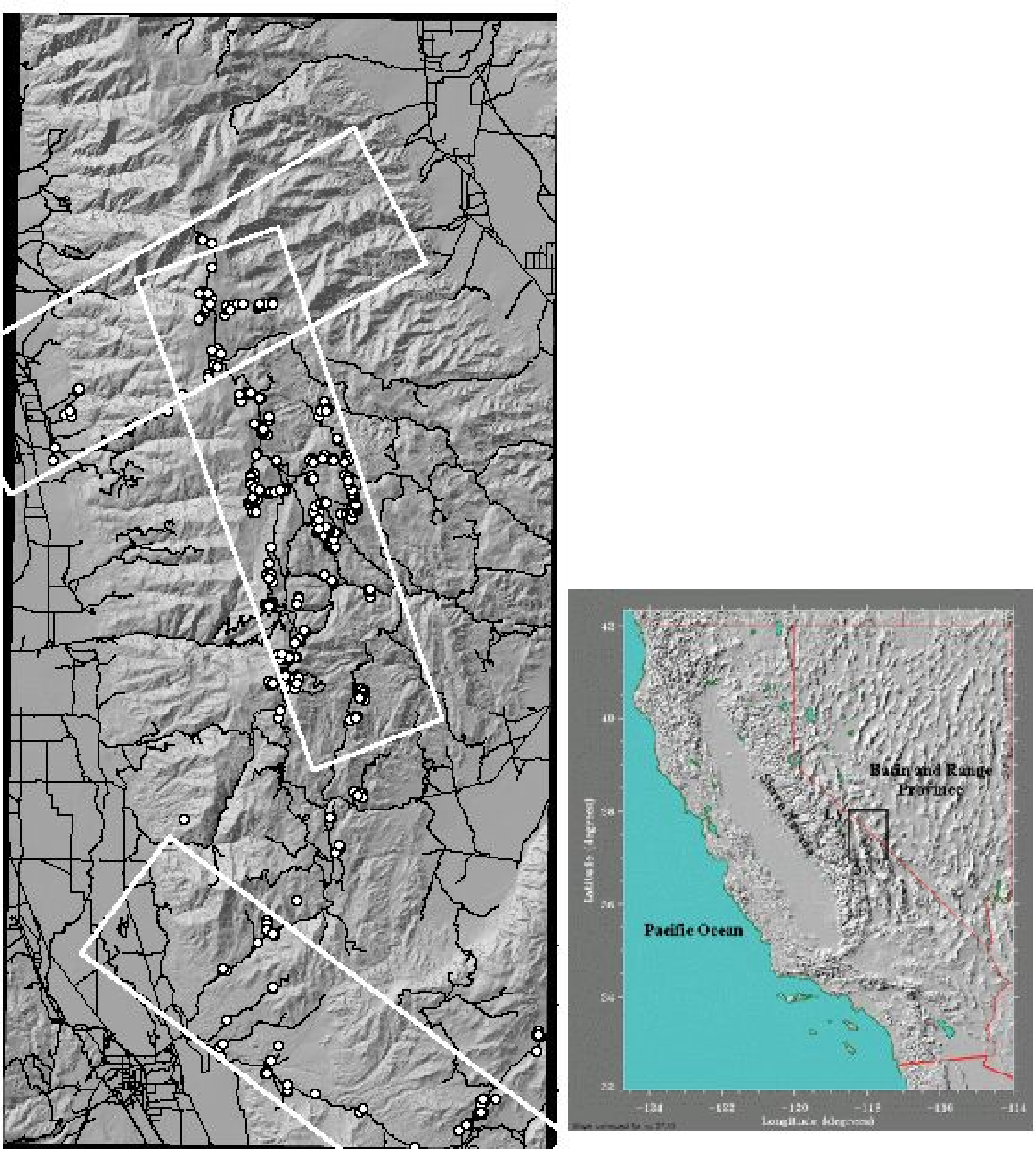

Figure 1. Shaded relief map of the White Mountains, with three October, 1996 AVIRIS flightlines (white boxes), roads and trails (black lines), and vegetation sites (white dots).

