

Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Remote sensing of structural complexity indices for habitat and species distribution modeling

L.D. Estes^{a,b,c,*}, P.R. Reillo^a, A.G. Mwangi^d, G.S. Okin^e, H.H. Shugart^b^a Rare Species Conservatory Foundation, P.O. Box 1371, Loxahatchee, FL 33470, USA^b Center for Regional Environmental Studies, Department of Environmental Science, University of Virginia, 291 McCormick Road Charlottesville, VA 22903, USA^c Woodrow Wilson School, Robertson Hall, Princeton University, Princeton, NJ 08544, USA^d Rhino Ark Charitable Trust, P.O. Box 181, Uhuru Gardens 00517, Kenya^e Department of Geography, University of California Los Angeles, Box 951524, Los Angeles, CA 90095, USA

ARTICLE INFO

Article history:

Received 5 November 2008

Received in revised form 18 November 2009

Accepted 27 November 2009

Keywords:

ASTER
 SRTM digital elevation model
 Spectral mixture analysis
 Topographic normalization
 Structural complexity index
 Distribution model
 Rare species
 Montane forest
 Microhabitat

ABSTRACT

Spatial distribution models are increasingly used in ecological studies, but are limited by the poor accuracy of remote sensing (RS) for mapping microhabitat (<0.1 ha) features. Mapping accuracy can be improved by combining advanced RS image-processing techniques with microhabitat data expressed as a structural complexity index (SCI). To test this idea, we used principal components analysis (PCA) and an additive SCI method developed for forest ecology (calculated by re-scaling and summing representative structural variables) to summarize 13 microhabitat-scale (0.04 ha) vegetation structure attributes describing the rare mountain bongo antelope's (*Tragelaphus eurycerus isaaci*) habitat in Kenya's Aberdare mountains. Microhabitat data were collected in 127 plots: 37 related to bongo habitat use, 90 from 1 km-spaced grid points representing overall habitat availability and bongo non-presence. We then assessed each SCI's effectiveness for discerning microhabitat variability and bongo habitat selection, using Wilcoxon Rank Sum tests for differences in mean SCI scores among plots divided into 4 vegetation classes, and the Area Under the Curve (AUC) of receiver operating characteristics from logistic regressions. We also examined the accuracy of predicted SCI scores resulting from regression models based on variables derived from a) ASTER imagery processed with spectral mixture and texture analysis, b) an SRTM DEM and c) rainfall data, using the 90 grid plots for model training and the bongo plots as an independent test dataset. Of the five SCIs derived, two performed best: the PCA-derived Canopy Structure Index (CSI) and an additive index summarizing 8 structural variables (AI_8). CSI and AI_8 showed significant differences between 5 of 6 vegetation class pairs, strong abilities to distinguish bongo-selected from available habitat (AUCs = 0.71 (CSI); 0.70 (AI_8)), and predicted scores 60–110% more accurate than reported by other studies using RS to quantify individual microhabitat structural attributes (CSI model $R^2 = 0.51$, RMSE = 0.19 (training) and 0.21 (test); AI_8 model $R^2 = 0.46$, RMSE = 0.17 (training) and 0.19 (test)). Repeating the Wilcoxon tests and logistic regressions with RS-predicted SCI values showed that AI_8 most effectively preserved the patterns found with the observed SCIs. These results demonstrate that SCIs effectively characterize microhabitat structure and selection, and boost microhabitat mapping accuracy when combined with enhanced RS image-processing techniques. This approach can improve distribution models and broaden their applicability, makes RS more relevant to applied ecology, and shows that processing field data to be more compatible with RS can improve RS-based habitat mapping accuracy.

© 2009 Elsevier Inc. All rights reserved.

1. Introduction

Swift conservation action is needed to halt accelerating biodiversity losses caused by climate change and direct human impacts. Effective conservation requires a sound understanding of target species' habitat relationships and distributions, which are increasingly studied using

quantitative distribution models (Guisan & Zimmermann, 2000; Rushton et al., 2004). To provide accurate ecological insight, distribution models must represent the habitat features that define a species' niche, and characterize these at appropriate spatio-temporal scales (Vaughan & Ormerod, 2003). Remote sensing (hereafter RS) is the most efficient way to collect multi-scaled habitat data, and is thus increasingly used to provide distribution model input (Kerr & Ostrovsky, 2003; Turner et al., 2003; Rushton et al., 2004).

RS is useful—often indispensable—for providing habitat data at larger patch- and landscape-scales that are difficult to cover in field surveys (Estes et al., 2008). However, ecological research is often

* Corresponding author. Center for Regional Environmental Studies, Department of Environmental Science, University of Virginia, 291 McCormick Road Charlottesville, VA 22903, USA. Tel.: +1 202 431 0496.

E-mail address: ldc2c@virginia.edu (L.D. Estes).

focused on the microhabitat scales (here defined as ≤ 0.1 ha) where fundamental habitat selection decisions are made (Turner et al., 2003; Hilden, 1965). Unfortunately, RS tends to be less effective at such fine scales, where it is more prone to errors related to terrain shadowing, geo-locational discrepancies, and other factors. The imprecision reported by the few ecological modeling studies that have used RS to quantify microhabitat features (e.g. Stickler & Southworth, 2008; Estes et al., 2008; Jeganathan et al., 2004) appears to vindicate the skepticism most ecologists have for RS-based microhabitat assessment (Turner et al., 2003). This perception means that ecologists generally choose one of two modeling options: 1) use a non-spatial model that includes microhabitat-scaled predictors collected from field plots (e.g. Estes et al., 2008); 2) use a spatial distribution model based entirely on coarser-scaled GIS- or RS-derived predictors (e.g. Osborne et al., 2001; Schadt et al., 2002).

The potential for distribution models to provide ecological insight therefore appears to be limited by the ability of RS to characterize microhabitat variables. Improved RS technology and image analysis techniques may mitigate this limitation, as can expressing microhabitat features in a manner that is more compatible with RS data. The RS literature contains numerous techniques for improving remote measurements of vegetation structure (e.g. Hyde et al., 2006; Hansen et al., 2001; Wulder et al., 1998), but these are primarily tested at stand-scales (≥ 1 ha) and are rarely applied to solve specific ecological modeling problems. Ecologists, who are generally unfamiliar with advanced RS, tend to use simple spectral indices (e.g. NDVI) or classifications to generate model predictors, which may be unrelated to or fail to capture the range of variation in important habitat factors (St-Louis et al., 2006).

Microhabitat data may be more compatible with RS when expressed as a structural complexity index (SCI), a composite variable widely used in forest ecology to summarize several vegetation structural measures (e.g. McElhinny et al., 2005; Neumann & Starlinger, 2001; Pommerening, 2002). A few RS studies, conducted at forest stand-level scales (≥ 5 ha), suggest that SCIs may correlate better with RS data than individual structural measures (Cohen & Spies, 1992; Hansen et al., 2001; Danson & Curran, 1993). Despite this potential, SCIs have been used in just a handful of RS-based ecological assessments (Coops & Catling, 1997; Coops & Catling, 2002; Gibson et al., 2004), and have not been tested at scales finer than 0.25 ha.

Our objective is to make RS more relevant for applied ecology and conservation by improving the accuracy of RS-based microhabitat mapping, thereby allowing distribution models to be more effectively used. We pursue this objective by using SCIs in conjunction with several effective RS techniques seldom used in ecological studies to address a need identified by Estes et al. (2008)—to quantitatively map microhabitat-scale (0.04 ha) vegetation structural features important to a rare antelope, the mountain bongo *Tragelaphus eurycerus isaaci*. The mountain bongo is the subject of an active ongoing conservation project that depends on distribution modeling to target appropriate, ecologically informed management actions. We assess our approach's ability to address this need by answering three questions:

1. Are SCIs effective for discerning microhabitat variability?
2. Are SCIs effective for distinguishing between selected and available microhabitats?
3. Do SCIs combined with advanced RS analysis techniques boost microhabitat mapping accuracy?

We use two different SCI construction techniques to summarize a suite of microhabitat-scale vegetation structure measures describing bongo habitat. We then predict and map SCI values with variables derived from inexpensive optical RS data using spectral mixture (SMA) and texture analysis (TA), together with several supplementary climate-related measures. SMA and TA effectively extract image information related to vegetation and habitat structure (Kayitakire et al., 2006; Lu et al., 2003; Sabol et al., 2002; St-Louis et al., 2006; Theau et al., 2005; Wulder et al., 1998). We answer the three questions regarding our

approach's effectiveness using univariate statistical tests, simple logistic regression, and the error resulting from RS-based predictions of SCI values.

