



Rapid sampling of plant species composition for assessing vegetation patterns in rugged terrain

Ross K. Meentemeyer and Aaron Moody

Department of Geography, University of North Carolina at Chapel Hill, NC 27599-3320, USA

Received 9 April 1999; Revised 30 September 1999; Accepted 17 November 1999

Key words: chaparral, landscape scale, remote sensing, vegetation pattern, vegetation sampling

Abstract

Detailed species composition data are rapidly collected using a high-powered telescope from remote vantage points at two scales: *site level* and *patch level*. Patches constitute areas of homogeneous vegetation composition. Multiple samples of species composition are randomly located within the patches. These data are used as site-level data and are also aggregated to provide species composition data at the patch level. The site- and patch-level data are spatially integrated with high resolution (10 m), topographically-derived fields of environmental conditions, such as solar radiation, air temperature, and topographic moisture index in order to evaluate the applicability of the sampling method for modeling relationships between species composition and environmental processes.

The methodology provides a balance between sampling efficiency and the accuracy of field data. Application of the method is appropriate for environments where terrain and canopy characteristics permit open visibility of the landscape. We evaluate the nature of data resulting from an implementation of the remote sampling methodology in a steep watershed dominated by closed-canopy chaparral. Analyses indicate that there is minimal bias associated with scaling the data from the site level to the patch level, despite variable patch sizes. Analysis of variance and correlation tests show that the internal floristic and environmental variability of patches is low and stable across the entire sample of patches. Comparison of regression tree models of species cover at the two scales indicates that there is little scale-dependence in the ecological processes that govern patterns of species composition between the site level and patch level. High explanatory power of the regression tree models suggests that the vegetation data are characterized at an appropriate scale to model landscape-level patterns of species composition as driven by topographically-mediated processes. Patch-level sampling reduces the influence of local stochasticity and micro-scale processes. Comparison of models between the two scales can be useful for assessing the processes and associated scales of variability governing spatial patterns of plant species.

Introduction

Sampling strategy is a critical element of field-based studies in landscape ecology (e.g., Stohlgren et al. 1997). We present a method for sampling vegetation composition in environments where field access is limited, or where rapid surveys are required. The methodology is designed to provide numerous, well distributed samples for modeling functional relationships between patterns in species composition and topographically regulated processes, such as fluxes of energy and moisture across a landscape. Models of this type can be used to better understand ecolog-

ical processes – such as resource-use, competition, seed dispersal, and germination – that govern spatial patterns in plant composition, as well as to predict vegetation redistribution due to environmental change (reviewed in Franklin 1995).

In general, modeling relationships between plant-community composition and environmental conditions requires vegetation samples and distributed environmental data that are of sufficient resolution to capture the underlying ecological processes, and of sufficient accuracy so that results are not obscured by noise associated with data error (Davis and Goetz 1990; Bolstad et al. 1998). However, implicit trade-

offs exist between data resolution and data accuracy on the one hand, and data collection costs and the need for adequate sample sizes on the other.

In situ sampling methods, such as nested plots, provide accurate and floristically detailed data (Kent and Coker 1992) but are labor intensive and time consuming, especially with rugged terrain or impenetrable vegetation. Other sampling methods, such as relevés, circular plots, line-intercepts and transects, provide faster alternatives, but still require *in situ* access. These limitations prohibit high sampling densities and often lead to a concentration of samples along roadsides and trails. As a result, it can be difficult to obtain samples of species composition that are sufficient in number and adequate in distribution for assessing spatial patterns in species composition at landscape scales. While rapid ecological assessment can be accomplished via satellite remote sensing, aerial photography, or airborne video imagery (e.g., Pickup et al. 1995), these data may be too generalized to resolve species composition.

We propose a method for rapidly sampling overstory species composition in environments where terrain and canopy characteristics prohibit access, but permit open visibility of the landscape. Species composition data are collected using a high-powered telescope from remote vantage points at two scales: *site level* and *patch level*. Randomly located points within field-defined patches of homogeneous vegetation composition are used as site-level data. The site-level samples are aggregated to provide species composition data at the patch level. In a raster geographic information system (GIS), patches represent multi-cell spatial units, and sites represent individual cells. These data are spatially integrated with a high resolution (10 m) digital elevation model (DEM) and terrain-derived fields of environmental variables, including solar radiation, air temperature, precipitation, and topographic wetness index. The GIS database provides the basis for evaluating the sampling methodology and demonstrating the method's applicability for analyzing vegetation patterns.

Application of the proposed methodology is appropriate for environments with open visibility of surrounding terrain. For example, woodland and shrubland types, semi-desert scrub, desert vegetation, and most Mediterranean-type vegetation formations often permit good visibility due to sparse and/or low stature canopies and rugged terrain. Although the method would be less effective in forested landscapes, it could be useful along roadside openings

for quick surveys. In this paper, we evaluate the nature of data resulting from an implementation of the remote sampling methodology in a steep watershed dominated by closed-canopy chaparral in the Santa Ynez Mountains, California. Our analysis is focused upon four questions which address the effectiveness of the sampling methodology for studying process-driven relationships between species composition and environmental conditions at the landscape scale:

- How reliable is remote species identification?
- Are patches defined such that their internal environmental and floristic variability is low and that intra-patch variability is stable across the entire sample of patches? That is, is there bias in environmental variance and species composition due to patch characteristics (e.g., size)?
- Are patches defined such that modeled relationships between species composition and environmental conditions reflect *landscape-scale* processes, as opposed to finer- or broader-scale processes?
- Is there scale dependence in relationships between species composition and environmental conditions within the range of scales that we sample (site vs. patch)? If these relationships are not scale dependent this implies that the processes governing species composition are consistent within our scale range.

The following sections describe and evaluate the sampling methodology as applied to the test site.

Description of the sampling method

Study area

The study area is a 7.1 km² watershed (Romero Canyon) located on the southern flank of the Santa Ynez Mountains, California (Figure 1). Romero Canyon is dominated by closed-canopy chaparral, typical of southern California and Baja, which develops a heavy thicket of dead understory biomass, distributed on steep, and often loose terrain. These conditions discourage movement, and have prevented most chaparral researchers from collecting a large number of field samples. Although chaparral is difficult to sample *in situ*, rugged topography and the low canopy (~1–3 m) provide open visibility for remote sampling of vegetation on nearby terrain.

