USING CART MODELING AND AVIRIS IMAGERY FOR ASSESSING RISK OF WEED INVASION AT VANDENBERG AIR FORCE BASE

Karen Olmstead,[1] Susan Ustin,[2] and Ted Ernst[3]

1. Introduction

Invasion by noxious weeds is a serious and rapidly accelerating ecological to the long-term environmental health of natural ecosystems. Current estimates indicate that in the U.S., weeds infest 70,000 hectares and cost more than $100 billion annually, causing severe economic losses and ecological degradation (Babbitt 1998; Pimentel et al. 2000).

One of the major problems faced by environmental managers is the difficulty in mapping the location of invasive weeds; traditional mapping methods using extensive ground surveys can be very costly and time-consuming. New remote sensing technologies, especially hyperspectral imagers, such as the Airborne Visible and Infrared Imaging Spectrometer (AVIRIS), may improve monitoring and mapping capability and improve early detection and management of invasive weeds. Our research addresses two broad objectives: 1) develop methods to inventory the spatial extent of established populations of noxious weeds; and 2) develop methods to estimate the likelihood of expansion in spatial distribution of invasive noxious weeds into additional habitat areas.

In a previous paper (Underwood et al. 2002), we presented our results for mapping iceplant and jubata grass at Vandenberg Air Force Base (VAFB) using classified AVIRIS images. In this paper we are presenting the preliminary results of our modeling efforts using the previously developed AVIRIS weed maps and classification and regression tree (CART) analyses to predict areas vulnerable to weed invasion.

2. Study Area

Vandenberg Air Force Base covers 98,400 acres (Fig. 1) in Santa Barbara County, and represents one of the last relatively undisturbed areas in coastal California (Keil and Holland 1998). The climate is

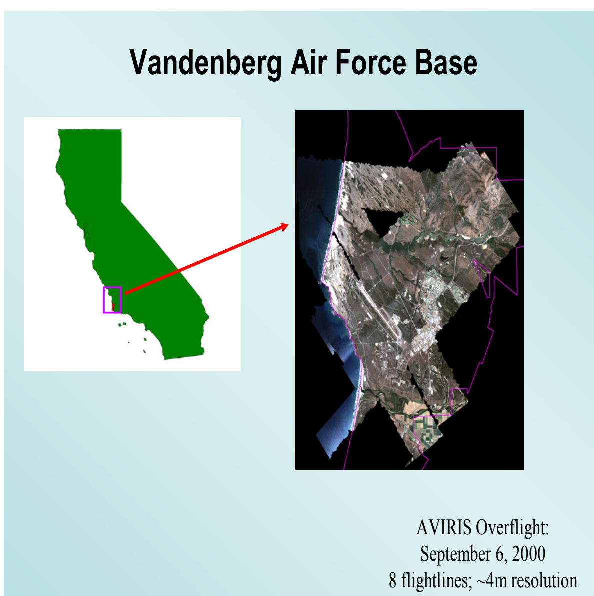

Figure 1. Map showing location of Vandenberg Air Force Base and area covered by AVIRIS overflight.

---

[1] Center for Spatial Technologies and Remote Sensing, University of California, Davis (ksolmstead@ucdavis.edu)
[2] Center for Spatial Technologies and Remote Sensing, University of California, Davis (slustin@ucdavis.edu)
[3] Dynamac Corp., Corvallis, Oregon (ernst@peak.org)

semiarid, but moderated by maritime influence and fog drip, and the landscape is typified by complex and varied topography, geology, and soils (The Nature Conservancy 1991).

The location of the base is in the ecological transition zone between northern and southern California, resulting in diverse environmental conditions and the presence of several unique and highly localized plant communities (Keil and Holland 1998) (Fig. 2). Since the military mission at VAFB requires extensive land space as a buffer for surrounding communities, much of the base remains relatively undisturbed (Coulombe and Cooper 1976). Despite this isolation, alien plant species have become established and invasive at VAFB, resulting in several communities dominated by introduced plants (Keil and Holland 1998).

Of particular concern, are *Carpobrotus edulis*, *C. chilensis*, and their hybrids (iceplant, also known as hottentot fig or sea fig, respectively) and *Cortaderia jubata* (jubatagrass, also known as purple pampas grass or Andean pampas grass) (Fig. 3). Both pose serious threats to two sensitive vegetation types, coastal dune scrub (hereafter, scrub) and maritime chaparral (hereafter, chaparral).

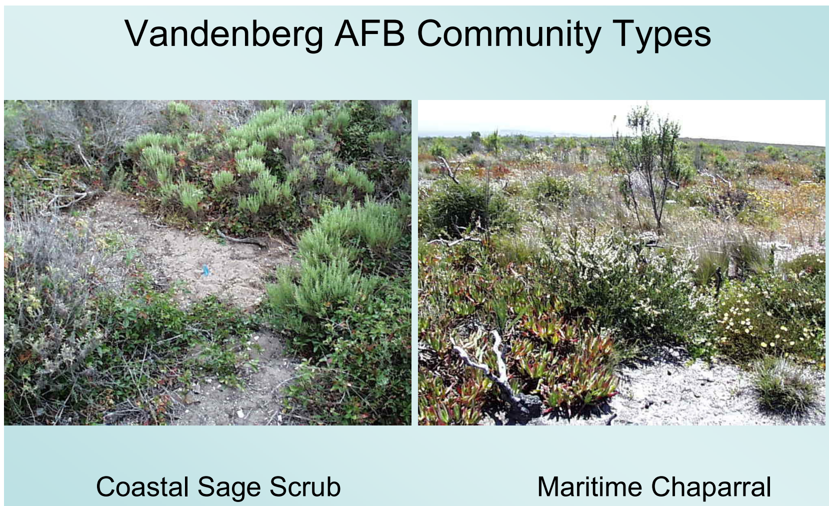

Figure 2. Two plant communities were studied: coastal sage scrub and maritime chaparral.