2. Methods

2.1. Background

This study concerns Kenya's Aberdare Mountains (Fig. 1 and described by Schmitt (1991)), an equatorial, volcanic massif reaching 3998 m along the Great African Rift. The Aberdares' climate is characterized by alternating wet and dry seasons linked to the shifting Inter-Tropical Convergence Zone. Mean annual precipitation varies between ~ 950 and ~ 2250 mm, depending on exposure to prevailing winds. Temperature is primarily influenced by elevation and time of day, rather than season.

Above dense rural settlements reaching to 2000–2500 m, the Aberdares support a diversity of habitats and wildlife protected by the Kenyan government. Vegetation is primarily organized along elevation and precipitation gradients. Woody forests interspersed with disturbance-maintained shrublands predominate below 2500 m, characterized by the genera *Podocarpus*, *Juniperus*, *Cassipourea*, *Neoboutonia*, *Macaranga*, and *Olea*. An extensive bamboo *Yushania alpina* zone (2500–3000 m) lies above this, followed by *Hagenia abyssinica*–*Hypericum revolutum* forest (3000–3300 m), and finally ericoid heathlands and alpine grasslands (> 3300 m). Micro-topography and various disturbances create significant structural variations within these vegetation types at scales ranging from 100s to 1000s m².

The mountain bongo is an endangered antelope sub-species endemic to Kenyan montane forest (Kingdon, 1982; East, 1999). The approximately 100 remaining wild bongo are mostly found in the 1800 km² Aberdare forests. At the microhabitat scale (0.04 ha), bongo (a browsing antelope) select sites with abundant tall forbs and lianes, which are primarily found in edge habitats or mature woody forests (Estes et al., 2008). Mapping and identifying such sites are necessary for properly understanding bongo spatial ecology (Estes et al., 2008). RS is the only practical tool for vegetation structure mapping in this study area, but the rough terrain and fine spatial scale both undermine accuracy. This problem was indicated by earlier work in which individual microhabitat-scale vegetation structure measures important to bongo were found to have relatively weak pair-wise correlations with variables derived using spectral mixture and texture analysis from 5 to 30 m resolution optical satellite imagery (Estes et al., 2008).

2.2. Data

2.2.1. Field data

We collected data on vegetation structure and composition within 127 0.04 ha circular plots (22.6 m diameter) over three seasons in 2005 and 2006 (see Fig. 1). Information gathered included diameter at breast height (DBH) of trees > 7.5 cm diameter, canopy height, density of shrubs and forbs > 1.5 m, canopy and ground cover proportions, the herbaceous layer's average height, and species composition within each structural measure. Thirty-seven of these plots were located on sites that bongo had used (as identified by expert trackers), and 90 plots were collected from a grid of pre-located points with 1 km spacing that were created in a geographical information system (GIS). Bongo plots were found by tracking along a general, pre-determined heading while following paths of least resistance (the "reconnaissance" approach (Walsh & White, 1999)). While searching for bongo sign we also collected grid points, which were sampled when the tracking party passed within 200 m of a point, as determined by a global positioning system (GPS). Grid plots were converted to bongo plots if habitat use signs were found (this occurred in only three instances). To ensure that our sampling represented the predominant vegetation types, we

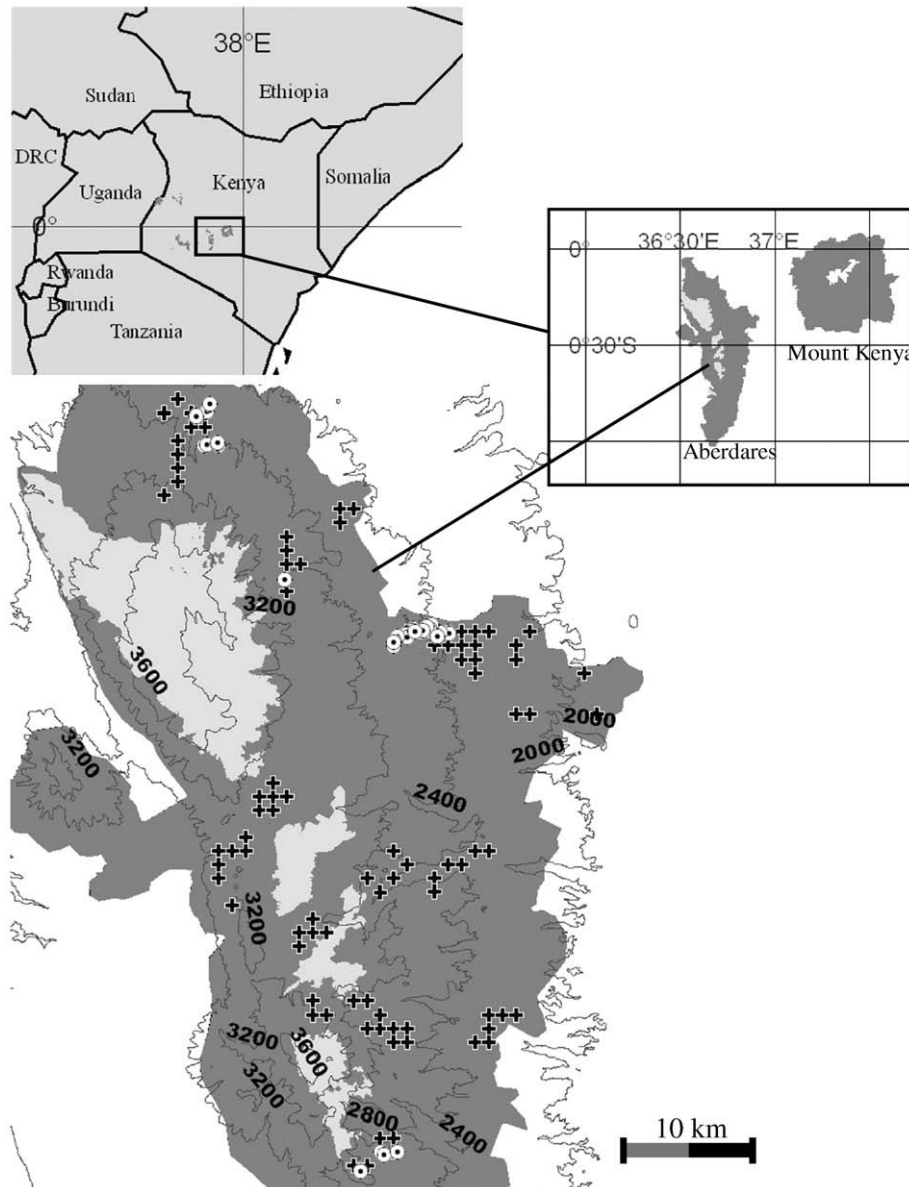


Fig. 1. The study area (the Aberdare Mountains) in relation to the region (top) and other montane forests in Kenya (dark grey patches in regional map). The close up view (main map) shows the location of the collected non-presence grid (black crosses with white outlines) and bongo presence plots (white circles with black dots). The dark grey area describes the habitats of interest, while the light grey indicates high elevation grassland and sparse shrubland that falls outside the scope of this study.

targeted tracking surveys to capture the study area's range in elevation and precipitation gradients. By combining grid point collection with reconnaissance surveys, we were able to search widely for bongo sign with minimal interruption, while obtaining data representing the microhabitat structural variation found within the study area. Since the grid plots we collected did not contain any bongo sign, they are henceforth referred to as “non-presence” plots. Sample plots were independent by virtue of their spatial (distance between non-presence plots) or temporal (interval between habitat use in neighboring bongo plots) separation. For additional description of field methods, see Estes et al. (2008).

We derived 13 structural and compositional variables from the collected data, described in Table 1 along with their basic statistics. Variables with highly kurtotic distributions were transformed for subsequent SCI construction.

2.2.1.1. SCIs. We used two techniques to create quantitative SCIs from the 13 microhabitat variables, principal components analysis (PCA) and a method developed by McElhinny et al. (2006) for forest biodiversity

assessments. We chose PCA because it was used in the few RS studies that have tested SCIs (Cohen & Spies, 1992; Hansen et al., 2001; Danson & Curran, 1993), and McElhinny et al.'s (2006) technique because it synthesizes—and thus represents—numerous SCI methods found in the forest ecology literature (McElhinny et al., 2006).

We applied PCA to the full set of 13 microhabitat variables, retaining individual principal components (PC) accounting for at least 15% of variance. We rotated the retained PCs to highlight the most influential microhabitat variables within each.