The steep, dissected topography of the Santa Ynez is carved from a young fault block, composed mainly

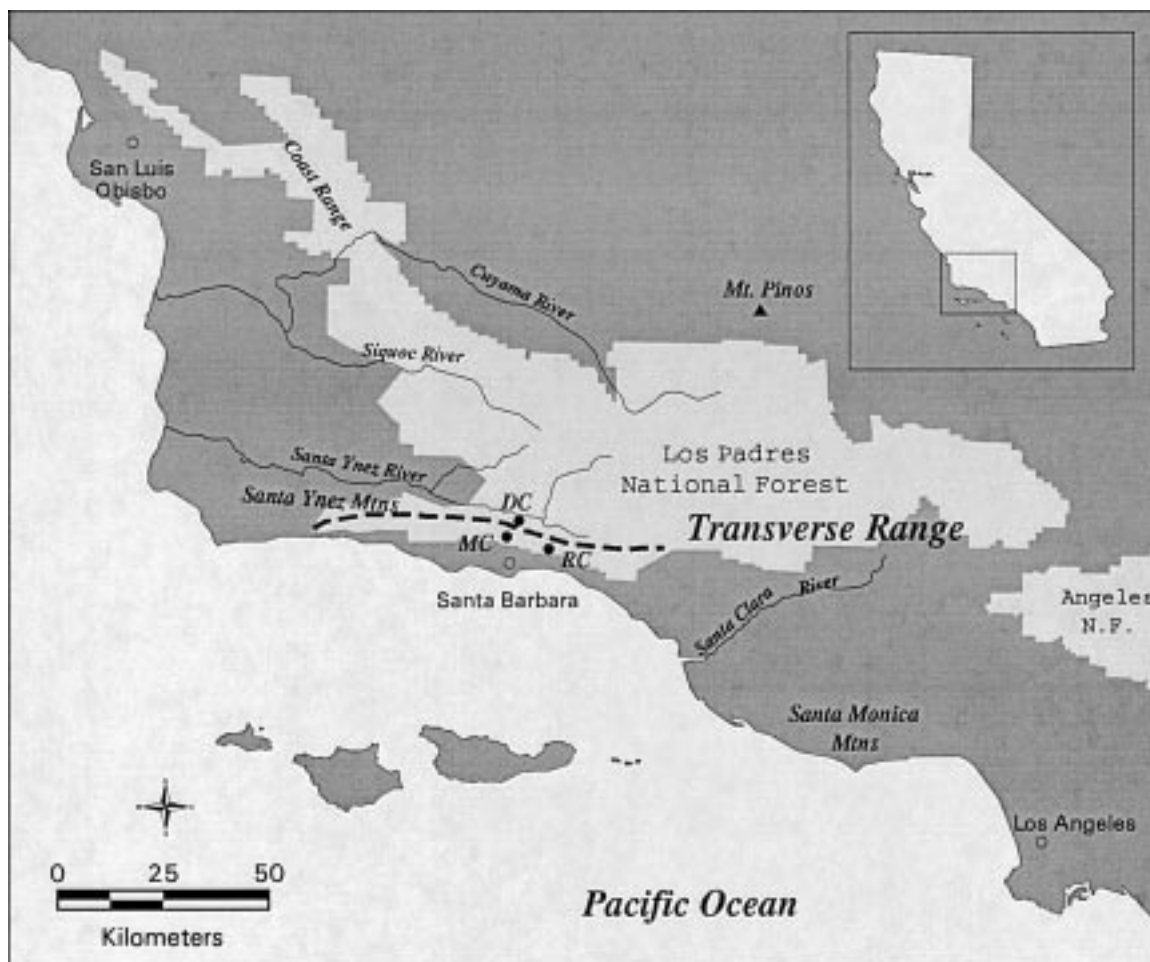


Figure 1. The study region. The test watershed (RC) is located on the southern flank of the Santa Ynez Mountains, California. The dashed line represents the main ridge of the Santa Ynez Range.

of Tertiary marine sediment and quaternary deposits, with common outcrops on steep sandstone slopes (Dibblee 1986). The Mediterranean-type climate is characterized by cool, wet winters and hot, dry summers. The base and upper ridge of Romero Canyon are 4 and 6.5 km from the Pacific coastline, respectively. This places the study within the maritime fringe that experiences onshore flow of cool moist air from the Pacific Ocean. The study area was completely burned in the Romero Canyon fire of 1971.

Eighty-seven patch-level samples of species composition were collected during June, August and December 1998. The samples are distributed from the boundary of the Los Padres National Forest, at an elevation of 244 m, to the upper ridge of the watershed, at 1067 m. The study area was chosen based on accessibility for sampling, topographic and geo-

logic variability, and common successional maturity of the vegetation. At our scale of interest, sampling within watersheds is practical because it provides a range of environmental variability that is constrained to an ecologically meaningful, yet manageable boundary. Additionally, sampling within watersheds can help control for the effects of broader scale patterns in climate, fire history and landuse history.

Field sampling

Mapping vegetation patches

Vegetation patches are identified in the field from a remote viewing location and mapped on digitized USGS elevation contours (1:24,000) overlaid with 20 m grids. Patch boundaries are delineated where pronounced transitions occur in either floristic composition or in the relative abundance of dominant species.

We use the following guidelines to identify patch boundaries. These criteria are not strictly quantified, but are used as guiding parameters to help maintain consistency in patch identification. Stands where a single species composes 60% or more of the relative canopy cover are considered dominant species patches. In these cases, a patch boundary is mapped if either a significant change in the dominant species occurs, or a transition to a mixed species composition occurs. Boundaries are also identified where the cover of any species changes by 20% or more. Finally, boundaries are mapped where significant changes occur in absolute cover, regardless of species composition. We use a minimum mapping unit (MMU) of 0.25 ha. For this environment, trial and error suggested that this MMU distinguishes topographically mediated resource gradients and filters out finer-scale floristic variability that does not reflect landscape-scale processes. We do not set a maximum size threshold. Patches are distributed in order to capture the range of elevations and topographic positions. Patch sizes range from 0.25 to 7.9 ha (mean = 1.2 ha).

Species composition sampling

Within each patch, two to five sites (depending on patch area) are randomly located from a regular array of grid points distributed at 20 m intervals. A zoom telescope (15× to 45×) is used to sample species composition for each site from a remote viewing location. Site-level species composition is determined by aggregating data collected within three to four telescope fields of view (FOV) that are clustered around the site. Three FOVs were occasionally used in cases where a FOV falls outside a patch boundary. This process is repeated for each randomly selected site within a patch. In order to reduce spatial autocorrelation among samples within a patch, a randomly selected site is rejected if it falls adjacent to a site that has already been sampled. The following data are collected for each FOV:

- A catalogue of species present.
- Relative percent canopy cover of each species.
- Total canopy cover.
- Percent of exposed bedrock.
- Disturbance characteristics.

The randomly selected sites within mapped patches can be used as site-level data or they can be aggregated (i.e., averaged) to estimate species composition for their parent patch. In the GIS database, patches are multi-cell (10 m) spatial units

and sites are represented as individual 10 m cells located with the parent patch.

Although FOV sizes vary, we attempt to constrain the size to roughly 100 m² by controlling viewing distance (< 300 m), avoiding overly oblique views (>45°), and using the zoom feature of the telescope. Locating FOVs is aided by navigating on high resolution elevation contours overlaid with the 20 m grid in conjunction with clinometer and compass measurements. The location of site-level samples, however, will still contain error. As an alternative, real-time differentially corrected global positioning system (GPS) measurements in combination with a reflectorless laser range finder, could be used to more precisely locate sites and measure view distances and view angles (Xu et al. 1999). Although this technology can be practical, we have not used this approach because the mountainous terrain of the study area frequently obstructs signals from GPS satellites.

Plant traits used for remote identification

Remote species identification is based on unique qualitative plant features that enable species differentiation and can be detected from a distance. These features include bark color and texture; canopy architecture; bud, flower, and new growth characteristics; and leaf shape, size, color, tone, and orientation. Table 1 lists some species that commonly occur in the Santa Ynez Mountains and the plant traits used for remote identification. The canopy of *Adenostoma fasciculatum*, for example, appears feathery with small, linear leaves.