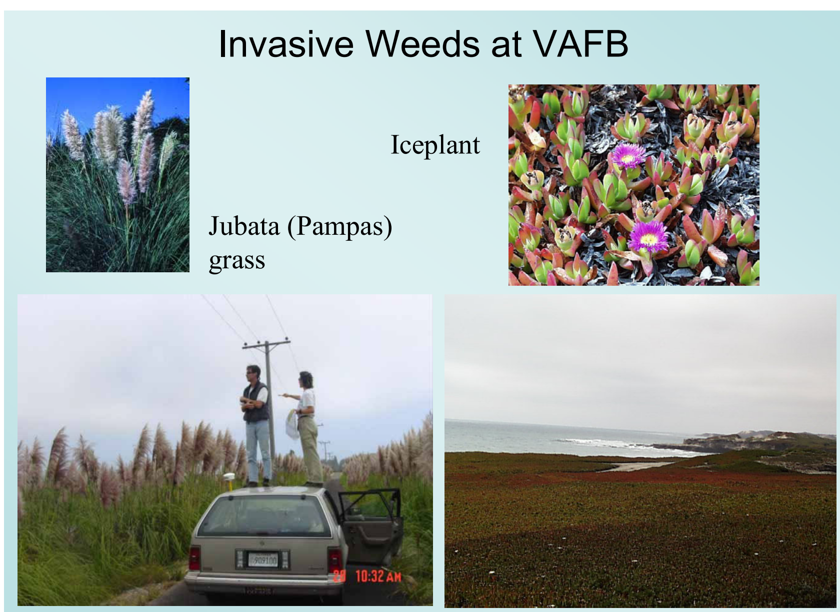

Figure 3. Jubata grass and iceplant are two invasive weed species threatening the coastal scrub and maritime chaparral communities.

3. Field Sampling and Analyses

We collected ecological data to 1) identify relationships between weed distributions to natural and anthropogenic disturbance at a small fine-grained scale, b) identify ecological conditions associated with target weeds and c) identify potential indicators or environmental correlates of target weed presence and abundance within specific natural communities.

Sampling was restricted to an area characterizing the range of invasion by the target weeds in scrub and chaparral communities from relatively intact native vegetation to large monocultures of iceplant or jubatagrass. Detailed species abundance data and rapidly assessed small-scale environmental/disturbance data were collected on the ground from 21 June to 26 July 2000. Sample sites were selected to represent the five community types: Scrub-Intact (n = 19), Scrub-Iceplant (n = 19), Chaparral-Intact (n = 19), Chaparral-Iceplant (n = 10), and Chaparral-Jubatagrass (n = 21).

In addition to the field data, we had access to the geographic information system (GIS) database maintained for VAFB. The GIS layers included hydrography, elevation, soil survey information, and fire history.

To characterize the vegetation, assess the relationships of communities and target species to environmental variables, and to identify potential fine-grain and landscape scale indicators for target species presence, we used an iterative multivariate approach integrating hierarchical clustering classification, nonmetric multidimensional scaling (NMS), multiple response permutation procedures (MRPP), Mantel tests, and classification and regression tree (CART) analysis using a common dissimilarity measure (Sorensen's distance) where appropriate (Urban et al. 2002). Plant community characteristics were described using classification procedures based on species cover data. The existence of associations between species composition, location, environmental variables and GIS variables were evaluated using Mantel tests. The first parameter space was composed of the site level environmental data collected on the ground at each site. The geographic parameter space was based on landscape scale information derived from GIS layers and included soils, hydrologic features, physiography, and fire history. The best suite of explanatory variables from the GIS subsets were aggregated into a final set for analyses.

Figures 4a and 4b are the NMS results for the chaparral and scrub communities. The numbering of NMS axes is arbitrary since axes scores are computed simultaneously. We present the pairs of axes that explain the greatest portion of the variation. Sites were ordinated on the basis of similar species composition. The position of the invasive species are coded on the ordination plot as: CORJUB= *Cortedaderia jubata*, CAREDU= *Carpobrotus edulis* and CARCHI= *Carpobrotus chilensis*. The radiating lines are vectors representing the direction and magnitude of the correlations of the environmental and GIS variables (R^2 >0.200) with the component scores for each axis (Note: The variable Hydrologic Segment Length in Figure 4b is the nearest distance to a stream, lake or other hydrologic feature on a GIS hydrology map). For example, the intact chaparral community was generally further away from fire breaks than the iceplant-invaded chaparral community. MRPP tests showed that all community types were significantly different at the 95% level. Mantel tests showed that sites with similar species composition were also similar with respect to the environmental and GIS variables. In addition, Mantel tests also showed that sites with similar environmental conditions were also similar with respect to the GIS variables.

After establishing the relationships between species composition, the environmental conditions at the site and the more spatially extensive GIS parameters we used CART analyses to investigate the relationships between the presence of the jubata grass and iceplant with the GIS variables across the full extent of the flight path.

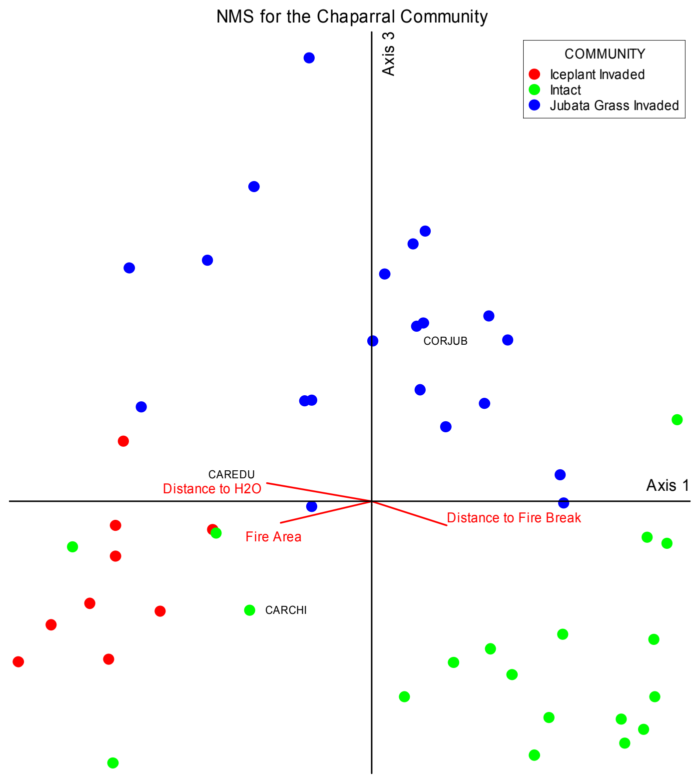

Figure 4a. Nonmetric multidimensional scaling ordination plot for the chaparral community showing sites, invasive weeds and significantly correlated variables.