2. METHODS

2.1 Canonical Correspondence Analysis

Canonical Correspondence Analysis (CCA) is a direct gradient ordination that relates site/species and site/environment matrices (ter Braak 1998). It extracts synthetic axes that are linear combinations of environmental variables, indicated by bold names (topographic position, insolation, elevation, slope, and rock types – granite, quartzite, limestone, and dolomite) and numbers for MNF bands in Figure 1, that maximize separation of species and sites. Environmental variables were chosen by forward selection with Monte-Carlo significance tests. Six axes were extracted from 21 input variables. Each species has a centroid and tolerance (standard deviation) on each axis that defines its distribution envelope in canonical space as a multi-dimensional ellipsoid known as a species envelope whose centroids are indicated by abbreviated species' names in Figure 2. Maps of all six CCA axes were produced (four are shown in Figure 3). Statistical distances of each pixel to species centroid in six-dimensional canonical space were calculated. Results were expressed in standard deviation units - closer to the centroid means higher probability of occurrence, and beyond three standard deviations was taken as absence. Axes maps show the spatial realization of the linear combinations of environmental variables. Each cell has six axis values that are then used to determine whether that cell falls into environmental envelopes of selected species based on the species' centroids and tolerances in six-dimensional canonical space.

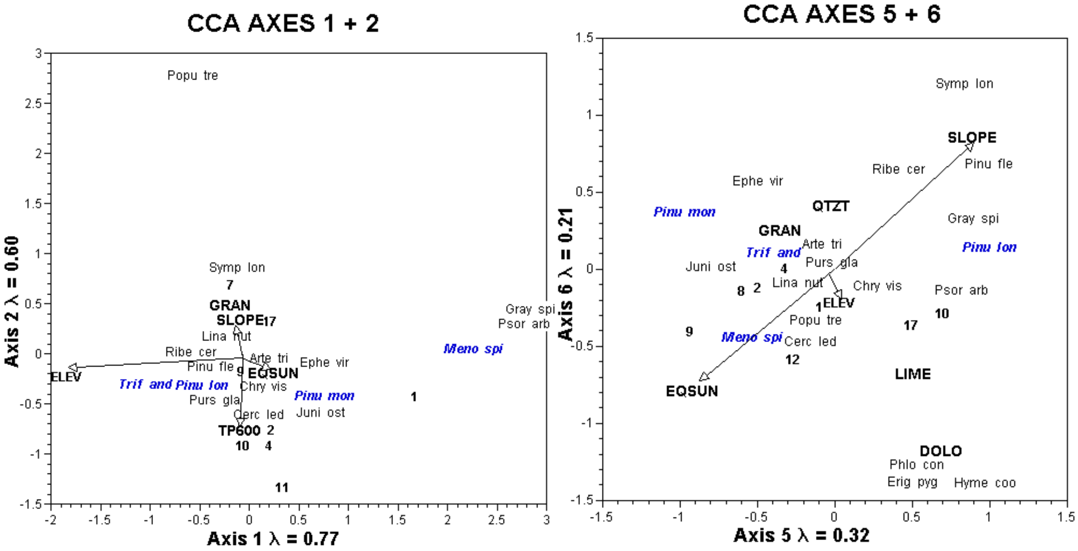

Figure 2 Left: Axis 1 is the long elevational gradient, running from desert shrubs on the right, to pinyon-juniper woodlands near the center, to subalpine bristlecone-limber pine forests, and alpine fell-fields to the left. Elevation is modified by EQsun; higher insolation lowers the effective elevation. Axis 2 is dominated by MNFs 11, 10, 4, 2, 7, and TP600 (topographic position) and strongly separates aspen (Popu tre) at the top of the diagram. Italicized, bold species are mapped in Figure 2. The eigenvalue (λ) for each axis is given next to the axis title.
Right: Axes 5 and 6 illustrate strong topographic effects (slope and EQsun) and geologic factors (the gradient from dolomite, limestone, through granitoid/quartzite). Lower eigenvalues for these axes indicate that fewer species (toward the edges of the diagram) are effectively separated, and gradients are relatively short. Three dolomite restricted herb species are shown lower right. MNF scores are more difficult to interpret, but their spread along these axes show that they add significant information to the axes.

2.2 CCA Methods

Axes and species parameters were derived from 202 calibration plots. 146 plots were used to evaluate presence/absence of 17 species. Projections were smoothed by 3x3 cell moving average (to account for registration errors between the different environmental and spectral grids). Kappa statistics were calculated for each species. Kappa is the relative increase in correct assignments over that expected by chance. A threshold standard deviation was chosen to maximize kappa.