Remote identification requires field experience, and species vary in the ease with which they can be identified. In chaparral, for example, species of *Arctostaphylos* are conspicuous due to their smooth, pale-green leaves and smooth, red bark (Table 1). In contrast, some species of *Quercus* (*Q. berberidifolia*, *Q. wislizenii*, and *Q. chrysolepis*) exhibit high phenotypic variability and are more difficult to differentiate from each other remotely. Understanding the phenology of species can be helpful for remote identification. *Ceanothus crassifolius*, for example, is conspicuous during late summer when the lighter toned lower leaf epidermis is visible due to increased leaf inclination in response to elevated heat loads. Reproductive cycles are also useful. *Ceanothus megacarpus*, *C. spinosus*, and *C. oliganthus* have staggered blooming periods through late winter and spring months. *Heteromeles arbutifolia* produces large clusters of red berries during December and January. The flowers of *A. fas-*

Table 1. Some common plant species in the Santa Ynez Mountains, California and the traits used for remote identification.

Code	Latin binomial	Common name	Remote characteristics
ADFA	<i>Adenostoma fasciculatum</i>	Chamise	Leaves small, linear, fascicled. Canopy feathery, Panicles of white and dried flowers concurrently through late spring and early summer. Commonly associated w/ <i>Arctostaphylos</i> spp. or is in homogeneous stands. Frequently pokes thin stems through stouter canopies of other shrubs.
ARGN	<i>Arctostaphylos glandulosa</i>	Eastwood Manzanita	Dramatic contrast w/ other spp. Stems smooth and red. Leaves ovate-lanceolate, coin-like, pale-green. Canopy broad, rounded and dense. White flowers in winter. Flowers hang down in clusters. May be confused w/ ARGU.
CEBE	<i>Cercocarpus betuloides</i>	Mountain Mahogany	Canopy composed of silvery vertical and thin stems, esp. at top. Tall. Bark gray. Leaves med. to dark green, small, deep veins. Flowers in spring w/out petals leaving an achene w/ long, curly tail. May be confused w/ CEME.
CEME	<i>Ceanothus megacarpus</i>	Bigpod Ceanothus	Leaves appear grayish-green remotely. Similar tone as CEBE. Canopy scrubby. No verticality of upper stems. Tall. Leaves small. Bark gray. White flowers usually late winter.
CECR	<i>Ceanothus crassifolius</i>	Hoaryleaf Ceanothus	Leaves opposite, thick-leathery, med. small, round, dull due to white hairs, distinct green. Pale underside. Prominent veins on underside. Blooms March-April. Leaves inclined late summer, whitish.
CESP	<i>Ceanothus spinosus</i>	Greenbark	Bark smooth and bright green on young branches. Branches terminate as distinct spines. Leaves greener than CEME. Frequently associated w/ CEME.
HEAR	<i>Heteromeles arbutifolia</i>	Toyon	Canopy tree-like. Leaves elliptical-oblong, relatively large, vertical reaching, dark green, toothed, and not shiny. Berries red in winter.
MALA	<i>Malosma laurina</i>	Laural Sumac	Leaves med.-large, lance-oblong, arched back, wavy and somewhat folded along midrib. Dead flower stalks persist. Leaves often killed by frost – yellowish.
PRIL	<i>Prunus ilicifolia</i>	Hollyleaf Cherry	Leaves shiny, dark-green, usually wavy, and oblong to flat and round 1–3 cm in size. Flowers dense, white clusters in early spring. Fruit large and round (1 cm).
UMCA	<i>Umbellularia californica</i>	California Bay	Leaves oblong, simple, med.-large. Leaf clusters star-like. Crown is symmetrical, tapers to a point. Canopy bends and leaves shutter in wind. Tree-like to tree. Tall.

ciculatum are white from April to September, but die and turn a distinct rusty color during autumn.

Field-validation of species identification

Because the sampling method represents a trade-off between sampling efficiency and data accuracy, a certain degree of species misidentification is expected. In order to determine the reliability of remote identification, we have performed validation tests on a set of 15 FOVs located across a range of topographic and floristic settings. For the tested FOVs, a worker located inside the FOV made a list of species present and indicated individual plants (using a long pole with a red flag attached) for identification by the remotely positioned interpreter. The tested FOVs have an average of 5 species each, and a total of 18 species were found within the entire set. The individual plants were identified by the remotely positioned interpreter with an accuracy of 90% (Table 2). For comparison to the in situ worker's list of species present, the remote interpreter also tabulated all species observed within the FOV. Species composition was correctly interpreted in all 15 FOVs.

Analysis methods

Derivation of environmental variables

Seven environmental variables were derived from climate data and a high resolution (10 m) DEM in order to evaluate the suitability of the sampling method for assessing relationships between species composition and landscape-scale variability in environmental conditions. The derived variables include two topoclimate variables: average minimum monthly temperature for the coldest month (January; TMIN) and average maximum temperature for the warmest month (July; TMAX); two solar radiation variables: clear-sky solar insolation integrated over winter (December–February; PSI_W) and summer (June–August; PSI_S); and three terrain variables: topographic moisture index (TMI), slope angle (SLP), and landform curvature (CURV). Percent exposed bedrock (ROCK) is estimated for each telescope FOV and aggregated to the site and patch levels. These variables were chosen to identify primary environmental regimes that control or are related to the distribution of chaparral shrubs (Wells 1962; Miller et al. 1983; Franklin 1998). In the GIS database, site-level variables are derived on a cell-level basis and patch-level variables are determined

by averaging cells that lie within a patch. Intra-patch variance is also calculated for each variable. These are referred to as: TMINvar, TMAXvar, PSI_Wvar, PSI_Svar, TMIvar, SLPvar, and CURVvar. Variance is not calculated for ROCK.

Lapse rates – required to estimate TMAX and TMIN – are derived using mesoscale fields of 850 m average monthly air temperature and precipitation (1965–1995) provided by J. Michaelsen (unpublished data). These data are interpolated from a network of climate stations over the southwest ecoregion using location (latitude, longitude), elevation, and a topographic barrier index to capture rainshadow effects. We regressed the climate fields against elevation in the study watershed ($n = 14$) to yield elevational lapse rates for TMIN (3.94 °C/1 km) and TMAX (3.6 °C/1 km). R-square values of TMIN and TMAX versus elevation are 0.67 and 0.80, respectively. The lapse rates were then applied to each cell of the 10 m DEM. Following the equations of the Mountain Microclimate Simulation Model (MTCLIM; Running et al. 1987), TMAX was further adjusted at each 10 m grid cell in order to account for the effect of solar radiation on air temperature.

Incident radiation for each grid cell in the DEM is estimated using the Atmospheric and Topographic Model of topographically distributed solar radiation (Dubayah 1992). Maps of solar radiation are estimated for the winter (PSI_W) and summer seasons (PSI_S) by integrating direct and diffuse radiation at hourly intervals over a season. The model is run using simplifying assumptions of uniform surface albedo for shrub vegetation and clear-sky atmospheric transmissivity. The topographic moisture index (TMI) (Beven and Kirkby 1979) is used to characterize topographic effects on soil moisture distribution. TMI is the natural logarithm of the ratio between a location's upslope drainage area and the slope gradient of that location. The curvature variable (CURV) characterizes topographic convexity (positive values) and concavity (negative values).