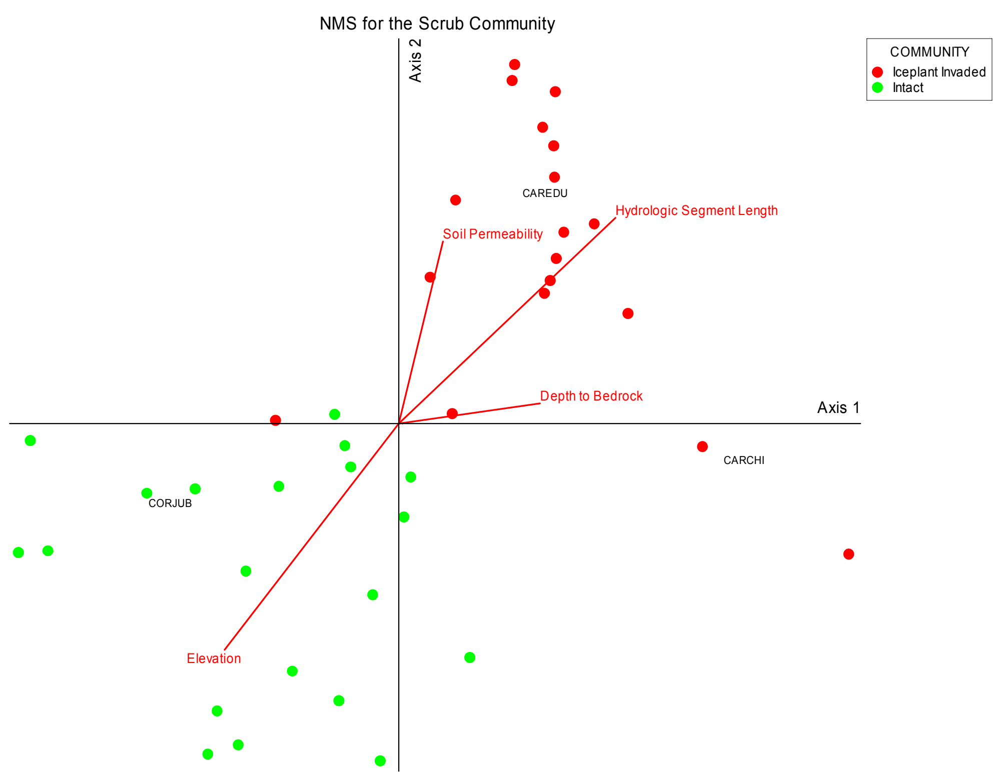

Figure 4b. Nonmetric multidimensional scaling ordination plot for the chaparral community showing sites, invasive weeds and significantly correlated variables.

4. CART Analyses

CART analysis has proven useful for description and predictions of ecological patterns and processes (Michaelson et al. 1994, De'ath and Fabricius 2000, Urban et al. 2002) and is a powerful tool for classifying multiple alternative environmental states. We chose CART for the reason that it can readily translate the predictor variables from the parameter space into a GIS where the results can be displayed geographically. The procedure offers a non-parametric, recursive approach well suited to complex, non-linear ecological data (De'ath and Fabricius 2000, Urban et al. 2002) including continuous, ranked and categorical variables. CART is tolerant of missing data and easy to interpret visually. CART analyses are performed using target species presence or abundance as the response variable and the environmental and GIS variables as explanatory variables.

CART trees explain the variation of a single response variable through repeated partitioning of a dataset into increasingly homogeneous groups using combinations of categorical or numeric explanatory variables (Breiman et al. 1984, Vayassieres et al. 2000). Trees are represented graphically beginning with the complete data set at the root node. Samples are divided into groups that form branches with each node characterized by a mean or typical value for the response variable, the number of samples in the group, and values of the explanatory variable defining the split that separates samples into the next lower branches (De'ath and Fabricius 2000, Urban et al. 2002). A single variable is identified at each branch and a variable can occur at multiple levels in the tree allowing for very complex interactions. Splitting continues until the tree is over-fitted to the data and the final groups occupy the leaves or terminal nodes of the diagram; the overlarge tree is then pruned back to the desired size using complexity plots (based on 10-fold cross-validation) and research goals (De'ath and Fabricius 2000, Urban et al. 2002). Pruning the tree minimizes over-generalizations often due to spurious or coincidental relationships peculiar to a sample set and not general to the population overall. The total percentage of correct classification is reported for the classification trees based on presence or absence of the target weeds.

To create the CART decision trees from which the rule-based models are derived we used the R statistical package (http://www.r-project.org). We used the decision tree module in ENVI to apply the CART rules to create invasion probability maps for the study area.

5. Development of the Uninvaded Plant Communities Map

Our ecological analysis indicated that the chaparral and scrub communities had different responses to the encroachment of the invasive weeds and should be modeled separately. Our first step in the modeling procedure was to develop a community-level map to use as a mask to separate the scrub areas from the chaparral areas during the modeling. We used the CART techniques with the ecological field plots as the dependent variable and the GIS layers as the independent variable.

Figure 5 shows the decision tree and resulting GIS image map depicting the spatial distribution of the scrub and chaparral plant communities in the area of the flight paths. This "uninvaded" map shows the distribution of the two communities without consideration of the target invasive species. The CART decision tree correctly classified 96.3% of the 644 plots (a combination of the field plots and ground-truthing plots where the plant community identification had been verified on the ground). Differences in elevation and soil type were significant predictors for the scrub and chaparral communities in this area. A rule-based model was constructed based on the set of "if-then-else" rules derived from the classification tree. Each pixel in the image was then input into the model and classified as to scrub or chaparral.