Using the same 13 variables, we followed three steps to construct SCIs using McElhinny et al.'s (2006) approach:

1. Identify the “core” variables, namely those a) with well distributed values (kurtosis < 2), b) that help distinguish between represented vegetation types, and c) that are not redundant, i.e. a variable's correlation with another should not exceed $|r| = 0.7$;
2. Standardize core variables to a 10-point scale by regressing their quartile midpoint values against the scalar quartile midpoint values;
3. Sum each plot's rescaled core variables to obtain its SCI score.

Table 1

The derived microhabitat structural variables and the basic statistics of their pre-transformed values, segregated according to bongo and non-presence grid data points. The associated transformation type and resulting kurtosis value is also presented. Canopy structure measures are presented in italic typeface, understorey measures in plain text.

Variable ^a	Units	Bongo (n = 37)				Grid (n = 90)				Transform	Kurtosis
		Mean	Std dev	Min	Max	Mean	Std dev	Min	Max		
<i>Basal area</i>	m ² ha ⁻¹	35.00	31.33	0.00	127.46	17.16	23.95	0.00	138.32	sqrt	-0.02
<i>Quadratic mean DBH</i>	cm	40.67	27.07	0.00	123.00	26.54	29.45	0.00	187.95	sqrt	-0.13
<i>Stem density</i>	n ha ⁻¹	251.98	233.46	0.00	772.78	223.52	277.98	0.00	1121.77	sqrt	-0.67
<i>Canopy height</i>	m	17.44	8.97	1.55	42.41	12.18	6.26	0.90	26.10	-	-0.45
<i>St dev canopy height</i>	m	4.21	3.95	0.15	14.47	2.25	2.31	0.06	12.78	-	0.79
<i>N tree species</i>	n	2.22	1.55	0.00	5.00	1.51	1.80	0.00	9.00	sqrt	1.62
<i>Canopy cover</i>	-	0.79	0.25	0.00	1.00	0.73	0.29	0.00	1.00	arcsine	0.28
<i>Bamboo density</i>	n ha ⁻¹	13,499.12	22,624.01	0.00	110,687.97	9172.73	15,464.03	0.00	72,932.51	sqrt	-0.09
<i>Shrub density</i>	n ha ⁻¹	1803.85	2172.30	0.00	9208.65	4380.19	5129.44	0.00	24,815.91	sqrt	-0.95
<i>Forb density</i>	n ha ⁻¹	5497.95	5442.15	0.00	20,259.03	3331.82	3915.02	0.00	22,865.36	sqrt	0.41
<i>Herbaceous height</i>	cm	48.41	20.09	12.13	97.63	43.30	23.28	4.75	127.50	sqrt	1.14
<i>Understorey richness</i>	n	13.35	5.38	5.00	24.00	12.29	5.65	3.00	29.00	-	0.31
<i>Ground cover</i>	-	0.82	0.14	0.35	1.00	0.79	0.17	0.30	1.00	arcsine	-0.12

^a Key to variables and their derivations: Basal area (BA) = $\pi \text{StemDens}(\text{DBH}_q/200)^2$. Quadratic mean DBH (DBH_q) = $\sqrt{(\sum \text{dbh}^2/N) \text{Stems} > 7.5 \text{ cm DBH in plot}}$. Stem density = \sum (n trees per 10 cm DBH class per plot X (10,000 m²/Plot area)). Canopy height = Mean height of the 3 tallest trees/shrubs/bamboo stems. St dev canopy height = Standard deviation in 3 tallest trees/shrubs/bamboo stems. N tree species = Total number of woody tree species greater than 7.5 cm DBH. Bamboo/Shrub/Forb density = n (shrubs, tall forbs, or bamboo) × (10,000 m²/Area of plot strip transects (2 × 22.6 m × 1.2 m)). Note: Bamboo and shrub density variables can be considered both understorey and canopy measures, depending on vegetation type and growth stage. Stem density includes woody heath and liane species, forb density includes all non-woody species (including lianes) taller than breast height. Herbaceous height = The mean of 20 height measurements based on 7 different height classes. Understorey richness = The number of non-woody canopy species. Canopy/ground cover = Cover variables represent the proportion of plant to non-plant cover, based on 20 intersect readings recorded along within-plot transects.

We modified this last method to suit our study's circumstances. Certain recommended variables were excluded because they are irrelevant to bongo (e.g. percent litter cover; hollow-bearing tree density). We also did not follow McElhinny et al. (2006) in creating a separate SCI for each vegetation class, as this would have introduced additional error by requiring an image classification to define vegetation class boundaries and multiple models to predict within-class SCI values. We thus developed a single SCI applicable across all vegetation classes.

To avoid biasing coefficients towards particular microhabitat structures favored by bongo, we performed PCA and Step 2 of the McElhinny et al. (2006) approach using the 90 non-presence plots, which provide an unbiased representation of study area microhabitat structure. The resulting coefficients were used to score all 127 plots.

2.2.2. Remote sensing

We used readily obtained, inexpensive optical RS imagery with 15–30 m native resolution, and rigorously pre-processed these to minimize the increased geometric and radiometric errors associated with mountainous, forested terrain. We extracted variables from these data using techniques that improve RS measurements of forest structural properties compared to conventional techniques such as NDVI (Sabol et al., 2002). Climate-related data from other sources were also employed to enhance our capacity to predict SCI values. These data provided useful, supplementary information on factors that we expected to influence plot-scale vegetation structure, despite being coarser in resolution than both the 22.6 m field plots and the optical imagery.

2.2.2.1. Imagery and pre-processing. The optical data we used were the visual/near-infrared (three 15 m bands covering 0.52–0.86 μm) and shortwave infrared (six 30 m bands covering 1.6–2.43 μm) bands of three cloud-free ASTER L1a scenes acquired January 23, 2007 (source: US Geological Survey, edcimswww.cr.usgs.gov). An SRTM digital elevation model (DEM, 90 m resolution, source: CGIAR, srtm.csi.cgiar.org) and an interpolated rainfall surface (1 km resolution, source: World Resources Institute, www.wri.org/publication/content/9291#rainfall, based on Hijmans et al. (2005)) provided supplementary data.

We ortho-rectified the ASTER images to minimize their substantial geometric distortions. Following Estes et al. (2008), positional accuracy (measured as root mean square error (RMSE)) was assessed against an ortho-rectified SPOT scene (RMSE = 10 m), resulting in values of 7 m or

less. We then converted these images to surface reflectance using the COST method (Chavez, 1988) modified for off-nadir viewing angles (Wu et al., 2005). We reduced image terrain shadowing using the SCS + C correction (Soenen et al., 2005), which also required a look angle adjustment (Estes et al., 2008). The necessary slope and aspect variables were derived from the SRTM DEM re-sampled to 15 and 30 m resolutions using a splining function. The corrected ASTER images were then mosaicked into a single, seamless scene of the study area.

2.2.2.2. RS-derived predictor variables

2.2.2.2.1. Spectral endmembers. Spectral mixture analysis (SMA) uses a linear model to estimate the relative abundance of the physical materials (endmembers) within each image pixel, using those materials' characteristic spectra (Adams et al., 1989). The resulting endmember (EM) fractions are better correlated with vegetation structural properties than other multispectral derivatives, e.g. NDVI (Sabol et al., 2002; Lu et al., 2003).

In forested habitats, solar reflectance is primarily influenced by green vegetation (GV), non-photosynthetic vegetation (NPV), soil, and canopy shadow (Sabol et al., 2002; Lu et al., 2003; Estes et al., 2008). For this analysis, we used available library EMs (ENVI, 2007) to represent GV, NPV, and soil because 1) site specific reference spectra were unavailable, and 2) finding pure image endmembers was difficult. Since canopy shadow is a mixture of pure shade and light transmitted through or reflected from canopy leaves, we obtained an EM that represented these two elements by averaging the least reflective pixels found in deep, shadowed valleys in the image (Sabol et al., 2002).

We performed SMA using all nine ASTER bands at 30 m resolution (VNIR bands were aggregated). EM fractions were constrained to sum to unity, but individual EM fractions could be less than zero or greater than one. Following Estes et al. (2008), fractional values between -1.05 and 1.05 were acceptable, and the criteria for selecting a model to represent the study area were: 1) >99.5% of pixels should fall within the acceptable fractional range; 2) model error should be <0.03 RMSE in >99% of pixels. In our selected model, 99.7 and 99.9% of pixels met the respective criteria.