Statistical procedures

Ideally, patches are defined such that their internal floristic and environmental variability is low and that intra-patch variability is stable across the entire sample of patches. Analysis of variance tests are used to determine if intra-patch variability in the environmental variables and species richness (the number of woody species present at a site) is low relative to between patch variability. We also examine cor-

Table 2. Validation results of species identification tests from 15 FOVs and 18 species.

Code	Latin binomial	Number correct	Number of observations
ADFA	<i>Adenostoma fasciculatum</i>	5	5
ARGN	<i>Arctostaphylos glandulosa</i>	3	3
ARGU	<i>Arctostaphylos glauca</i>	1	1
BRNI	<i>Brassica nigra</i>	1	1
CAMA	<i>Calystegia macrostegia</i>	2	2
CEBE	<i>Cercocarpus betuloides</i>	7	8
CEME	<i>Ceanothus megacarpus</i>	11	14
CESP	<i>Ceanothus spinosus</i>	9	9
FRDI	<i>Fraxinus dipetala</i>	3	4
HEAR	<i>Heteromeles arbutifolia</i>	6	6
LOSC	<i>Lotus scoparius</i>	0	1
PENS	<i>Penstemon spp.</i>	1	1
PHCI	<i>Phacelia cicutaria</i>	2	2
PRIL	<i>Prunus ilicifolia</i>	8	8
MALA	<i>Malosma laurina</i>	3	3
SAME	<i>Salvia mellifera</i>	3	3
TODI	<i>Toxicodendron diversilobum</i>	1	1
UMCA	<i>Umbellularia californica</i>	1	1
	Totals	66	73 (90.4%)

relations between patch area and each environmental variance variable to determine whether there is bias in intra-patch environmental heterogeneity due to variable patch size. Correlations between patch area and species richness (the number of woody species present within a patch) are used to test whether we have avoided sampling in the range of patch sizes over which species richness and area are fundamentally related (Kershaw 1974; Greig-Smith 1983).

We also evaluate if species composition is related to intra-patch environmental variance using correlations between species richness and each variance variable (TMINvar, TMAXvar, PSI_wvar, PSI_svar, TMIvar, SLPvar, CURVvar) and between the percent cover of ten common species and each variance variable. Samples where species are absent are omitted to avoid the inclusion of zero values in the correlation tests. The ten species (Table 5) are chosen in order to represent a range of habitats in the study area and three life history strategies related to seedling recruitment processes (Keeley 1998).

Correlations are compared among the environmental variables at the site and patch levels to determine whether the processes controlling environmental con-

ditions are stable across the two sampling scales and to determine whether patches are defined such that the patch-level environmental variables reflect topographically mediated processes (e.g., redistribution of soil moisture between ridges, hillslopes, and drainages).

Regression tree models are developed to evaluate scale dependence in functional relationships between species cover and environmental conditions at the site vs. the patch level. The site ($n = 87$) and patch level ($n = 87$) models predict the relative percent cover of ten species based on the environmental variables (TMAX, TMIN, PSI_w, PSI_s, TMI, and ROCK). Scale dependence is addressed by evaluating the consistency of modeled relationships between the two sample levels. The performance of the regression tree models is also evaluated in order to illuminate whether the patch-level data are suitable in scale for modeling relationships between species composition and environmental conditions driven by landscape-scale processes. Model performance is based on correlations between actual species cover and tree-predicted values.

Regression trees recursively partition a dependent variable into increasingly homogeneous subsets based

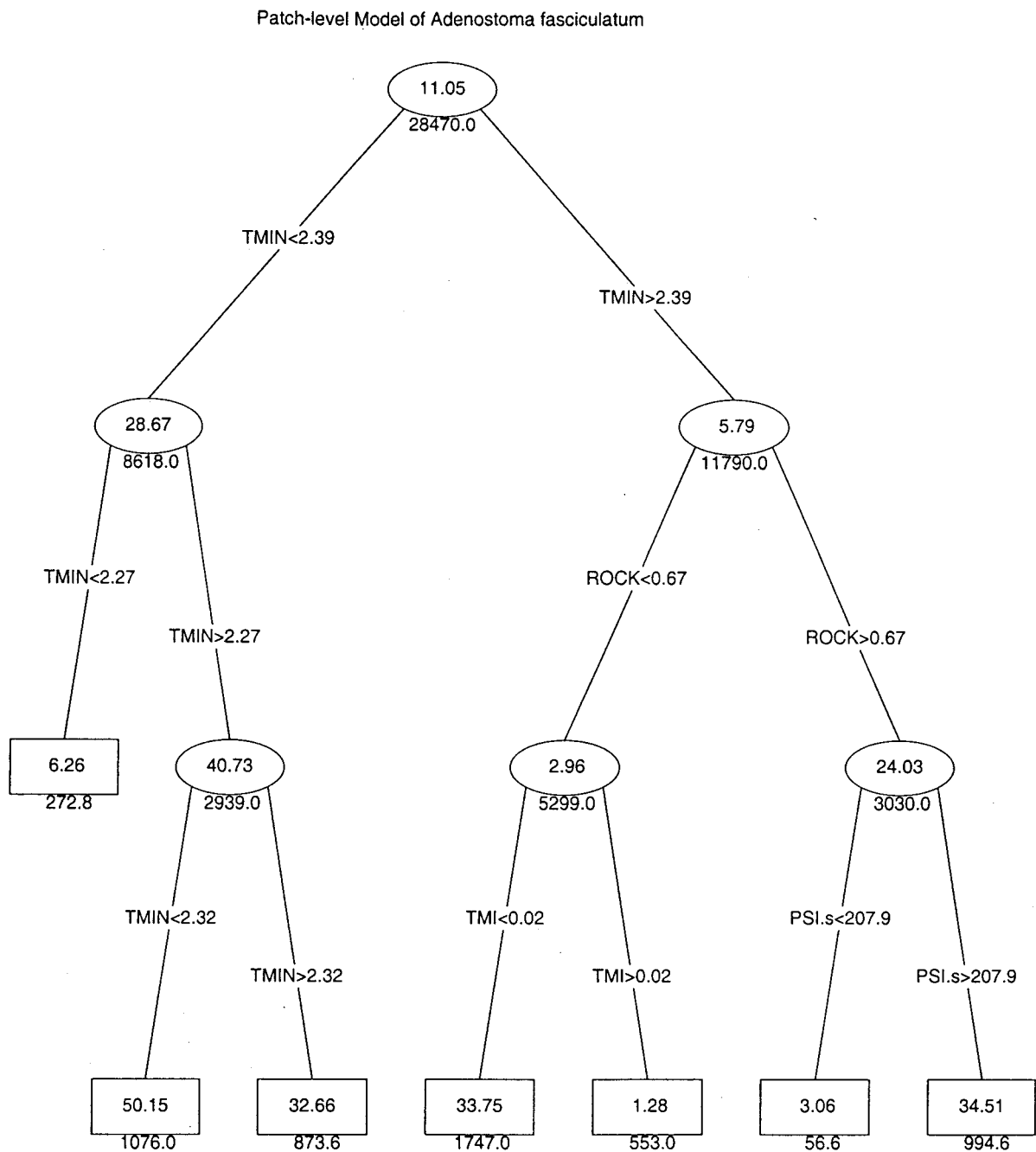


Figure 2. Patch-level regression tree for *Adenostoma fasciculatum* percent cover based on environmental variables (TMAX, TMIN, PSI_w, PSI_s, TMI, and ROCK). Abbreviations used for variables are given in the Derivation of environmental variables section. Ovals and squares represent non-terminal and terminal nodes, respectively. The values inside the ovals or squares are the predictions (means) of species percent cover. The values beneath the predictions are the RMS-errors associated with using the mean as the estimate for all samples which flow through that node. Values along the internode connections are critical thresholds of given variables that provide the basis for subsequent split and calculation of the predictions and RMS-errors.