We visually compared our results for the uninvaded scrub and chaparral communities map based on the CART model with a field ecology map provided to us by VAFB personnel (Figure 6). The two maps are in remarkably close agreement.

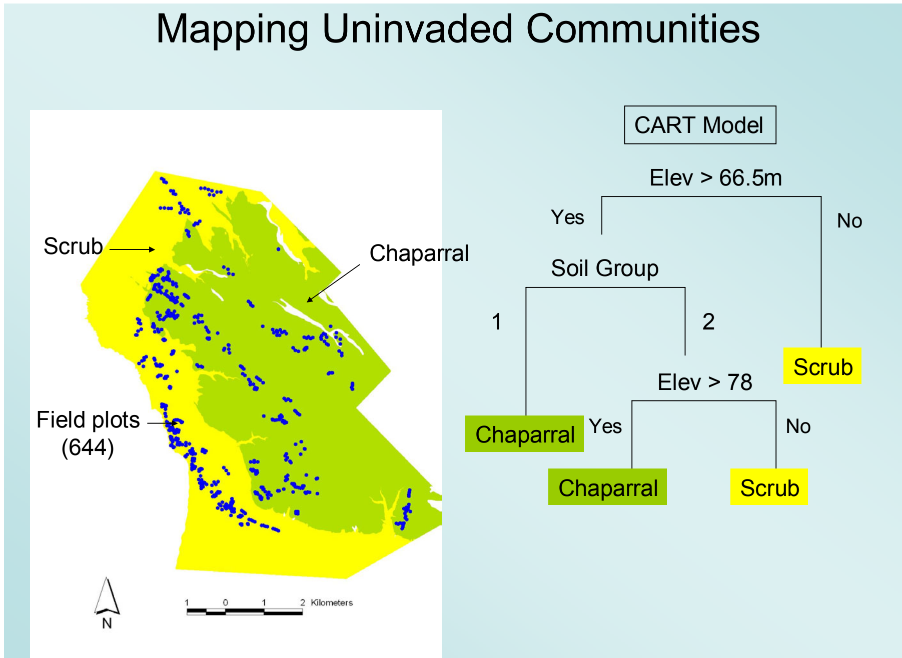

Figure 5. Predicted scrub and chaparral distribution map derived from the CART analysis decision tree.

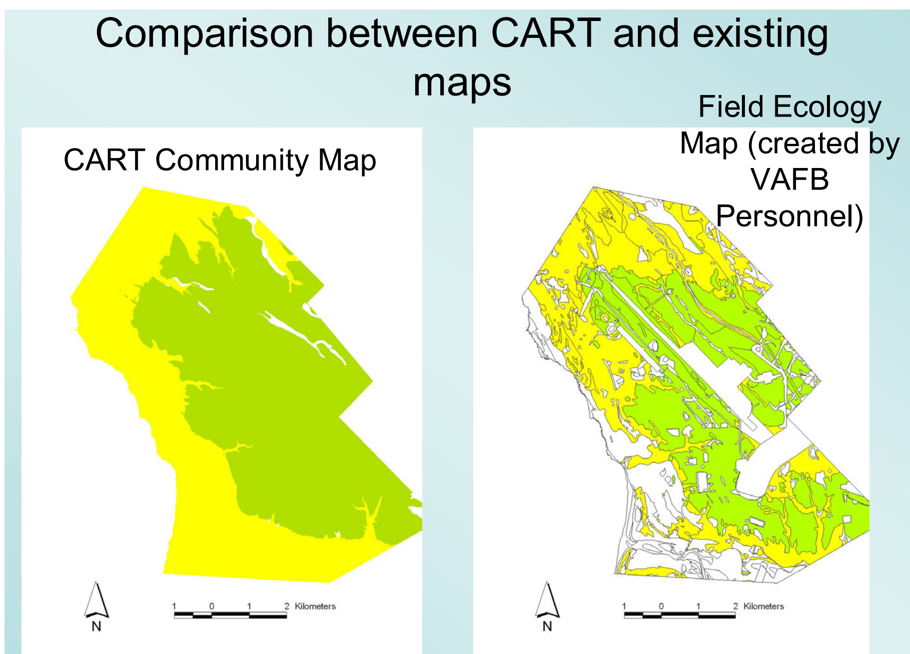

Figure 6. Comparison of the predicted community map and a vegetation map for VAFB.

6. Development of the Hyperspectral Imagery Map

We used AVIRIS imagery (3.7 m pixels) collected on September 6, 2000 to create a map of the location of the pampas grass and the iceplant as described in Underwood et al, 2002. The classified map is shown in Figure 7. The classification was 70% successful in identifying pampas grass and 98% successful in identifying iceplant.

7. Development of a Map of Invasion Risk

We used the AVIRIS map as the dependent variable for the invasion risk modeling. Each pixel in the map has a possible value of ”uninvaded”, pampas, or ice plant. To create a statistical sampling set, we used a “virtual sampling” technique. We developed a sampling grid of points spaced 100 m apart that could overlaid on top of the AVIRIS to extract the value at each grid point as shown in Figure 8. We then used the same grid with the GIS layers to extract the independent (predictor) variables. This virtual