Figure 4 shows the predicted distributions of four selected species overlain on the hillshaded DEM Because MNF bands derived from the AVIRIS data were included as variables, species predictions could only be made within the three flightlines. Darker red indicates a higher probability of occurrence (closer to the centroid of the species envelope). The predicted species distributions includes details beyond simple elevation limits or geologic affinities, but incorporates environmental, geologic, and spectral constraints derived from all the input variables.

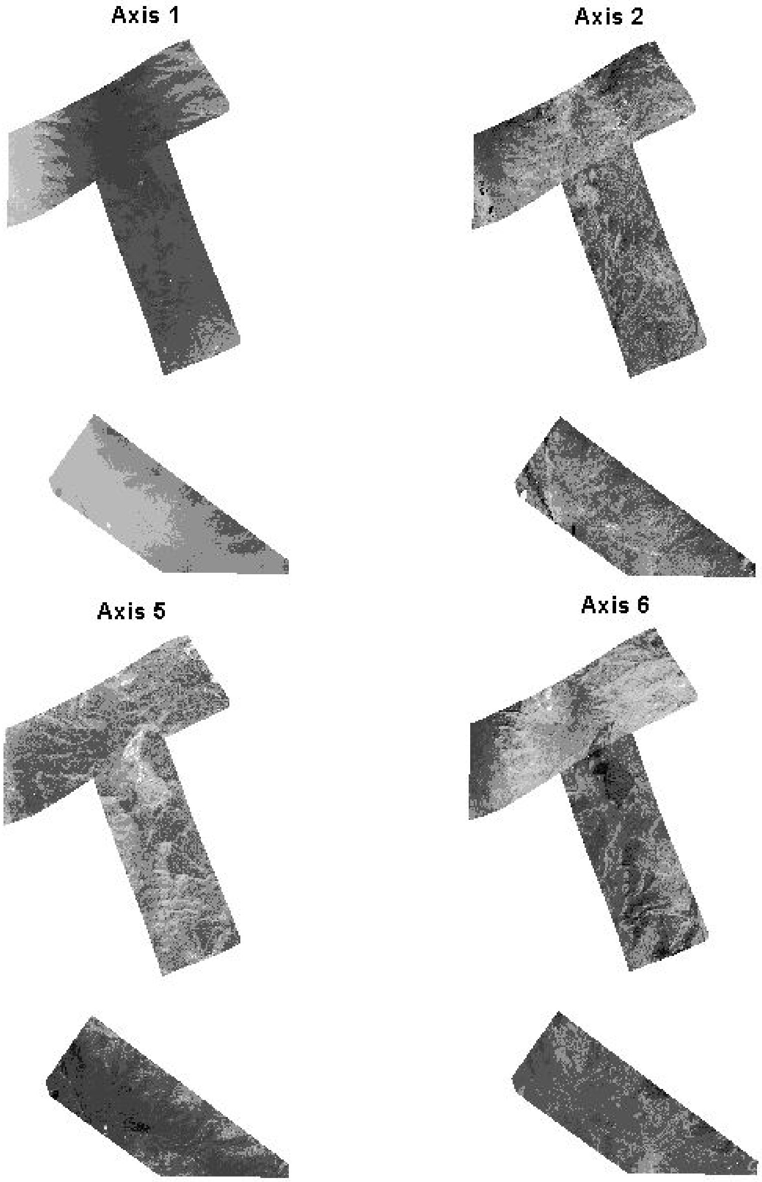

Figure 3: Axis 1 is the primary elevational gradient (dark gray is higher elevation), but the finer-scale variations include differences between the north and south sides of canyons. Axis 2 is more difficult to interpret, but is effective in separating out aspens. Axes 5 and 6 have large geologic components -- dolomite shows up particularly strong on Axis 6 as darker lighter gray at mid-elevations. Each cell has 6 axis values that are then used to determine whether that cell falls into environmental envelopes of selected species based on the species' centroids and tolerances in six-dimensional canonical space.