2.2.2.2.2. Texture measures. Texture analysis assesses the spatial variation in tonal values around each image pixel, using a variety of statistical measures (Baraldi & Parmiggiani, 1995). We applied texture analysis to ASTER's 15 m red (0.63–0.69 μm) and near-infrared (0.76–

0.86 μm) channels (green (0.52–0.60 μm) was redundant with red), to calculate the mean and standard deviation of neighborhood reflectance. These two easily calculated measures correlate significantly with forest structural properties in this study area (Estes et al., 2008) and elsewhere (Wulder et al., 1998; Franklin et al., 2001). We used windows of 3×3 , 7×7 , 11×11 , 15×15 , and 21×21 pixels, informed by an earlier analysis suggesting that larger window sizes produced better correlations with bamboo density, while smaller windows were best for canopy structure measures, e.g. basal area (Estes et al., 2008).

2.2.2.2.3. DEM-derived climatic measures. We used the re-sampled 30 m DEM to extract several additional micro-climatic variables that we selected because of their possible influence on forest understorey structural properties not measurable by the ASTER sensor.

The first was the difference in noontime illumination between the two solstice dates (December 21st and June 21st), an easily calculated proxy measure of potential solar radiation, which influences plant physiology and diversity (Kumar et al., 1997). Using Riaño et al.'s (2003) equation for solar illumination (IL, based on variables for slope, aspect, solar azimuth, and solar zenith angles), we calculated this measure by subtracting IL_{June21} from $IL_{\text{December21}}$. The second measure was a modified version of the topographic relative moisture index, which characterizes soil moisture by calculating, re-scaling, and summing four DEM-derived measures: aspect, slope, slope shape, and relative slope position (Parker, 1982). The modification entails replacing aspect with the product of aspect and slope, in order to reflect the interaction of these two elements in determining solar exposure, which affects soil moisture (Manis et al., 2001). The final measure was elevation, a proxy for temperature.

2.3. Model development and performance evaluation

2.3.1. Predicting SCI values

The complete set of RS-derived variables is listed in Table 2. These were used to predict SCI values, with the exception of the GV SMA fraction because it was highly correlated with the more informative shade fraction (Sabot et al., 2002). We also squared several variables to test for possible non-linear relationships, including the texture variables, the topographical relief moisture index (TRMI), and the IL difference (ΔIL), orthogonalizing these to prevent collinearity (Kleinbaum et al., 1988).

We used least squares multiple regression to model SCIs with the RS-derived predictors. For each SCI, we selected a model that explained the most variation in SCI scores, provided that each predictor included 1) was significant at the $p < 0.1$ level, and 2) was not strongly correlated with any other (i.e. $|r| > 0.7$ and/or variance inflation factors (Belsley

et al., 1980) > 4). These last two criteria prevented over-fitting and multi-collinearity.

Following the same rationale we used to develop SCIs, we fit regression models to the SCI response values from the 90 non-presence plots to avoid biasing parameter estimates towards microhabitat structures selected by bongo. To screen for violations of the key regression assumptions of homoscedasticity, linearity, and normality, we examined plots of residuals versus predicted values and residual normal quantiles. We used the final models to predict SCI values for both non-presence and bongo plots, which respectively serve as the training and test datasets.

2.3.2. Evaluating the performance of the SCI-RS approach

We assessed the first two questions regarding the effectiveness of our approach using the observed SCI values. To see whether SCIs could effectively discern variation in microhabitat structure, we divided non-presence plots into four broad vegetation classes and tested class separability within each SCI using pair-wise Wilcoxon Rank Sum tests with Bonferroni-corrected alpha values (Sokal & Rohlf, 1995). To assess each SCI's ability to distinguish bongo-selected from available habitat, we used simple logistic regression to individually test how effectively SCIs discriminated between bongo and non-presence plots. We assessed model performance using the area under the curve (AUC) of the Receiver Operating Characteristic (ROC), which is a threshold-independent method for evaluating classification accuracy (Fielding & Bell, 1997). AUC measures the overlap in the distributions of predicted probabilities calculated for presence and non-presence observations, varying between 0.5 (superimposed) and 1.0 (completely separated). Put differently, AUC is the probability that a randomly selected presence observation will outscore a randomly drawn non-presence plot; 0.7 or higher indicates good predictive capacity (Fielding & Bell, 1997; Pearce & Ferrier, 2000).

To answer the third question, we examined the RS-based multiple regression models' (see Section 2.3.1 above) coefficients of determination (R^2) and RMSE values to determine how accurately each SCI was predicted. We used RMSEs calculated for the 37 bongo plots' predicted SCI values to independently assess model fitting error. Lastly, we re-ran the Wilcoxon Rank Sum Tests and simple logistic regressions with predicted SCI values to see how effectively they preserved the observed patterns of vegetation class and bongo/non-presence plot separation.

3. Results

3.1. Derived SCIs

The first three principal components (PCs) of the PCA were retained as SCIs because each accounted for at least 15% of the microhabitat structural data's variation (Table 3). The first PC (score

Table 2

Key to variables derived from the remotely-sensed data.

Variable name ^a	Description
GV	ASTER SMA GV fraction
NPV	ASTER SMA NPV fraction
Soil	ASTER SMA soil fraction
Shade	ASTER SMA shade fraction
AB2sd(3,7,11,15,21;sq)	Standard deviation texture images based on the ASTER red channel
AB3sd(3,7,11,15,21;sq)	Standard deviation texture images based on the ASTER near-infrared channel.
TRMI(sq)	The Topographic Relative Moisture Index, derived from SRTM DEM
ΔIL (sq)	The local illumination differences between solstices, derived from SRTM DEM
Elev(sq)	Elevation from 30 m re-sampled SRTM DEM
MAP	Mean annual precipitation (Hijmans et al., 2005)

^a Numbers in brackets indicate the diameter of the window size used to calculate the neighborhood statistic. "Sq" indicates that a squared-transform of the variable was also derived.

Table 3

Eigenvectors of the rotated principal components. Loadings greater than |0.5| are presented in bold, and the proportion of variance explained by the principal component is listed.

Variable	PC1	PC2	PC3
sqrt(Basal area)	0.90	0.12	−0.02
sqrt(DBHq)	0.78	0.07	0.11
sqrt(Stem density)	0.76	−0.23	0.01
Canopy height	0.69	0.53	−0.03
sqrt(Std dev canopy height)	0.46	0.06	0.15
sqrt(Tree species)	0.84	−0.12	0.11
sqrt(Bamboo density)	−0.19	0.78	−0.04
sqrt(Shrub density)	0.13	− 0.71	0.36
sqrt(Forb density)	0.10	−0.02	0.78
sqrt(Herbaceous height)	−0.14	−0.02	0.85
Understorey richness	0.25	−0.38	0.69
arcsine(Canopy cover)	0.36	0.78	0.01
arcsine(Ground cover)	−0.01	−0.23	0.29
Proportion variance	0.28	0.18	0.16