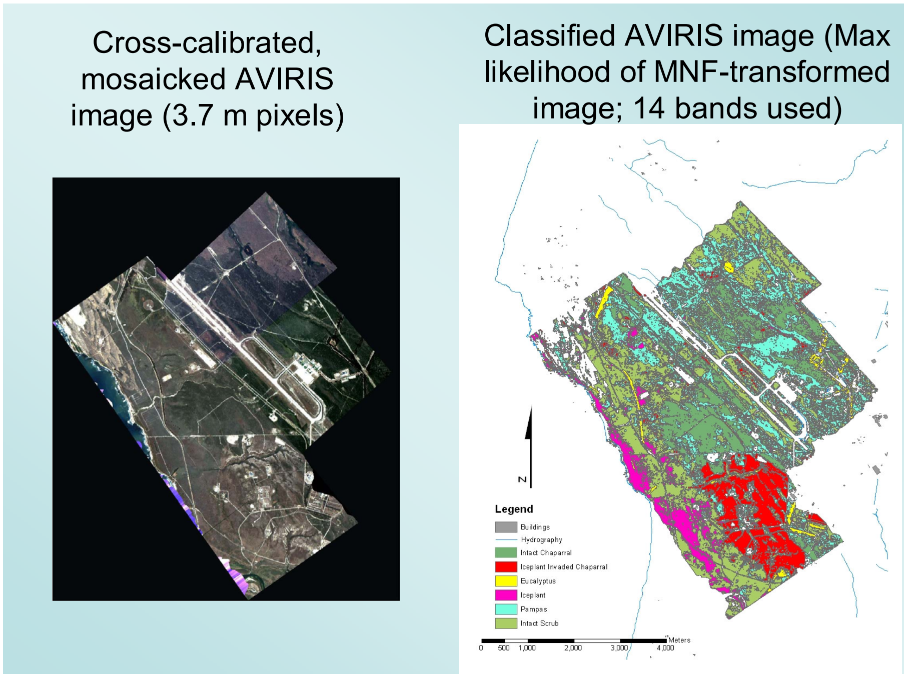

Figure 7. AVIRIS imagery and vegetation classification map for Vandenberg AFB.

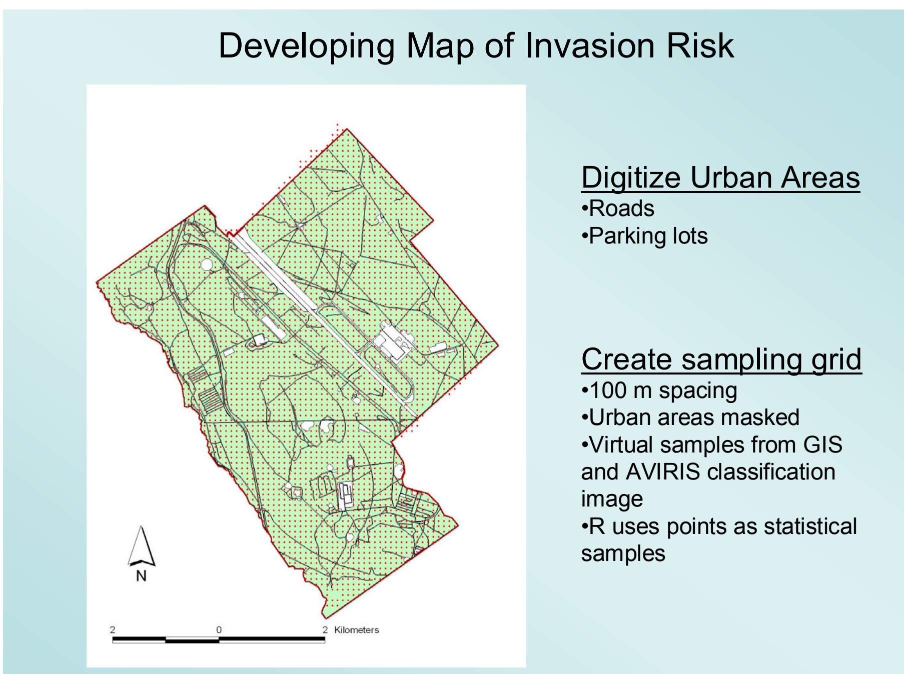

Figure 8. Virtual sampling technique for CART analysis.

sampling grid gave us complete coverage over the flight path areas. We digitized polygons around the urban areas to remove these grid points from the analysis.

We hypothesized that the invasion of weeds could be predicted by two general classes of predictor variables. First are those parameters that are the natural landscape variables not under human control (Figure 9). This includes such things as elevation, soil type, soil permeability, slope, aspect, etc. The second class is those parameters that are anthropogenic (Figure 10). This second class includes, control burns, roads, buildings, pastures, ditches, etc. The analysis described below is that for the natural

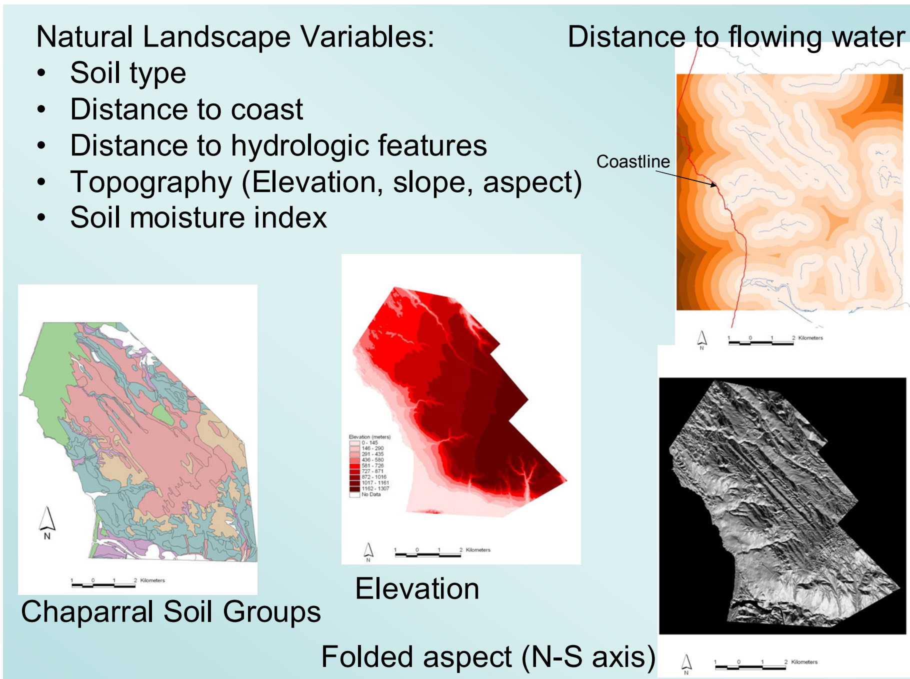

Figure 9. Examples of natural landscape GIS layers used in the CART analysis at Vandenberg AFB.