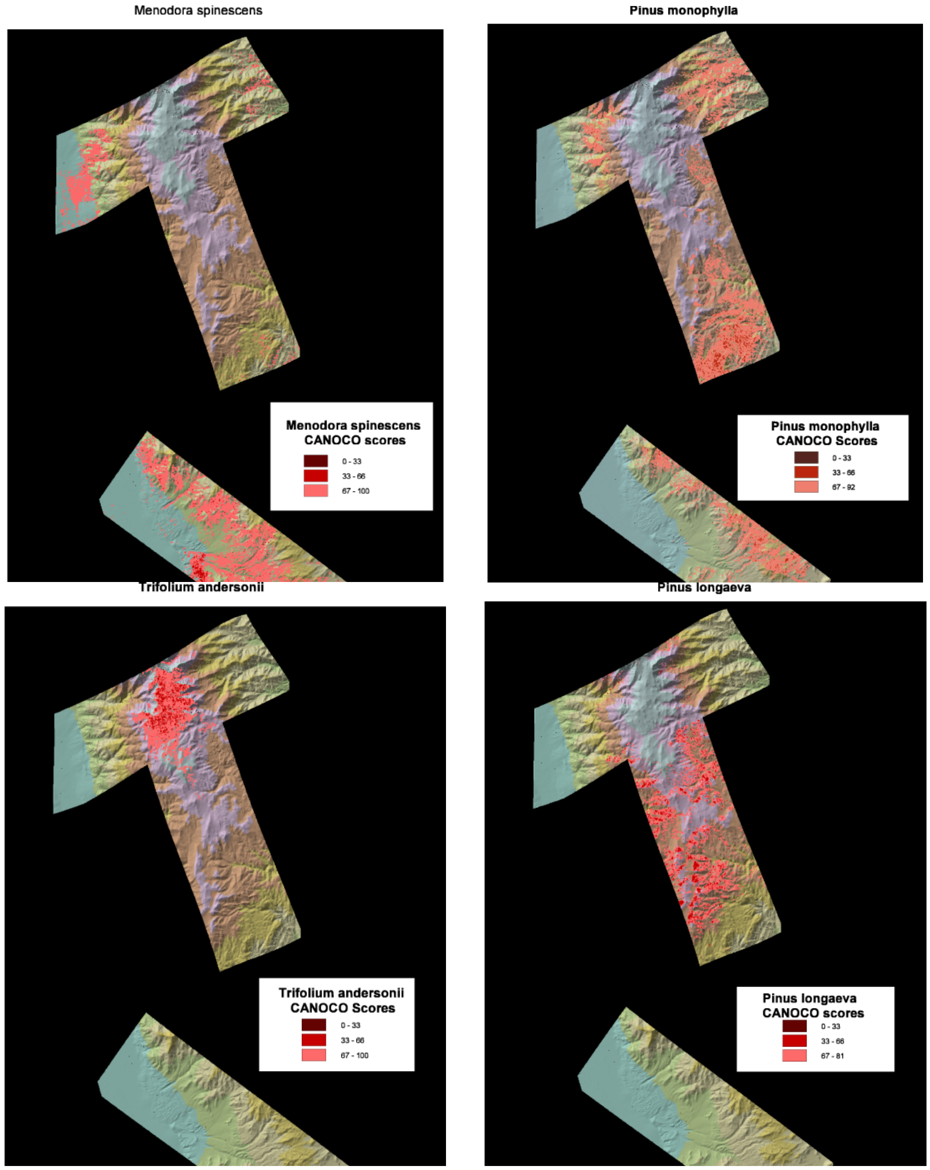

Figure 4 Top left: Predicted distribution of *Menodora spinescens* (Spiny Menodora), a low elevation desert shrub. *Menodora spinescens* occupies south-facing slopes at its extreme upper range limit (2300 m), all aspects at mid-elevations (1600-2000 m), and north-facing slopes at lower range limits (1300 m). The species has no obvious association with geology,

Top right: Predicted distribution of *Pinus monophylla* (Pinyon Pine) the lower timberline tree in the White Mountains. *Pinus monophylla* occupies south-facing slopes at its extreme upper range limit (3100 m), all aspects at mid-elevations (2300-2600 m), and north-facing slopes at lower range limits (1800 m). It also avoids dolomite, even in the middle of its elevation range.

Bottom left: Predicted distribution of *Trifolium andersonii* (Mono Clover), a common herb in the alpine-fell fields above timberline. *Trifolium andersonii* occupies south-facing slopes at its extreme upper range limit (4000 m), all aspects at mid-elevations (3700-3900 m), and north-facing slopes at lower range limits (3600 m). It has a preference for flat areas, and strongly avoids dolomite.