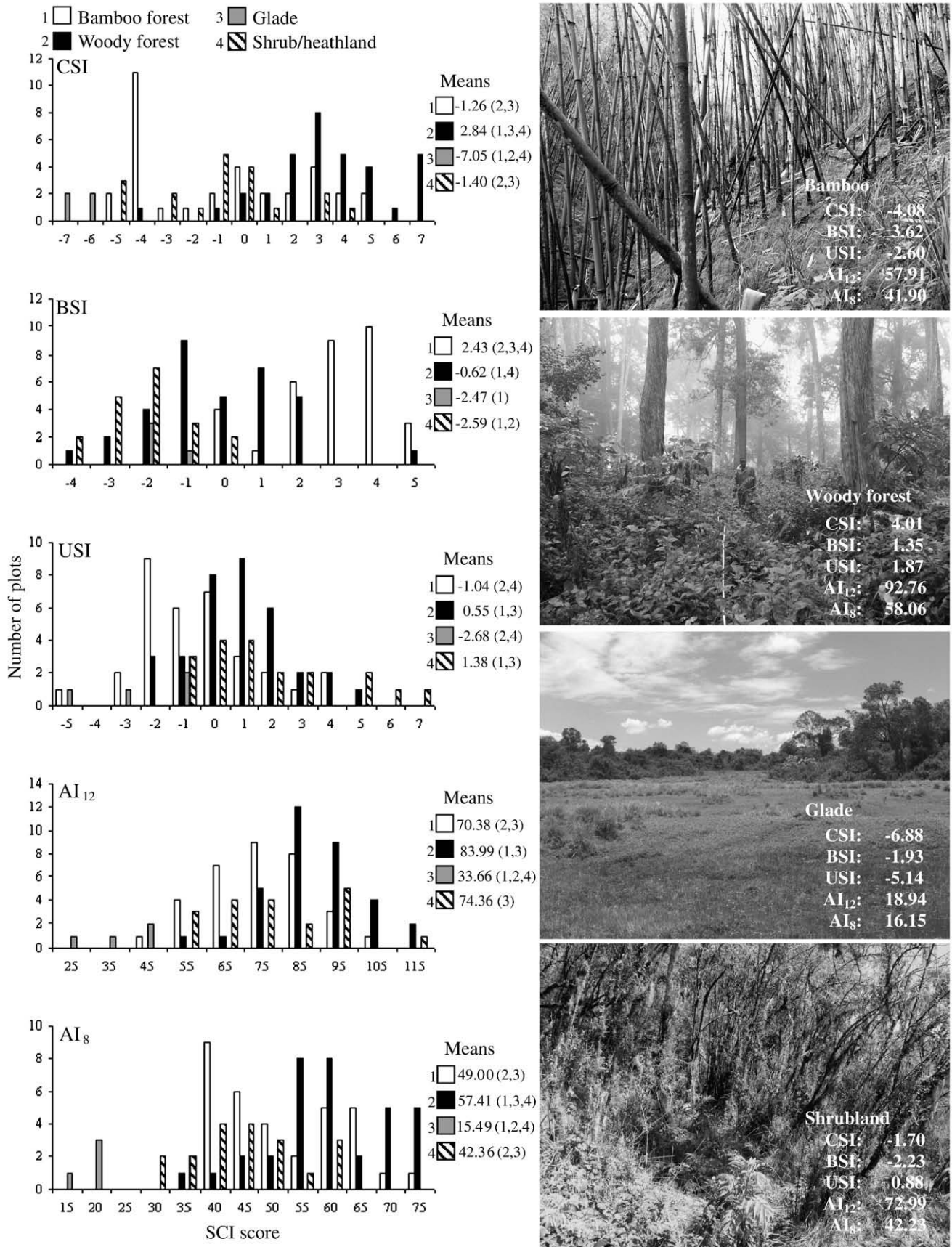


Fig. 2. SCI (canopy structure index (CSI); bamboo-shrub index (BSI); understorey structure index (USI); full and reduced additive indices (AI₁₂ and AI₈)) score distributions within four primary vegetation structural types (1. bamboo forest (n = 33); 2. woody forest (n = 34); 3. glades (n = 4); 4. shrublands and heathlands (n = 19)). Each class' mean SCI value and the corresponding classes (in parentheses) with which it differs significantly are presented. Photographs from sample plots illustrate the four vegetation types, together with recorded SCI scores (top to bottom: mature bamboo *Yushania alpina* forest; mature *Podocarpus latifolius* forest; short grass glade; *Hypericum revolutum* shrubland).

range = -7.2 to 6.7), the Canopy Structure Index (CSI), describes a gradient between treeless microhabitats (negative scores) and those with mature, woody forest cover (positive scores). Scores of the second PC, the bamboo-shrub index (BSI, range = -5.7 – 6.4), increase as woody shrub density declines and bamboo density rises. Scores of the third PC, the understory index (USI, range = -5.2 – 6.7), increase with the diversity and abundance of herbaceous understory plants.

We used 12 of 13 variables to construct McElhinny et al.'s (2006) additive index (hereafter AI_{12}), excluding the quadratic mean DBH because it was redundant with basal area (Spearman $r = 0.9$). We also constructed a second version of this index (hereafter AI_8), excluding four understory attributes that were least detectable with RS data (forb density, herbaceous height, understory richness, and ground cover). These two SCIs' scores ranged between 18.9–107.1 (AI_{12}) and 14.2–74.3 (AI_8).

3.1.1. SCIs and microhabitat structure

We divided non-presence plots into four classes representing the general structural patterns in the study area's vegetation: woody forest, bamboo forest, shrubland/heathland, and grassland. Fig. 2 presents the class-wise plot distributions for the five SCIs and indicates the classes whose means were significantly different.

CSI plot scores for woody forests were mostly positive, while glade plots were strongly negative. Bamboo and shrub/heathland plots had similar score distributions straddling zero, and comparable, slightly negative means that were not significantly different. All other pairwise comparisons of classes (e.g. woody forest versus bamboo) showed significant differences between mean plot scores.

For the BSI, bamboo plots were positive and had a significantly higher mean than the other three classes, which were mostly negatively scored. The only other significantly distinguishable class pair was woody forest and shrub/heathlands.

USI class scores showed the most overlap. Woody forest and shrub/heathland plots were primarily positive and statistically indistinguishable, but significantly higher than bamboo and glade plots. These last two classes were not significantly different.

AI_8 's class patterns were the same as CSI's: woody plots were significantly higher and glade plots were significantly lower than all other classes, while bamboo and shrub/heathland plots had intermediate values with statistically indistinguishable means. AI_{12} was similar except for the non-significant difference between the forest and shrub/heathland class pairing.

3.1.2. SCIs and microhabitat selection

Simple logistic regression models were run to test each SCI's ability to predict the probability of bongo presence (P). The model based on CSI, which was positively related to bongo presence, was highly significant (based on the likelihood ratio chi-square test) and correctly distinguished between bongo presence and non-presence plots 71% of the time ($P = -1.160 + 0.194 * CSI$, $p < 0.0004$, $AUC = 0.71$). Although significant, the BSI model was much weaker, both in p -value and the ability to distinguish bongo-selected from available habitat ($P = 0.974 + 0.173 * BSI$, $p < 0.04$, $AUC = 0.62$). USI was even less effective, producing a non-significant model not much better than chance for distinguishing between plot types ($P = -0.947 + 0.150 * USI$, $p < 0.08$, $AUC = 0.59$). Models using AI_{12} and AI_8 were similar to CSI in significance, sign, and plot classification success rate ($P = -4.984 + 0.050 * AI_{12}$, $p < 0.0001$, $AUC = 0.73$; $P = -3.805 + 0.054 * AI_8$, $p < 0.0004$, $AUC = 0.70$).

3.2. RS prediction of SCI values

3.2.1. Models and fit

The model selection criteria resulted in unique predictor sets for each RS-based multiple regression model. Fig. 3 details these models' structures, parameter estimates, and measures of fit, and plots predicted SCI values against observed SCI scores.

The models for CSI, BSI, AI_{12} and AI_8 were significant at the $p < 0.0001$ level, while the USI model was significant at $p < 0.01$. The CSI, AI_{12} , and AI_8 models each contained the same three SMA variables, with the shade and soil fractions being their most significant predictors ($p < 0.0001$). These models also contained the B3sd variable, but the window sizes used to calculate this measure varied in each model (3×3 for the CSI model up to 21×21 for AI_{12}). The CSI model had the most predictors (6) and was the only one to include the precipitation variable (MAP). The BSI model was primarily based on non-linear transforms of B2sd21 and B3sd21, but also included the squared elevation variable. USI was best predicted by DEM-derived variables, including the squared terms of TRMI (its strongest predictor) and elevation, and ΔIL .

The model for CSI explained the greatest amount of variation in the response (51%), followed by AI_8 (46%), BSI (41%), AI_{12} (34%), and USI (17%). The five models' (normalized) RMSE values for predicted non-presence plot values were more comparable: AI_{12} was lowest (0.16), followed by AI_8 and USI (0.17 each), CSI (0.19), and BSI (0.21). The BSI and USI models were least accurate, as bongo plots had substantially higher RMSEs than non-presence plots, jumping from 0.21 to 0.28 for BSI and from 0.17 to 0.41 for USI. In contrast, the other three models' had similar training and test accuracies, with bongo plot RMSEs increasing by only 0.01 (CSI and AI_{12}) or 0.02 (AI_8).

Scatter plots for the two best-fitting models, CSI and AI_8 , show that the range of predicted SCI scores was reduced relative to observed values; higher scores were somewhat under-predicted and lower scores over-predicted. This tendency increased as model R^2 decreased, and was most severe in the USI model.