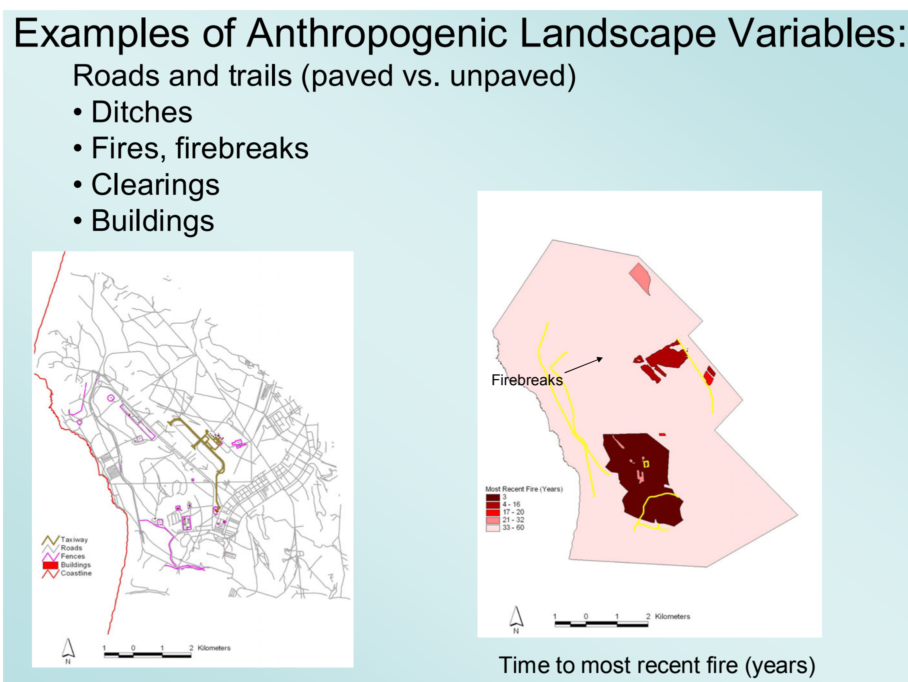

Figure 10. Examples of anthropogenic GIS layers to be used in the CART analysis at Vandenberg AFB.

landscape variables. We will be investigating the anthropogenic variables at a later date. Our approach was to first develop maps predicting the distribution of the invasive weeds using a CART model with the natural landscape variables and then investigate changes in the distribution after including the anthropogenic variables. Using points on the sampling grid we populated a database by first extracting the plant community predicted by the AVIRIS image and then extracted variables of interest from the various GIS layers provided by VAFB. Separate CART models were developed for the scrub and chaparral associated communities.

Figure 11 shows the classification decision tree resulting from the CART analysis for the scrub community. Percentages at the termination of each node represent the percentage of sites that follow that branch of the tree to its terminus. Of the sites that were within 245 m of the coast 76% of them were not invaded. The remaining sites beyond 245 m of the coast were separated using a complex suite of predictors that included the type of soil, distance to flowing water, elevation, percent slope and aspect. This decision tree correctly classified 86% of the 1068 sampling grid points.

Figure 12 shows the classification decision tree resulting from the CART analysis for the chaparral community. This tree is a complex interaction between the soil types, distance to coast, and elevation. Sixty-seven percent of the 1864 sampling grid points were correctly classified by this tree.

The CART rules shown in Figures 11 and 12 were used in the ENVI Decision Tree module as shown in Figure 13. Each pixel was assigned a probability of invasion based on which endpoint on the tree the pixel represents. A comparison between the original AVIRIS image and the CART results for iceplant and pampas grass is shown in Figures 14 and 15, respectively.

8. Summary

The CART modeling technique provides a powerful, highly intuitive method to look at spatial ecological data. Once the variables promoting weed invasion have been identified by the CART model, high risk areas can be plotted out in the GIS as in the examples shown in Figures 14 and 15. Alternative scenarios can be explored by modifying the GIS layers, then running the Decision Tree to see the effects of a planned action.

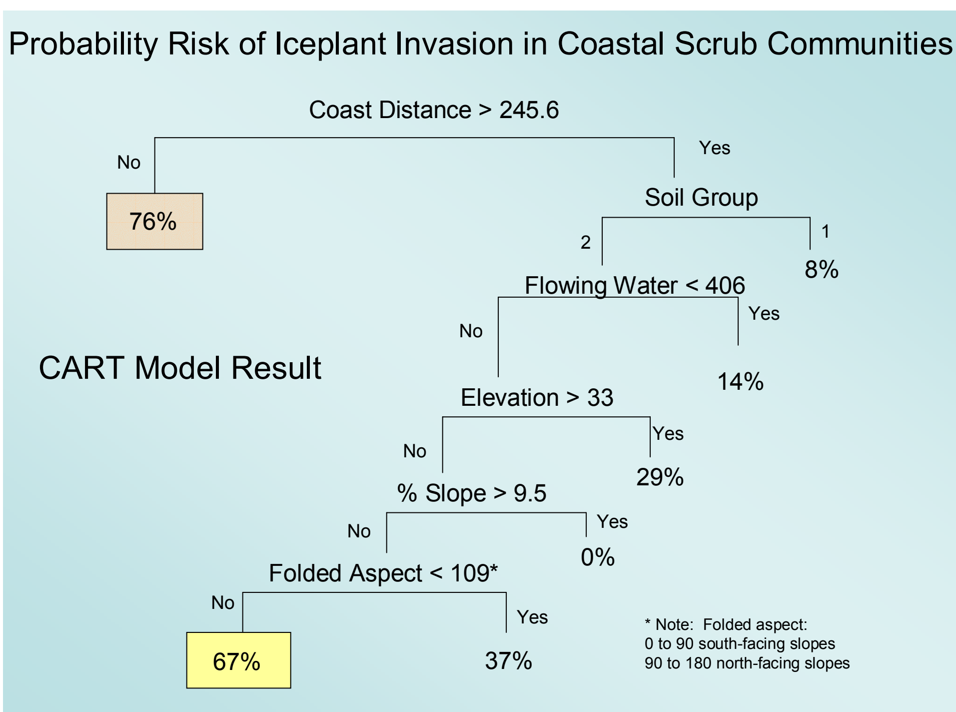

Figure 11: CART tree developed for iceplant invasion into scrub community at Vandenberg AFB.