Bottom right: Predicted distribution of *Pinus longaeva* (Bristlecone Pine), the upper timberline tree and the most famous species from the mountain range. *Pinus longaeva* occupies south-facing slopes at its extreme upper range limit (3600 m), all aspects at mid-elevations (3100-3400 m), and north-facing slopes at lower range limits (2900 m). Bristlecone pine is found on all geologic substrates, contrary to the common perception of a tight association with dolomite.

3. EVALUATION AND RESULTS

Each species map can be considered a testable hypothesis, and the 146 evaluation plots provide an initial test. Our main evaluation criterion is the kappa statistic. As shown in Table 1, evaluation kappa ranged from 0.81 (a very good fit) for *Pinus monophylla* to 0.09 for *Purshia glandulosa*. Five of seventeen species models produced kappa > 0.60, a very satisfying value (Guisan et al. 1999). Five more were between 0.4 and 0.6, a moderate improvement over random. Seven species models produced poor fits (<0.40).

Table 1

Species	MaxKappa	Threshold	Absence	Presence	ProdAcc	UserAcc	OverallAcc
Pinus monophylla	0.81	87	136	10	0.7	1	0.98
Juniperus osteosperma	0.75	100	135	11	0.73	0.73	0.96
Trifolium andersonii	0.74	100	116	30	0.67	0.95	0.92
Pinus longaeva	0.72	81	123	23	0.78	0.75	0.92
Psorothamnus arborescens	0.67	100	126	20	0.6	0.86	0.93
Cercocarpus ledifolius	0.42	82	130	16	0.5	0.47	0.88
Pinus flexilis	0.41	93	127	19	0.79	0.38	0.81
Artemesia tridentate	0.41	100	37	109	0.72	0.91	0.73
Grayia spinosa	0.40	99	128	18	0.28	1	0.91
Linanthus nuttallii	0.40	69	108	38	0.5	0.59	0.78
Menodora spinescens	0.36	96	126	20	0.3	0.67	0.88
Ephedra viridis	0.35	79	127	19	0.47	0.41	0.84
Populus tremuloides	0.31	86	142	4	0.5	0.25	0.95
Symphoricarpos longiflorus	0.25	70	139	7	0.43	0.23	0.9
Chrythamnus viscidflorus	0.25	93	60	86	0.73	0.68	0.64
Ribes cereum	0.20	89	119	27	0.78	0.32	0.65
Purshia glandulosa	0.09	100	135	11	0.64	0.12	0.63

The species with good fits include 3 trees, 1 desert shrub, and 1 alpine herb. Species with well-defined environmental envelopes, and/or a well-defined spectral signature such as trees are naturally easier to model with these methods. Many of the poorer fits include shrub species that are disturbance-dependent (i.e. *Chrysothamnus viscidiflorus*), wide environmental tolerances (i.e. *Artemesia tridentata*), or have few occurrences for calibration and evaluation (i.e. *Populus tremuloides*). The quality of the models and evaluation data are strongly dependent on having many presences among the sample points.

Evaluation data provided an independent measure of fit (kappa). High evaluation kappas (>0.60) were obtained for 5/17 sample species, moderate kappas (0.4-0.6) were obtained for another 5 species, and the remaining 7 species with poor fits tended to be disturbance dependent or relatively rare. Combining hyperspectral imagery, topography, geology, and field data provide a powerful tool for understanding the major factors determining the distributions of specific plants in the arid White Mountains, but finding the correct method for effectively combining the data remains a research challenge.

The producer, user, and overall accuracies are strongly weighted by absence of species, and are misleadingly high. Kappa weights both omission and commission errors, and is a better-integrated measure of model success (or failure), and therefore is our primary evaluation criterion.

More than 60 species were abundant enough to model (> 2 sites present), but many had very few occurrences for both calibration and evaluation. For these species, any projected distribution is, at best, a broad hypothesis. Full evaluation of all species where it is appropriate is presently underway, and collection of more evaluation (and calibration) data is a priority for the upcoming field season.

4. DISCUSSION

The inclusion of topographic factors was effective in determining the fundamental temperature/moisture gradient. All species occupy proper elevational ranges, and major features of their microdistributions are captured. CCA Axis 1(the strongest axis) is primarily elevation but is influenced by insolation. Each additional MJ/m^2 of potential March 21 (equinox) insolation added approximately 30 meters to the effective elevation - the full range of insolation values from a steep north-facing slope to a steep south-facing slope (15 MJ/m^2) produces about a 500 m equivalent shift in elevation. Other topographic factors create more subtle patterns, but several species demonstrate distinct preferences for flats, steep slopes, canyons or ridgetops. However, many species depend on topographic features well below the resolution of the DEM, such as small depressions or rocky outcrops.