3.2.2. Predicted SCI values and microhabitat structure

Fig. 4 (excluding the poorly-predicted USI) indicates that predicted SCI values generally preserved the relative positions of the observed class score distributions (Fig. 2), but their statistical separability was reduced. Predicted CSI woody forest plot scores remained significantly higher than all other classes, but the distinctions between bamboo, shrub/heathlands, and glades became non-significant. Glade plots were significantly overestimated, and thus statistically indistinct from bamboo and shrub/heathland plots. The BSI model maintained the separation between bamboo and other classes, but substantially over-predicted shrub/heathland plot values such that this class became superimposed with woody forest. Bamboo plot values predicted by the USI model (results not shown) remained significantly lower than forest plots and significantly higher than glade plots, but the separation between glade and forest classes was lost. As with the observed index, mean predicted AI_{12} glade plot scores were significantly lower than the other three classes' means, but the separation between the bamboo and woody forest classes was not preserved. The AI_8 model retained the most separability of the five predicted SCIs, losing just one significantly different pairing between the woody forest and shrub/heathland classes.

We mapped predicted SCIs to further assess their ability to distinguish microhabitat classes. Fig. 5 presents maps for CSI, BSI, and AI_8 over a portion of the study area containing woody forest, shrubland, and grassland, together with the locations of four non-presence plots and their predicted and observed values. The two left-hand plots (1 and 3) were collected in tall woody forest: Plot 1 in a dense *Cassipourea malosana* stand, and Plot 3 in mature *Podocarpus latifolius* forest. Both produced high CSI and AI_8 scores, although Plot 1's were higher because of its greater basal area and stem density. The respective regression models predicted correspondingly high scores, but eliminated the two plots' score differences: the CSI model under-predicted plot 1's score by 8% (of the observed CSI range), and over-predicted plot 3's value by 6%; the AI_8 model closely matched plot 1's score, but over-predicted plot 3's score by 12%. Both plots' predicted BSI values were 10–13% lower than their moderately positive observed values.

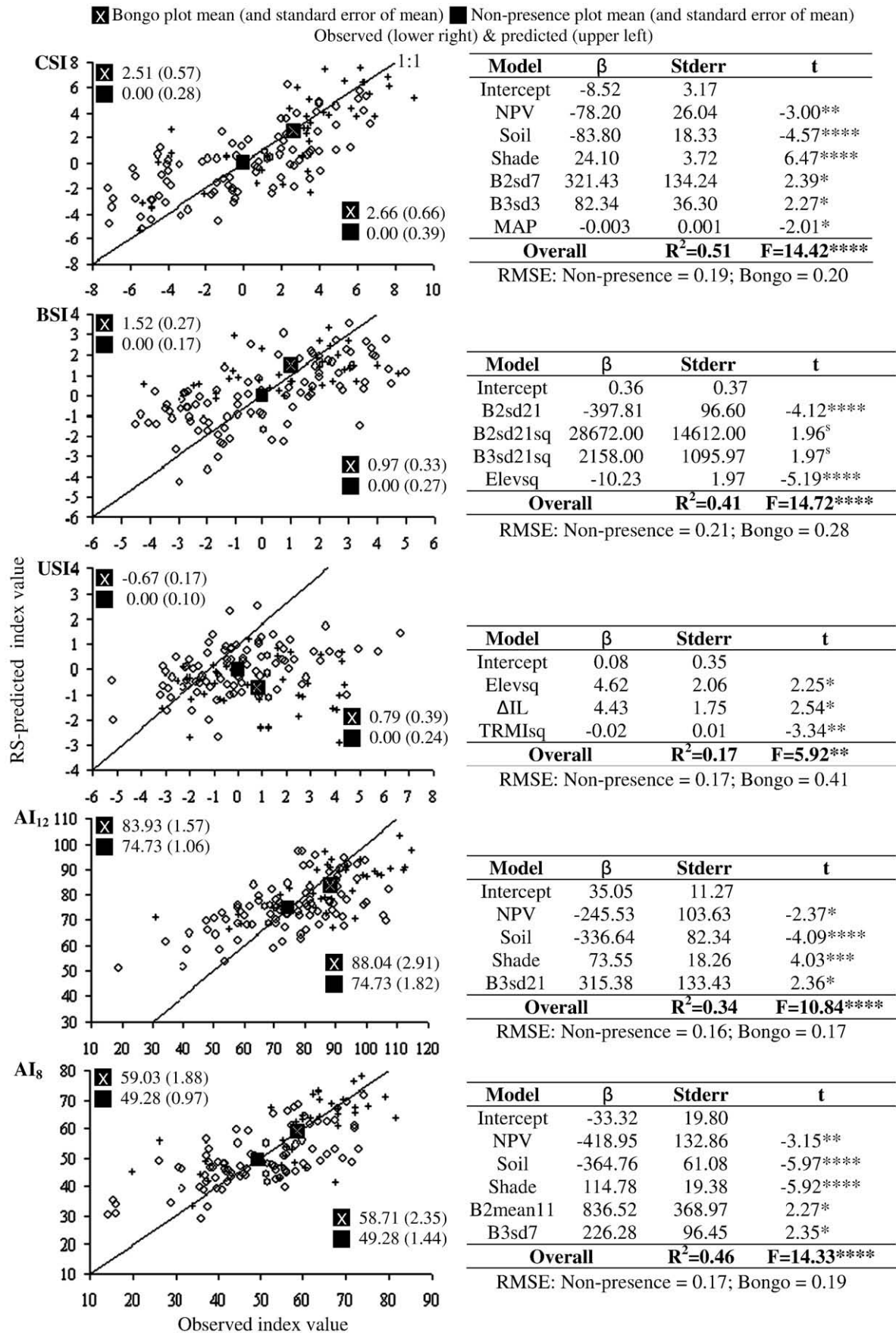


Fig. 3. Scatter plots of observed and RS-predicted SCI values for bongo (+) and non-presence plots (o), accompanied by plot means and standard errors. Tables display model parameter estimates (β), their standard errors, t and F -statistics with significance codes (**** $p < 0.0001$; *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; ^s $p < 0.10$; ns = non-significant), model R^2 and normalized RMSEs for non-presence (training) and bongo (test) plots.

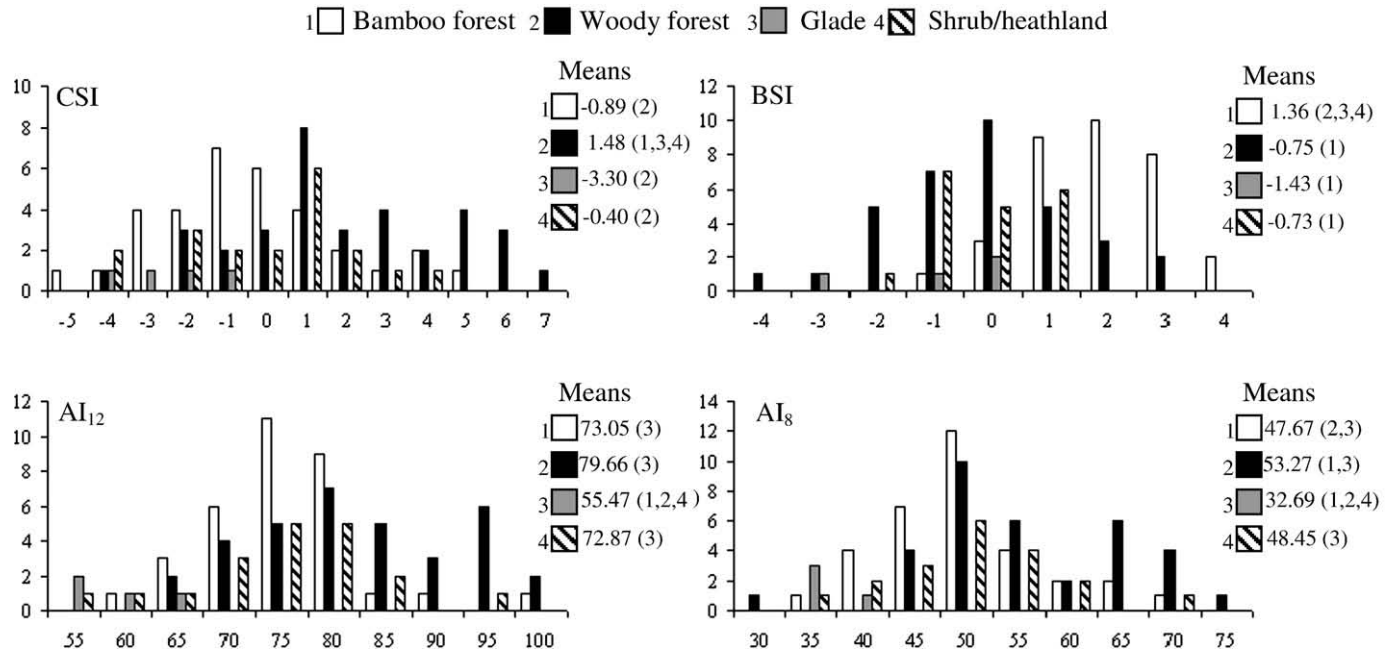


Fig. 4. Distributions of the predicted scores for the canopy structure index (CSI), bamboo-shrub index (BSI), and full and reduced additive indices (AI₁₂ and AI₈) at the 90 non-presence points, classified according to the four main vegetation structural types (1. bamboo forest (n = 33); 2. woody forest (n = 34); 3. glades (n = 4); 4. shrublands and heathlands (n = 19)). Each class' mean SCI value and the corresponding classes (in parentheses) with which it differs significantly are presented. Predicted SCI scores are on the x axis, and plot count is on the y axis.