Table 3. Analysis of variance results. Floristic and environmental variability *within* patches is significantly lower than variability *between* patches ($P > 0.001$). For species richness, between and within patch calculations have 252 and 86 degrees of freedom, respectively. For the seven environmental variables, between and within patch calculations have 86 and 13,599 degrees of freedom, respectively. The 'MS' column indicates the sum of squares divided by their corresponding degrees of freedom. The between and within mean squares form the numerator and denominator for the F test.

Variable	Source	MS	F
RICHNESS	Between	20.54	9.04
	Within	2.27	
TMAX	Between	326.96	885.46
	Within	0.37	
TMIN	Between	68.56	3198.47
	Within	0.02	
PSI _W	Between	952831.30	836.57
	Within	1138.97	
PSI _S	Between	203268.60	145.20
	Within	1399.91	
TMI	Between	67.13	37.98
	Within	1.77	
SLP	Between	7971.91	116.42
	Within	68.47	
CURV	Between	781.51	5.39
	Within	145.02	

on critical thresholds in continuous or categorical predictor variables. Tree-based models are graphically structured so that one can follow the tree node (*root*), through a series of binary splits on the predictor variables (*branches*), to an end node (*leaf*) (e.g., see Figure 2). The estimate for all observations that follow the same lineage of branchings from root to a given leaf is given by the mean x -value for that set of observations. Tree-based models are well suited for use in vegetation analysis where non-linear, hierarchical, or interactive relationships may exist (Clark and Pregibon 1993, Michaelsen et al. 1994). Other techniques such as generalized linear models (Hastie and Pregibon 1993) and generalized additive models (Hastie 1993), are also commonly used to predict vegetation patterns. (Austin and Meyers 1995; Austin et al. 1995; Franklin 1998). Franklin (1995) reviews literature in

which these and other approaches have been used for predictive vegetation modeling.

Trees can be pruned by removing splits that contribute little to the predictive power of the model or that overfit the data. To ensure consistency for model comparison and to prevent overfitting, we restrict the growth of each regression tree to the expansion point where the node deviance (a measure of node heterogeneity) explains less than 5% of the root node deviance, or where there are less than six observations at that node.

Results and discussion

Analysis of variance and correlation tests

Analysis of variance tests show that variability of species richness and environmental conditions is significantly lower within patches than between patches (Table 3). This indicates that the floristic criteria used to identify patches maintains homogeneity in intra-patch floristic and environmental variability, but captures considerable variability across the entire watershed. Correlation tests of patch area versus the variance variables are insignificant for all variables except TMINvar ($r = 0.74$; $P < 0.01$). This indicates that – except for minimum temperature – environmental heterogeneity is unbiased with respect to patch area within the range of patch sizes sampled. The relationship between TMINvar and patch area exists because the derivation of TMIN is based on a lapse rate that follows elevation. In this environment, the range of elevation values is almost necessarily greater within large patches than small patches. In contrast, intra-patch variability in solar radiation may be low for a large patch if it encompasses a uniform slope aspect. TMAX, although also governed by a lapse rate, has low intra-patch variability because it is also adjusted by slope aspect.

The relationship between patch area and species richness is insignificant ($r = 0.01$). This result indicates that we have sampled in the range of patch sizes over which species richness and area are independent in this environment. There are also no relationships between the environmental variance variables and species richness, or between the variance variables and species cover, except *Prunus ilicifolia* vs. CURVvar ($r = 0.41$; $P < 0.01$). The results of these three correlation analyses suggest that patch-level species composition is independent of intra-patch

Table 4. Correlations between environmental variables on a patch-level and site-level basis.

Variable	TMAX	TMIN	PSI _W	PSI _S	TMI	SLP	CURV
PATCH-LEVEL							
TMAX	1.00	0.72*	0.85*	0.21	0.08	0.11	-0.01
TMIN		1.00	0.33	0.04	0.17	0.01	-0.05
PSI_W			1.00	0.16	0.02	0.13	0.04
PSI_S				1.00	0.48*	-0.80*	-0.13
TMI					1.00	-0.64*	-0.74*
SLP						1.00	0.19
CURV							1.00
SITE-LEVEL							
TMAX	1.00	0.63*	0.83*	0.32	0.12	-0.04	0.12
TMIN		1.00	0.28	0.04	0.15	0.03	-0.12
PSI_W			1.00	0.26	0.09	-0.05	0.08
PSI_S				1.00	0.31	-0.87*	0.11
TMI					1.00	-0.44*	-0.44*
SLP						1.00	0.01
CURV							1.00

*P<0.01.

environmental heterogeneity given the criteria used to map patches.

Finally, comparison of correlations among the eight environmental variables on a per-cell basis and on a patch-level basis indicate that environmental relationships are consistent across sample levels (Table 4). This is evident, for example, for TMI, which is significantly correlated with CURV at both levels. Although CURV is not explicitly used to calculate TMI, the TMI/CURV relationship is logical; the topographic moisture index increases as a function of increasing topographic concavity. Stability of these relationships suggests that the processes controlling environmental conditions are stable across the two sampling scales. In this environment, these processes are primarily topographically mediated (e.g., the distribution of solar irradiance and soil moisture). Low or insignificant correlations at the patch level would indicate that patch boundaries are distributed such that topographically driven patterns in environmental conditions are obscured. This would be the case, for example, with randomly distributed sampling units that overlap pronounced topographic transitions. Combined with the earlier conclusion that species composition is independent of intra-patch environmental variability, this suggests that patch-level data are suitable for modeling

relationships between species composition and topographically mediated processes at landscape scales.