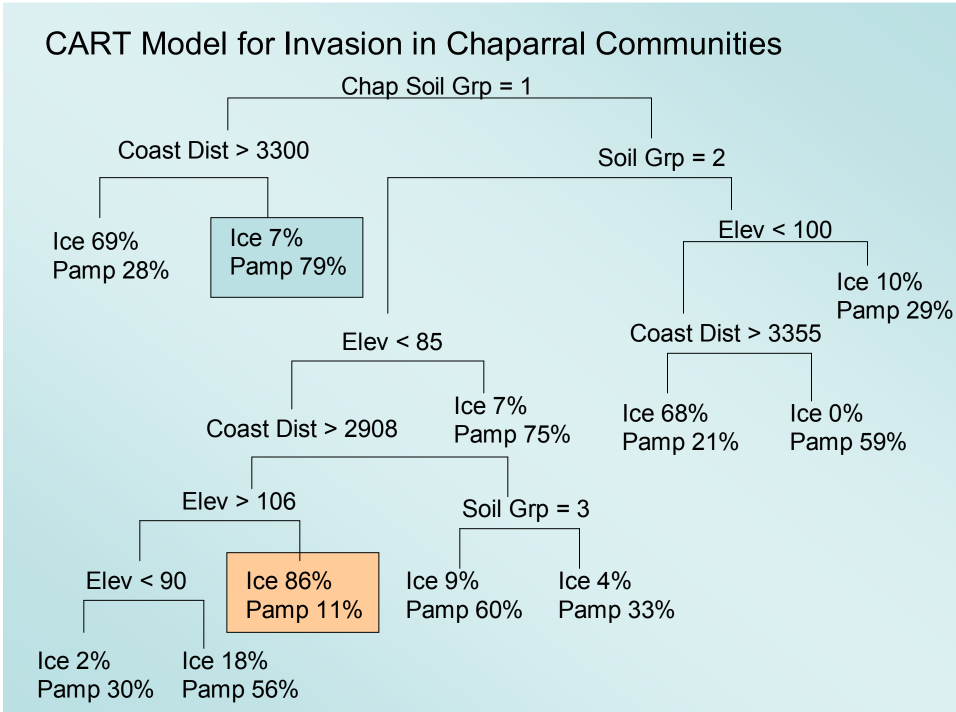

Figure 12. CART tree developed for iceplant and pampas grass invasion into chaparral community at Vandenberg AFB.

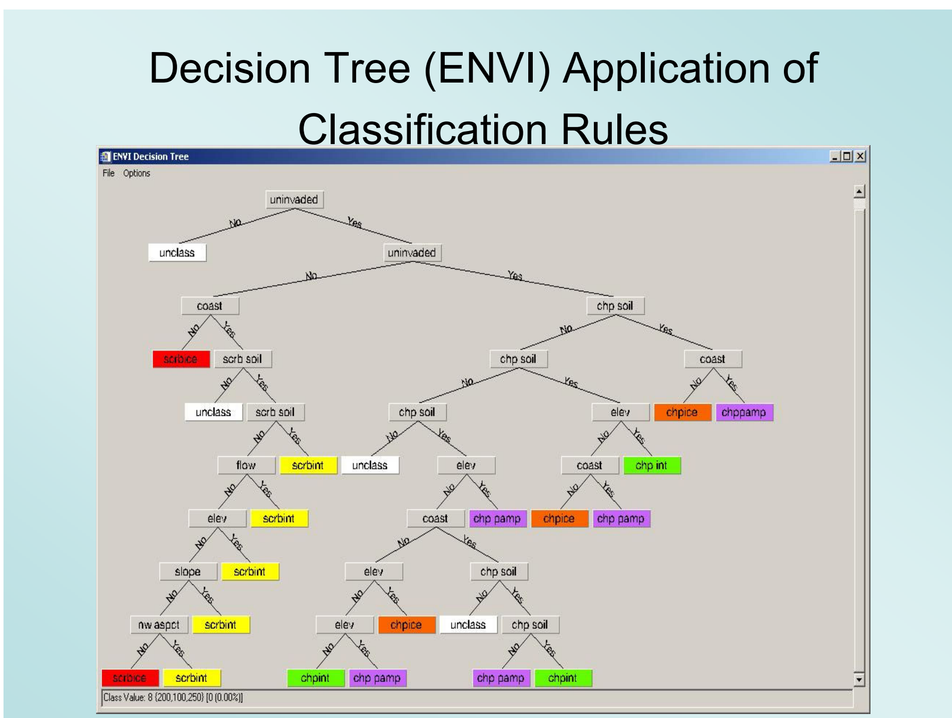

Figure 13. Example of application of CART rules in the Decision Tree module of ENVI.