Geologic factors were important for many species. Some site-specific errors could be attributed to geology not digitized at a fine enough scale. Many of the formations and members described as quartzite for the purposes of this study actually contain distinct members, layers, or lenses of carbonates. These carbonate lenses frequently support very different vegetation than the surrounding clastic rocks, but are generally too small to appear on geologic maps, or, if mapped, were too numerous and discontinuous to be digitized separately. For example, the Wyman Formation contains significant carbonate layers, but was grouped with the quartzite units due to the greater proportion of that rock type. Many *Cercocarpus ledifolius* (Mountain Mahogany) sites in the Wyman Formation were actually on limestone lenses, surrounded by sagebrush-covered quartzites.

AVIRIS data (MNF bands) were highly significant in the CCA ordination for most of the tree species, and apparently contributed to species models, but were difficult to interpret except by visual inspection. Because the AVIRIS data was obtained in October, most of the vegetation had senesced. However, the tree species were still green, helping distinguish forested from unforested areas, although not perfectly due to topographic shading and spectral similiarity to green riparian grasses. One of the challenges in mapping mature bristlecone pine groves was the wide spacing between trees with minimal foliage on individual trees, making spectral identification of areas such as the Patriarch Grove (3500 m) difficult. While MNF data with environmental variables improved prediction of tree species, it did not improve the prediction of other species, and, in a few cases of low elevation shrubs (*Ephedra viridis* and *Menodora spinescens*) seemed to hinder accurate predictions. Low elevation shrubs grow sparsely on a variety of rock types, and, at the time of the AVIRIS imagery, was very dry, resulting in little vegetative cover to be detected within a 20 m pixel. Conversely, trees are considerably larger, frequently dense, and had a healthy vegetative spectral signature to be detected within the imagery.

The MNF data improved the predictions for four of six tree species (*Cercocarpus ledifolius, Juniperus osteosperma, Pinus longaeva,* and *Pinus flexilis*), and did slightly worse for *Populus tremuloides* and *Pinus monophylla*. The addition of the AVIRIS spectral (MNF) data did not significantly improve the prediction of shrub or herb species, and seemed to do particularly poorly for low elevation species. It is likely that the tree species were large enough, dense enough, and green enough to comprise spectrally distinct pixels from the smaller, drier shrub and herb dominated communities. *Populus tremuloides*, whose prediction was slightly worse with the addition of MNF data, was found in few sites, and may have been in different stages of senescing at different elevations. The low elevation shrubs did particularly poorly with the MNF data likely due to the fact that by October, when the AVIRIS data was obtained, the low elevation shrubs were completely dried and are sparsely distributed. Hence, the AVIRIS data primarily sees the geology at low elevations, rather than the vegetation. Inclusion of June 2000 AVIRIS flightlines may provide a better test of hyperspectral data inputs.

Using all three data sources (DEM derived products, AVIRIS data products, and digitized geologic maps) proved challenging, as registration accuracy becomes important. Some degree of data-smoothing was necessary - our predictions are at a scale of 90 meters (3x3 pixel window). Input data, especially AVIRIS MNF bands, may need to be smoothed as well to account for registration errors. At a broader scale, the obvious boundary on the MNF8 map corresponds to the flightline boundary, so stitching together scenes (even those taken within minutes of each other) may introduce artifacts.

5. CONCLUSIONS

Combinations of topographic, geologic, and hyperspectral data can provide good models for some individual species. Other species are more difficult to model; they are disturbance dependent, influenced by fine-scale or other unmapped ecological factors, or are not spectrally apparent (senescent in October or characteristically sparse cover).

These species-specific models go beyond coarse vegetation "communities" that can be determined from spectral data alone. Each type of data provides unique information. Spectrally similar species (i.e. pines) can be separated by elevation and geology; species at similar elevations may be separated by spectral data.

Predictions can be made for species with low cover. Spectral data alone can only map communities or large, "macroflora," while predictions can be made for nearly all species noted in field data based on non-spectral environmental variables. Each species projection is a testable hypothesis and may lead to insights into environmental controls on distribution, even in the absence of a high kappa values.