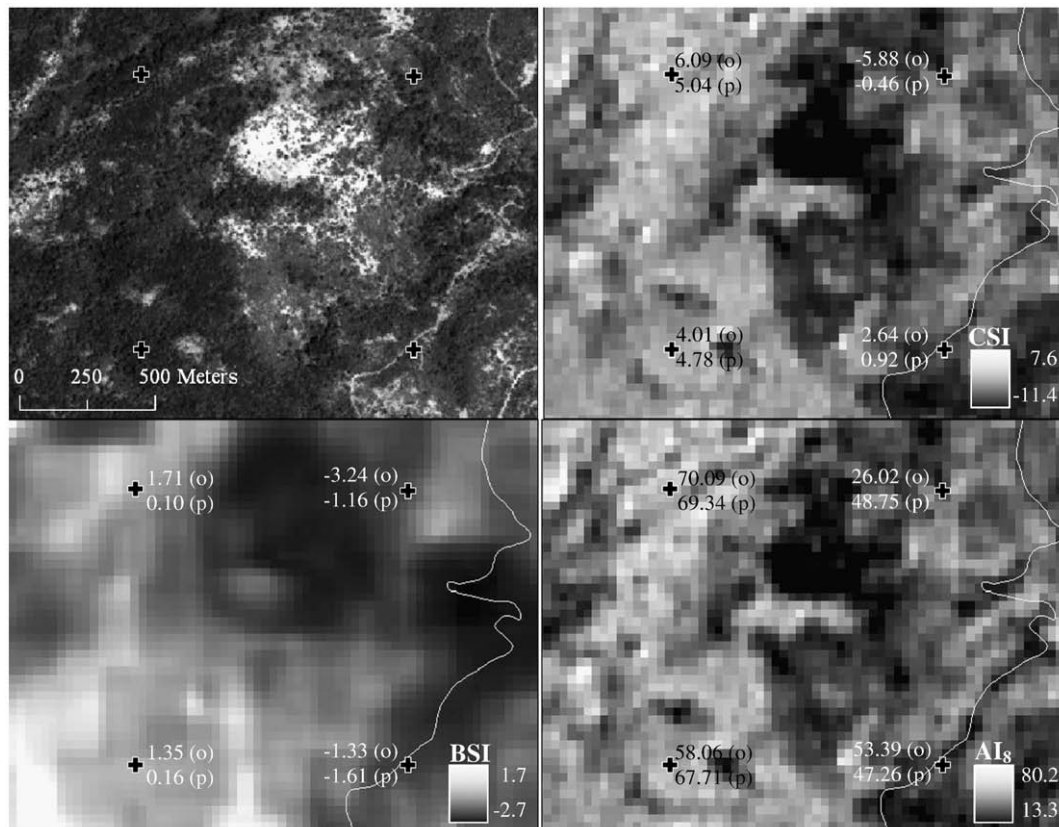


Fig. 5. CSI, BSI, and AI₈ maps for an area comprised of woody forest, shrubland, and grassland, as seen in a high resolution (5 m) SPOT image (top left)). Crosses indicate four non-presence plots. Plot 1 is located in dense *Cassipourea malosana* forest, plot 2 in open shrubland, plot 3 in mature *Podocarpus latifolius* forest (see second photo in Fig. 2), and plot 4 in a shrubland/forest mosaic next to a park road. Observed (o) and predicted (p) SCI values are listed.

Plot 2 (upper right) was collected in a small patch of open shrubland surrounded on three sides by mature forest. Plot 4 (lower right, adjacent to a park road), was in shrubland interspersed with

clumps of forest trees. Plot 2 produced low CSI and AI₈ scores (− 5.88 and 26.02), which the regression models over-predicted by 39 and 28%, respectively. Plot 4's CSI and AI₈ scores were higher because it

contained forest trees. In this case, the two models under-predicted scores by 12 (CSI) and 8 (AI₈) percent. BSI scores followed the same pattern at these two plots: plot 2's score was lowest and over-predicted, and plot 4's scores were higher and under-predicted, but by a smaller margin.

A visual assessment of mapped CSI and AI₈ values surrounding the four non-presence plot locations shows that predicted scores reflect the structural complexity patterns evident in the accompanying high resolution SPOT image (upper left in Fig. 5); CSI and AI₈ scores are low in open glades (the brightest areas in the SPOT image), intermediate in shrublands (lighter grey in the SPOT image), and high in forested areas (dark grey). BSI values are less easily interpreted; predicted BSI scores are generally lower in glades and shrublands, as expected, but show substantial variation in forests that cannot be explained by examining the SPOT image.

3.2.3. Modeling habitat selection with predicted SCI values

Simple logistic regressions were run to test each RS-predicted SCI's ability to discern bongo habitat selection. Predicted CSI values produced a slightly more significant model than the original, and had a one point improvement in AUC ($P = 1.2418 + 0.2848 * \text{CSI}$, $p < 0.0000$, $\text{AUC} = 0.72$). Predicted AI₁₂ and AI₈ showed the same tendency, and increased AUC by 1 and 4 points, respectively ($P = -8.24631 + 0.09275 * \text{AI}_{12}$, $p < 0.0000$, $\text{AUC} = 0.74$; $P = -5.87162 + 0.09221 * \text{AI}_8$, $p < 0.0000$, $\text{AUC} = 0.74$). Predicted BSI discriminated between bongo and non-presence plots much more effectively than observed values, having the highest AUC of all predicted and observed models ($P = -1.335 + 0.6277 * \text{BSI}$, $p < 0.0000$, $\text{AUC} = 0.75$). Predicted USI values showed a similar jump in classification performance ($P = -1.1296 - 0.7385 * \text{USI}$, $p < 0.0004$, $\text{AUC} = 0.69$), but the sign of the relationship between bongo presence and predicted USI was reversed compared to the observed USI logistic regression model.

4. Discussion

4.1. Effectiveness of SCIs in discerning microhabitat structure and selection

Multivariate statistical analyses are generally more sensitive than univariate approaches, thus condensing microhabitat data into SCIs could have undermined our ability to discern habitat structure and selection. Although the SCIs we tested presumably masked information that a multivariate analysis would reveal, both PCA and McElhinny et al.'s (2006) technique produced indices that could distinguish between most microhabitat structural classes and—most importantly for distribution modeling—between bongo-selected and available habitats.

CSI and AI₈ were the two most effective SCIs for separating between microhabitat classes, showing significant differences between all but the bamboo and shrub/heathland classes. AI₁₂, BSI, and USI were comparable to one another, with each producing significant differences between four of six class pairings. CSI and AI₈ failed to separate the bamboo-shrub/heathland pairing because both classes were defined by an absence or scarcity of trees, thus were not differentiable by the canopy structure measures that dominate these two SCIs. However, AI₈ showed more (albeit statistically non-significant) difference between the two classes (bamboo mean = 49.00; shrub/heathland mean = 42.36) than CSI (Fig. 2), because it was constructed with two variables (shrub and bamboo densities) that correlated weakly with CSI scores. This difference highlights the advantage of McElhinny et al.'s (2006) manual, additive approach over PCA, which gives the user less control because variable selection (determined by loading patterns) is automatic.

AI₈ and CSI, together with AI₁₂, were also the best predictors of bongo habitat selection, as shown by their AUC values of 0.7 or greater (Pearce & Ferrier, 2000). Contrary to expectations, based on our

understanding that bongo microhabitat selection is primarily driven by understorey structural characteristics, USI was ineffective for discriminating between bongo and non-presence plots. This unanticipated result may be attributed to the fact that USI was the third PCA axis, and thus explained less microhabitat structural variation than CSI and BSI. Furthermore, USI did not correlate strongly with shrub density, which significantly influences bongo habitat selection (Estes et al., 2008).