Regression tree analyses

Regression tree analyses further examine scale-dependent features in the data and whether the data are suitable for modeling ecological relationships between species composition and topographically mediated environmental conditions. Functional relationships expressed in the regression tree models of species cover are generally consistent between the patch-level and site-level models. For example, both the site-level model (Figure 2) and the patch-level model (Figure 3) for *Adenostoma fasciculatum* indicate that cover increases on dry, exposed, rocky sites with moderately cold minimum temperatures. Although the two models of *A. fasciculatum* cover are not identical, the architecture of the two trees is similar. In both models, the direction of the effect of each variable on cover is the same at all levels. For example, the first split occurs on minimum temperature (TMIN) in both regression trees. Where TMIN is high, both the site- and patch-level models split on percent bedrock (ROCK), predicting greater cover on rocky slopes. Where ROCK is high, both models predict greater cover at sites with high summer insolation (PSI_S). Where ROCK is

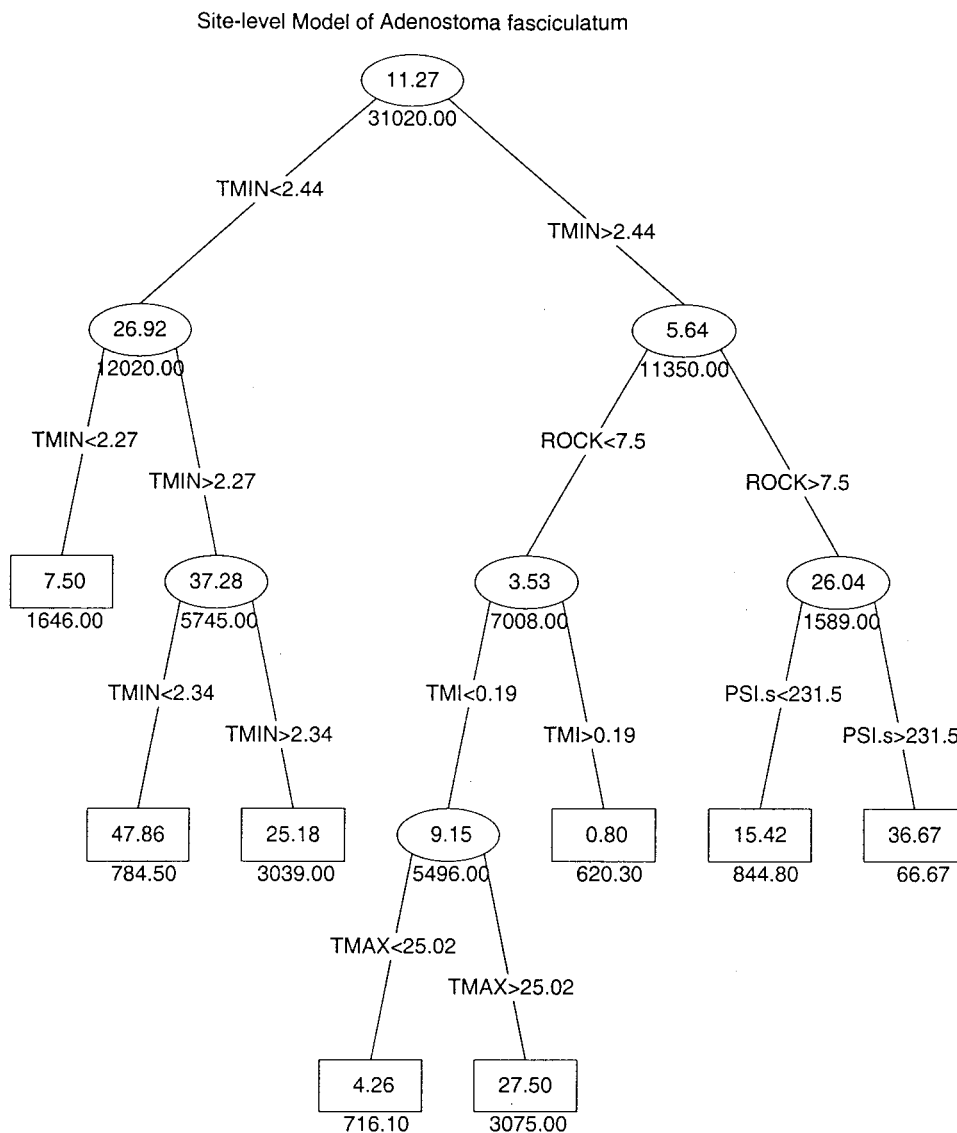


Figure 3. Site-level regression tree for *Adenostoma fasciculatum* percent cover based on environmental variables. See caption of Figure 2 for explanation of this figure.

low, both models are still consistent, predicting greater *A. fasciculatum* cover on dry sites characterized by low TMI. However, where TMI is low, the site-level model departs from the patch-level model, predicting an increase in cover on slopes with high summer temperatures (TMAX). In the left lineage of branches where TMIN is lower, both models indicate similar responses to topographic variations in minimum air temperature.

A similar evaluation can be made for *Arctostaphylos glandulosa*, for which both models indicate that cover increases on rocky, xeric sites with relatively

cold winter temperatures (trees not shown). In both models the first split occurs on percent bedrock (ROCK). Where bedrock is high, another split occurs on TMI in the patch-level model, but splits on PSI_S in the site-level model. However, both splits predict greater *A. glandulosa* cover on more xeric sites. Where bedrock is low, both models split on TMIN followed by splits on TMI and PSI.

The overall consistency in the nature of the species responses to environmental gradients is greater than expected given the difference in area that each sampling level encompasses (e.g., 100 m² cell of a site vs.

the mean patch area of 1.2 ha). This consistency suggests that – at scales between the site level and patch level – there are minimal scale dependencies in the environmental processes underlying species composition patterns in this system.

An implicit assumption in the above models is that the GIS-derived variables characterize environmental variability that is driven by landscape-scale processes, such as topographically regulated fluxes of energy and moisture. Furthermore, literature shows that chaparral shrub species are distributed in response to such processes (e.g., Franklin 1998; Miller et al. 1983). Good model performance at the patch level (Table 5), therefore, suggests that we have sampled patch-level species composition at a scale that is functionally related to landscape-scale processes, rather than finer or broader scale processes.

Although there are no obvious scale-dependent features in the data, the patch-level models perform slightly better than the site-level models. Correlations between actual species cover and tree-predicted cover are higher for every species except *Heteromeles arbutifolia* (Table 5). Aggregation of site-level data to the patch level may improve model performance by averaging out fine-scale floristic variability that develops in response to intra-patch biotic processes (e.g., seed dispersal) or smaller-scale physical processes (e.g. micro-topographic flows of soil moisture) that are important at sub landscape-scales. Averaging site-level data to the patch level may also reduce the impact of site-level data error. Multiple sampling and aggregation of vegetation samples to the patch level may thus improve the reliability of the species composition data, and in turn, improve model performance.

The reproductive features of *Heteromeles arbutifolia* may explain why the site-level model performs better than the patch-level model for this species. Bullock (1978) showed that chaparral species with fleshy fruits (such as *H. arbutifolia*) are often sparsely distributed in space. Bullock (1978) views them as interstitial species because they occur in small openings within a matrix of other species, random locations where frugivorous animals (e.g. birds) deposit seeds. In contrast to the other species, *H. arbutifolia* recruitment patterns may track environmental processes that vary at higher spatial frequencies, and hence, the site-level model performs better. This is an example of how comparison between models at these two scales can be used to make inferences regarding the processes governing spatial patterns of plant species.

Other ecological processes can also be inferred from the modeled patterns of species cover. The models of *Quercus agrifolia* and *Umbellularia californica*, for example, indicate that these species favor mesic conditions. The success of both species increases with increasing topographic moisture index and decreasing solar radiation loads. These results are consistent with literature indicating that resprouting species require moist soil conditions for successful seed germination and prevention of drought stress (Keeley 1991). Typically excluded from drier sites, these species have poor stomatal control and productivity is limited at much lower water stress levels (Davis et al. 1998; Keeley 1986). The model of *Ceanothus megacarpus*, in contrast, indicates greater success on xeric, well-illuminated slopes. Consistent with processes reported in the literature, obligate seeders – such as *C. megacarpus* – resist water stress on xeric sites by continuing carbon uptake during drought, despite very low leaf water potentials (Davis et al. 1998).