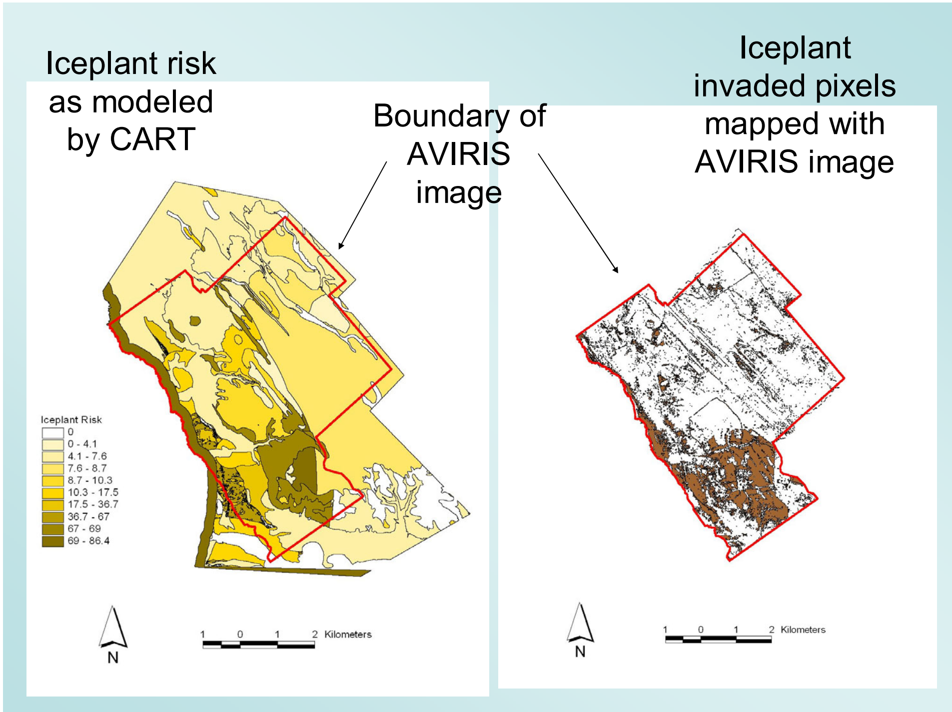

Figure 14. Comparison between CART-derived iceplant invasion risk and original AVIRIS map.

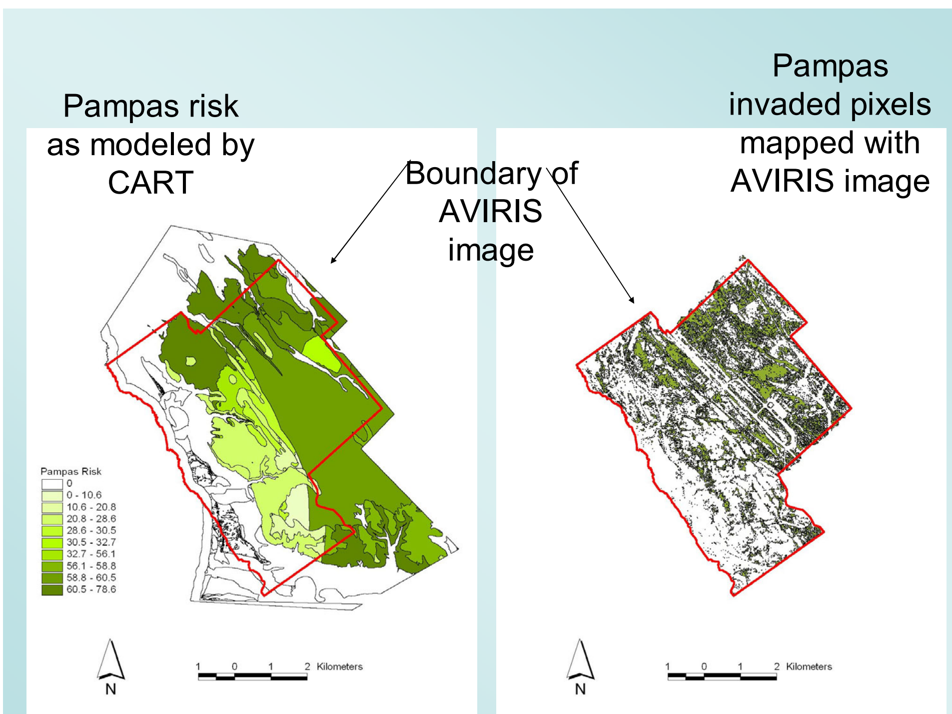

Figure 15. Comparison between CART-derived pampas grass invasion risk and original AVIRIS map.

9. References

Babbitt, B. 1998. Statement by the Secretary of the Interior on invasive alien species. *Proc. Int. Weed Symp*. BLM Weed Page, April 8-10, 1998.

Breiman, L., J.H. Friedman, R.A. Olshen, and C.J. Stone. 1984. *Classification and regression trees.* New York: Chapman and Hall.

Coulombe, H.N. and C.F. Cooper. 1976. *Ecological Assessment of VAFB, California: Evaluation and Recommendations, Volume 1 and 2*. San Diego State University, Center for Regional Environmental Studies for U.S. Air Force. HQ, SAMSO, AFCEC-TR-76-15.

De'ath, G., and K.E. Fabricius. 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. *Ecology* 81:3178-3192.

Keil, D.J., and V.L. Holland. 1998. *Documented flora of Vandenberg Air Force Base, Santa Barbara County, California*. The Nature Conservancy and Vandenberg Air Force Base

Michaelson, J., D.S. Schimel, M.A. Friedl, F.W. Frank and R.C. Dubayah. 1994. Regression tree analysis of satellite and terrain data to guide vegetation sampling and surveys. *Journal of Vegetation Science* 5:673-686.

Pimentel, D., McNair, S., Janecka, J., Wightman, J., Simmonds, C., O'Connell, C., Wong, E., Russel, L., Zern, J. Aquino, T., and T. Tsomondo. 2000. Economic and environmental threats of alien plant, animal and microbe invasions. *Agri. Ecosys. & Environ.* 84: 1-20.

TNC. 1991. *Fish and wildlife management plan for Vandenberg Air Force Base*; Revision Number 3 for Plan Period August 1991 to August 1996. The Nature Conservancy

Urban, D., S. Goslee, K. Pierce, and T. Lookingbill. 2002. Extending community ecology to landscapes. *Ecoscience* 9 (2):200-212.

Underwood E.C., S.L. Ustin and D. DiPietro. 2002. Mapping non-native plants (iceplant) using hyperspectral imagery. *11th Earth Science Airborne Workshop*, Jet Propulsion Laboratory, Pasadena, CA March 31-April 2, 2002.

Vayassieres, M.P., R.E. Plant, B.H. Allen-Diaz. 2000. Classification trees: an alternative non-parametric approach for predicting species distributions. *Journal of Vegetation Science* 11:679-694.