Topographic and geologic factors can be interpreted ecologically. Some MNFs can be interpreted, but many cannot due to incorporation of many visible (brightness or relief) and non-visible factors (water vapor or vegetation health).

6. NEXT STEPS

Mapping species distributions in complex mountain ranges is an inherently difficult problem. Many combinations of input data and analyses are possible. The MNF transformation may not be the ideal way to incorporate AVIRIS data. Spectral indices that relate directly to known features of vegetation (i.e. red-edge), or even standardized raw bands at key wavelengths may be more appropriate (and interpretable) spectral inputs. Because October is not an ideal month for vegetation (many species are senescent), the June 2000 AVIRIS data may provide better spectral information. AVIRIS flightlines are only 9 km wide but contain 224 spectral bands. Other imagery may also be used; LANDSAT images have coverage of entire range in different seasons, but only with seven spectral bands.

Further evaluation methods include black and white Digital Orthophoto Quads (DOQs) for broad-scale evaluation of tree cover. More field evaluation plots, especially where there are "interesting predictions," will allow for inclusion of more data into the calibration data set while still maintaining sufficient points for evaluation.

We have explored only one statistical method in this poster -- the use of CCA. There are many options within CCA itself. For example, running the ordinations with presence/absence data may provide a different picture (especially on higher order axes) because abundant species are downweighted. Species-specific General Linear Models (GLMs, Guisan et al. 1999), General Additive Models (GAMs), Regression Trees, Neural Networks, and other modeling techniques can be applied to this data set.

End member spectral classifications appear to work well for some species/communities (Van de Ven et al. 2000a and b), but inclusion of elevation/insolation (i.e. CCA Axis 1) as an input "band" may allow for a better representation of broad cover-types.

These sorts of analyses can be used to help understand implications for climate change on vegetation. Elevation and insolation, primary determinants of local temperature, are important variables in mapping present vegetation. For example, a 3 degree C rise in temperature is the equivalent of 500 meters elevation. Species may shift their elevation ranges, but the fine-scale patterns determined by insolation, slope, topographic position, and geology will still operate.

7. ACKNOWLEDGEMENTS

We gratefully acknowledge support from the White Mountain Research Station, Stanford University (McGee grant), NASA (grant NAG5-4888), Texaco (Exploration Technology and Production Dept.), the White Mountain Groupies, W. G. Ernst, R. J. P. Lyon, and the White Mountain Research Station chefs and staff.

8. LITERATURE CITED

Ernst, W. G., C. A. Nelson, and C. A. Hall, 1993, “Geology and metamorphic mineral assemblages of Precambrian and Cambrian rocks of the central White-Inyo Range, Eastern California,” Map Sheet 46, California Division of Mines and Geology, Sacramento, California.

Guisan, A., S. B. Weiss, A. D. Weiss.1999, “GLM versus CCA spatial modeling of plant species distribution,” Plant Ecology, vol. 143, pp. 107-122.

Hall, C. A. (ed.), 1991, Natural History of the White-Inyo range, Univ. of California Press, Berkeley, California.

Mooney, H. A. 1973, “Plant communities and vegetation,” in Lloyd, R.M., and R.S. Mitchell (eds.), Flora of the White Mountains, California and Nevada, Univ. of California Press, Berkeley, California, pp. 7-18.

ter Braak, C. J. F., 1996, CANOCO 4 Users Manual, Waginegen Press, Netherlands.

Van de Ven, C. M., W. G. Ernst, and R. J. P. Lyon, 2000a, “Relationships between the geologic substrate and vegetation in the White-Inyo Mountains, Eastern California as determined from AVIRIS and field data,” Proceedings of the Fourteenth International Conference Applied Geologic Remote Sensing, Veridian ERIM International, Ann Arbor, Michigan, pp. 36-43.

Van de Ven, C. M., W. G. Ernst, and R. J. P. Lyon, 2000b, “Analysis of the relationships between geologic substrate and vegetation in the White-Inyo Mountains, Eastern California,” Proceedings of the Ninth JPL Airborne Earth Science Workshop, JPL Publication 00-18, Jet Propulsion Laboratory, Pasadena, California. Also available at http://www.makalu.jpl.nasa.gov/docs/workshops/toc.htm Last accessed November 2001.