Methodological issues

The sampling approach represents a balance between sampling efficiency and the accuracy of field data. In this section, we discuss several concerns relating to data accuracy. Oblique views can cause geometric distortion in terrestrial monoscopic sensing in similar ways that it occurs in satellite and airborne imagery (e.g., Pickup et al. 1995). In addition, oblique viewing results in the masking of plants by other plants in the FOV. Consequently, there will be some degree of error in site-level canopy cover estimates due to variations in view angle. Sampling must also avoid view directions that are opposite to the solar zenith and azimuth in order to avoid excessive glare. Although species identification accuracy is high (Table 2), the accuracy of percent canopy cover estimation has not been tested. Attempts to compare remote and *in situ* estimates of percent cover have been impeded due to the difficulty of *in situ* estimation of species-specific percent cover within the large stature, closed-canopy.

Problems associated with canopy cover estimation, masked plants, species misidentification, and missing rare species are offset by using acute viewing angles and obtaining multiple samples at each nested level. At the patch-level, multiple site-level samples within each patch minimize the risk of missing species. At the site-level, clusters of FOVs around a randomly located point help ensure that site-level samples are characterized as thoroughly and accurately as possible. Failure

Table 5. Model performance of 20 regression-tree models relating species canopy cover (%) to 6 environmental variables ($P < 0.001$ for all models). Models are developed at the patch level and site level. Each species' seedling recruitment mode is listed. Each tree model is restricted at the expansion point where the node deviance explains less than 5% of the root node deviance and if there are at least six observations at that node.

Species	Correlation coefficient		Seedling recruitment mode
	Patch-level	Site-level	
<i>Artostaphylos glandulosa</i>	0.91	0.86	Facultative resprouter
<i>Adenostoma fasciculatum</i>	0.90	0.84	Facultative resprouter
<i>Cercocarpus betuloides</i>	0.83	0.66	Obligate resprouter
<i>Ceanothus crassifolius</i>	0.87	0.85	Obligate seeder
<i>C. megacarpus</i>	0.86	0.82	Obligate seeder
<i>C. spinosus</i>	0.84	0.83	Facultative resprouter
<i>Heteromeles arbutifolia</i>	0.77	0.88	Obligate resprouter
<i>Quercus agrifolia</i>	0.92	0.80	Obligate resprouter
<i>Prunus ilicifolia</i>	0.87	0.86	Obligate resprouter
<i>Ubellularia californica</i>	0.82	0.77	Obligate resprouter

to see understory plants poses an additional problem for remote sampling of dense vegetation canopies. Consequently, samples may need to be limited to overstory plant composition in these settings. In chaparral, the closed, single-layer canopy typically precludes understory plants (Christensen and Muller 1975). Where bedrock outcrops and/or inhospitable substrate conditions prevent canopy closure, subshrubs (e.g., *salvia spp.*) often take advantage of the increased illumination in canopy gaps. In such cases, it is possible to see into gaps and identify lower stature species.

The criteria for mapping vegetation patches are designed as a simple method to maintain consistent patch identification. However, because they are not applied quantitatively, error will exist in the delineation of boundaries. For example, it is possible to overlook visually inconspicuous floristic transitions, or to incorrectly estimate transitions. Non-homogeneous patches may impact modeled relationships between species composition and environmental conditions. However, the analyses of intra-patch variability indicate that internal floristic and environmental variability is low and stable across the entire sample of vegetation patches.

While the proposed approach can provide rapid samples of species composition, other methods may also be appropriate for certain investigations. Airborne videography, for example, can provide numerous, high resolution samples of surface characteristics (Mausel et al. 1992). It is relatively inexpensive (Pickup et al. 1995) and has been applied to resolve vegetative features at less than 1 meter pixel resolution (e.g., Pickup

et al. 1995; Biging et al. 1995). This method is useful in analyses that require data for training and validation of canopy attributes (e.g., Biging et al. 1995), and can use phenological change to enhance remote detection. However, such data may still be too generalized to identify individual plant species or estimate species abundance. Biging et al. (1995) evaluated videography for identifying tree species and crown sizes in northern California. While individual tree crown sizes were estimated with acceptable accuracy, identification of individual tree species was not feasible in that environment. Investigations that require accurate estimates of canopy closure and plant composition (e.g. wildlife habitat suitability) could perhaps benefit by integrating our sampling method with other high resolution approaches, such as videography, to provide effective, yet efficient landscape assessments.

Conclusion

GIS-based studies of relationships between species composition and environmental conditions require data that are of appropriate resolution and distribution given the scale at which underlying ecological processes operate. We have outlined a methodology for sampling plant species composition in environments with poor accessibility but high visibility. The method provides species-level data at two scales, and allows sample sizes that are large enough to outweigh noise associated with data error and stochastic vari-

ability. Our results suggest that the approach achieves a suitable balance between the cost of data collection and data accuracy. Other researchers have tried to strike this balance by sacrificing sample size and representative distributions in order to achieve detailed floristic data at the scale of small plots. Others have used remote sensing to collect vegetation data, thereby sacrificing floristic detail for sample size and density. Our approach strikes an intermediate balance between these two extremes.

Evaluation of the methodology permits several conclusions. Remote identification of overstorey species is feasible with an acceptable level of accuracy in environments – such as chaparral – with open visibility of surrounding terrain. Landscapes with woodland and shrubland types, semi-desert scrub, desert vegetation, and other Mediterranean-type vegetation formations, would also allow open visibility due to sparse and/or low stature canopies and rugged terrain. Our rules and scales used for sampling (e.g., MMU) can be adjusted to match the processes underlying vegetation patterns in other ecosystems.

In this application, we have defined patch boundaries such that spatial patterns in topographically regulated processes are preserved. In addition, there is minimal bias associated with scaling the data from the site level to the patch level. Patches are defined such that internal floristic and environmental variability is low and intra-patch variability is stable across the sample of patches. As a result, species composition is independent of patch area and intra-patch heterogeneity in environmental conditions.

Regression tree models of relationships between species cover and environmental conditions are generally consistent across the two sampling scales. Therefore, at scales between the site level and patch level, there is a minimal degree of scale dependence in the ecological processes that govern species composition patterns in this environment. Good performance of the regression tree models of species cover suggest that the data are characterized at an appropriate scale to assess ecological processes underlying species composition patterns at landscape scales. The patch-level models perform somewhat better than the site-level models. Aggregation of the site-level data to the patch level probably reduces the influence of stochasticity and processes that vary at a finer scale than the spatially distributed environmental variables. Given the slightly better model performance and the fact that we are more confident in the reliability of patch-level species composition (due to repeat sampling and

data aggregation), the patch-level data seems to be more appropriate than the site-level data for assessing vegetation patterns at the scale of interest. However, comparison between models at the two scales can be useful for assessing the processes and associated scales of variability governing spatial patterns of plant species.

Decisions regarding trade-offs in sampling design may impact research results. In response to challenges associated with floristic sampling in inaccessible environments, careful consideration has been given to the scale of investigation, level of taxonomic detail, the methods used to sample taxa, and the spatial distribution of vegetation samples. We trade a degree of accuracy for sufficient degrees of freedom and thorough representation of the range of floristic variability. Not only does the remote approach provide numerous observations, but it also provides vegetation composition data at the species level. Neither remote sensing nor field plots do both of these things well.

Acknowledgements

The authors thank three anonymous reviewers for valuable comments on this manuscript. We gratefully acknowledge the logistic and field support of Don Johnson, Sheila Johnson, Carl Moody, Lisa Meentemeyer, Jessica Dora, Sue Hegner, and Tom Moody. Our work benefited from generous input by Larry Band, Mark Borchert, Janet Franklin, and Joel Michaelsen. We also thank J. Michaelsen (Professor, Univ. of California at Santa Barbara) for providing climate data and Max Moritz (Professor, California Polytechnic Institute) for providing fire history data. This research is partially supported under NASA grant NAG5-6583.

References

- Austin, M.P. and Meyers, J.A. 1995. Modelling of landscape patterns and processes using biological data, Subproject 4: Real data case study. Consultancy report to ERIN, CSIRO, Division of Wildlife and Ecology, Canberra.
- Austin, M.P., Meyers, J.A., Belbin, L. and Doherty, M.D. 1995. Modelling of landscape patterns and processes using biological data: Sub-project 5, Simulated data case study. Division of Wildlife and Ecology, Commonwealth Scientific and Industrial Research Organization, Canberra.
- Beven, K.J. and Kirkby, M.J. 1979. A physically based, variable contributing area model of basin hydrology. *Hydrol Sci Bull* 24: 43–69.

- Biging, G.S., Dobbertin, M. and Murphy, E.C. 1995. A test of airborne multispectral videography for assessing the accuracy of wildlife habitat maps. *Can J Remote Sensing* 21: 357–366.
- Bolstad, P.V., Swank, W. and Vose, J. 1998. Predicting Southern Appalachian overstory vegetation with digital terrain data. *Landscape Ecol* 13: 271–283.
- Bullock, S.H. 1978. Plant Abundance and distribution in relation to types of seed dispersal in chaparral. *Madroño* 25: 104–105.
- Christensen, N.L. and Muller, C.H. 1975. Effects of fire on factors controlling plant growth in *Adenostoma* chaparral. *Ecol Monogr* 45: 29–55.
- Clark, L.A. and Pregibon, D. 1993. Tree-based models. *In* *Statistical Models*. Edited by Chambers, J.M. and Hastie, T.J. S. Chapman & Hall, Inc., London.
- Davis, F.W. and Goetz, S. 1990. Modeling vegetation pattern using digital terrain data. *Landscape Ecol* 4: 69–80.
- Davis, S.D., Kolb, K.J. and Barton, K.P. 1998. Ecophysiological processes and demographic patterns in the structuring of California chaparral. *In*: *Landscape Disturbance and Biodiversity in Mediterranean-Type Ecosystems*. Edited by Rundel, P.W., Montenegro, G. and Jaksic, F.M.
- Dibblee, T.W. 1986. Geologic Map of the Santa Barbara Quadrangle. Thomas Dibblee, Jr. Geological Foundation.
- Dubayah, R. 1992. Estimating net solar radiation using Landsat Thematic Mapper and digital elevation data. *Water Res Res* 28: 2469–2484.
- Franklin, J. 1995. Predictive vegetation mapping: geographic modeling of biospatial patterns in relation to environmental gradients. *Prog Phys Geogr* 19: 474–499.
- Franklin, J. 1998. Predicting the distributions of shrub species in California chaparral and coastal sage communities from climate and terrain-derived variables. *J Veg Sci*.
- Franklin, J., McCullough, P. and Gray, C. In press. Terrain variables used for predictive mapping of vegetation communities in Southern California. *In* *Terrain Analysis*. Edited by Wilson, J.P. and Gallant, J.C. GeoInformation International, Cambridge.
- Hastie, T.J. 1993. Generalized Additive Models. *In* *Statistical Models*. Edited by Chambers, J.M. and Hastie, T.J. S. Chapman & Hall, Inc., London.
- Hastie, T.J. and Pregibon, D. 1993. Generalized Linear Models. *In* *Statistical Models*. Edited by Chambers, J.M. and Hastie, T.J. S. Chapman & Hall, Inc., London.
- Greig-Smith, P. 1983. *Quantitative Plant Ecology*, third edition. University of California Press, Berkeley, California.
- Keeley, J.E. 1986. Resilience of mediterranean shrub communities to fires. *In* *Resilience in Mediterranean-type Ecosystems*. Edited by Dell, B., Hopkins, A.J.M., and Lamont, B.B.
- Keeley, J.E. 1991. Seed germination and life history syndromes in the California chaparral. *Bot Rev* 57: 81–116.
- Keeley, J.E. 1998. Coupling Demography, Physiology and Evolution in Chaparral Shrubs. *In* *Landscape Disturbance and Biodiversity in Mediterranean-Type Ecosystems*. Edited by Rundel, P.W., Montenegro, G. and Jaksic, F.M.
- Kent, M. and Coker, D. 1992. *Vegetation Description and Analysis: a Practical Approach*. Belhaven Press, London.
- Kershaw, K.A. 1975. *Quantitative and Dynamic Plant Ecology*, second edition. Elsevier, New York, NY.
- Mausel, P.W., Everitt, J.H., Escobar, D.E. and King, D.J. 1992. Airborne videography current status and future perspectives. *Photogr Eng Remote Sensing* 58: 1189–1195.
- Michaelsen, J., Schimel, D., Friedl, M., Davis, F.W. and Dubayah, R.C. 1994. Regression tree analysis of satellite and terrain data to guide vegetation sampling and surveys. *J Veg Sci* 5: 673–686.
- Miller, P.C., Hajek, E. and Miller, P.M. 1983. The influence of annual precipitation, topography and vegetation cover on soil moisture and summer drought in southern California. *Oecologia* 56: 385–391.
- Pickup, G., Chewings, V.H. and Pearce, G. 1995. Procedures for correcting high resolution airborne video imagery. *Int J Remote Sensing* 16: 1647–1662.
- Running, S.W., Nemani, R.R. and Hungerford, R.D. 1987. Extrapolation of synoptic meteorological data in mountainous terrain and its use for simulating forest evapotranspiration and photosynthesis. *Can J Forest Res* 17: 472–483.
- Stohlgren, T.J., Coughenour, M.B., Chong, G.W., Binkley, D., Kalkhan, M.A., Schell, L.D., Buckley, D.J., and Berry, J.K. 1997. Landscape analysis of plant diversity. *Landscape Ecol* 12: 155–170.
- Wells, P.V. 1962. Vegetation in relation to geological substratum and fire in the San Luis Obispo Quadrangle, California. *Ecol Monogr* 32: 79–103.
- Xu, X., Aiken, C.L. and Nielsen, K.C. 1999. Real time and the virtual outcrop improve geological field mapping. *EOS, Transactions, Am Geophys Union* 80 (20): 317–324.