4.2. Effectiveness of SCIs for mapping microhabitat structure

Having established that SCIs can be useful logistic regression predictors, we then determined their capacity for improving microhabitat mapping accuracy. Working in rough terrain that increases detection error, we predicted SCIs with substantially greater accuracy than reported by other studies that used RS to quantify individual, microhabitat-scale structural attributes. For instance, Tomppo et al. (2008) reported RMSEs greater than 40% for RS-predicted timber volumes at sub-hectare scales. Stickler and Southworth (2008) found a maximum R^2 of 0.24 in several models relating plot-scale (0.09 ha) basal area, canopy height, and understorey vegetation height to variables derived from Landsat and Quickbird images. Jeganathan et al. (2004) estimated bush density within 0.01 ha plots using raw Landsat reflectance in a model explaining 28% of variance. Using ASTER-derived SMA and texture variables, our models for CSI and AI₈ were ~60–110% more accurate (in terms of RMSE and R^2 values) than these studies' results. Even the AI₁₂ model's lower R^2 (0.34) was ~20–40% higher than those presented in these other examples. These performance improvements were validated by the closeness of bongo and non-presence plot RSMES calculated for each of the three SCIs (although the BSI model's error measurements equaled or exceeded AI₁₂'s, the higher bongo plot RMSE showed the model was not robust).

Our models were less accurate than Coops and Catling's (2002), which achieved an R^2 of 0.76 in predicting the habitat complexity score (HCS), an ordinal SCI constructed by adding several simple, categorical habitat structure measures (Newsome & Catling, 1979) that was also used in two other faunal microhabitat modeling studies (Coops & Catling, 1997; Gibson et al., 2004). However, these studies were conducted in small areas with gentler terrain, using a 0.25 ha field plot and multispectral airborne videography. Given these studies' coarser plot scale and higher image resolution, our results compare favorably. Furthermore, our methods are more broadly applicable in terms of image cost, availability, and footprint.

By re-running univariate tests for vegetation class differences and logistic regressions using predicted SCI values, we could more fully assess RS modeling accuracy. The logistic regressions based on RS-predicted SCIs confirmed what the independent RMSE error assessment showed for BSI and USI: that they were not suitable for RS-based mapping. Although both BSI's and USI's predicted values produced large increases in AUC, this "improvement" is attributable to these two models' larger prediction errors for bongo plots relative to non-presence plots. The negative coefficient for predicted USI is an obvious manifestation of this error, given that the relationship between observed USI and bongo presence was positive. The poor fit of the RS-BSI model was also evident in the inconsistent relationships between mapped BSI values and observable vegetation structure (Fig. 5). This inconsistency was an artifact of the standard deviation texture measures that were the primary predictors in the BSI model; such texture variables simply describe local variations in reflectance, and do not necessarily relate to specific vegetation structures (e.g., given a uniformly dense bamboo forest next to a large glade, a pixel adjacent to the glade would have a higher texture score than one in the interior, with correspondingly different predicted BSI scores).

The same textural artifacts were also evident in an AI₁₂ map (not illustrated). These irregularities, together with its lower model R^2 and relative ineffectiveness for separating vegetation classes, make AI₁₂

less useful for mapping microhabitat features, despite its consistent strength in distinguishing bongo-selected from available habitats (using both observed and predicted values) and the RS- AI_{12} model's robustness (based on the similarity of non-presence and bongo plot RMSEs). AI_{12} 's capacity to be accurately predicted with RS data was reduced by including the four understorey measures most poorly correlated with optical RS data in the index. The additional variation these measures added to the index could not be explained by the topographic and climate-related predictors we used, although we had expected that these variables, describing abiotic controls on plant growth (Parker, 1982; Kumar et al., 1997), would improve the predicted accuracy of SCIs incorporating understorey variables. The RS-USI model showed that these predictors explained a significant (albeit small) proportion of understorey variation, but they did not appear in the RS- AI_{12} model because it was dominated by the far stronger relationships between the canopy structure measures and the ASTER-derived predictors. Furthermore, the four additional understorey measures used to construct AI_{12} merely added "noise" within the index that might have been difficult to resolve even if the non-optical predictors were more effective. This noise problem was also evident in observed AI_{12} 's inferiority to AI_8 in separating vegetation classes.

Overall, RS-predicted AI_8 performed best in this second round of statistical tests, losing only one of the five significant observed vegetation class differences (woody forest-shrub/heathland) in the univariate analysis, and having a logistic regression predictive performance comparable to observed AI_8 's (the discrepancy of 0.04 between predicted and observed model AUCs was not sufficiently large to question the RS- AI_8 regression model's fit). Predicted CSI, on the other hand, could only distinguish woody forest from the other three vegetation classes, losing the separation between the bamboo-glade and glade-shrub/heathland class pairings. AI_8 , on the other hand, maintained the ability to distinguish between these classes because it included the bamboo and shrub density measures. Unlike the other excluded understorey measures, these two variables correlate significantly with several ASTER-derived texture and SMA variables (Estes et al., 2008). AI_8 thus struck the best balance between information content and RS predictability, because it included the two understorey variables most detectable with optical RS data and most important for distinguishing the bamboo and shrub/heathland classes, but excluded those variables that would obscure (as in AI_{12}) the stronger relationship between the ASTER data and the canopy structure components.

4.2.1. SCIs and RS prediction error

The greater accuracy of this RS-based microhabitat mapping approach compared to those used in other studies suggests that SCIs, when properly constructed, correlate better with RS data than individual vegetation structure measures. This improved correlation may be related to the functional similarity between an SCI and reflected solar radiation: reflectance expresses the interaction of radiation with the numerous structural and compositional properties within a habitat (Danson & Curran, 1993; Gemmell, 1995), much as an SCI numerically summarizes those multiple properties. Therefore, reflectance and an SCI both summarize habitat features, but in two different forms: one is photonic, the other numeric.

A portion of these accuracy gains should also be attributed to the image-processing techniques we applied. We combined several pre-processing and image analysis methods, such as SMA and texture analysis, which have individually been proven useful (e.g. Wulder et al., 1998; Sabol et al., 2002; Soenen et al., 2005), but have not been jointly applied to improve RS-based microhabitat mapping accuracy.

In terms of pre-processing, orthorectification removed the substantial geo-local error that can occur in mountainous terrain, reducing residual positional error to less than 17 m, while topographic normalization minimized illumination variability that would have further confounded microhabitat structural predictions.

For extracting image information, spectral mixture analysis (SMA), in particular, provided an advantage for measuring forest structural properties over more widely used image derivatives, such as NDVI (Sabol et al., 2002; Peddle et al., 2001). By using the known spectral response of common habitat elements, SMA directly connects RS data with biophysical properties, extracting the relevant information from reflectance data while consigning unexplained spectral variation to the error term (Sabol et al., 2002; Adams et al., 1989). This filtering characteristic not only improves habitat structure measurements, but also permits resulting RS-habitat relationships to be generalized over broader areas (Sabol et al., 2002). Hence SMA-derived variables are more advantageous for mapping habitat features than raw reflectance (a more widely used approach), as habitat-reflectance relationships can be harder to interpret and are less transportable because local field data are needed to calibrate parameter estimates to each scene's idiosyncrasies (Jeganathan et al., 2004). Combining texture measures with the SMA variables in the SCI models further boosted accuracy by adding habitat structure information from the spatial domain to the within-pixel spectral data (Wulder et al., 1998).

Of course, not all RS prediction error could be removed using these combined techniques, particularly given the fine-scale of our study. One error source that presumably decreased the accuracy of the RS-SCI models was mismatches between field plot and pixel centers, which can occur even in perfectly geo-located images. For example, a small forest opening of 30–50 m diameter could be divided between two image pixels. A plot centered in the middle of this opening would thus record low structural diversity, but each pixel's signature would include reflectance from more complex forest habitats. The model would thus overpredict the plot's habitat complexity, which may account for the inflated SCI predictions reported for Plot 2 in Fig. 5.

Another potential error source was residual terrain shadowing resulting from the mismatch between the native resolution of the DEM (90 m) and the ASTER sensor, which could have caused SCI models incorporating the shade fraction to slightly overpredict SCI index scores. Field data collection methods may also have caused error. For instance, time constraints prevented us from recording the spatial arrangement of habitat elements, which contains important structural information (Pommerening, 2002). Plots with different configurations (e.g. 5 clumped versus 5 evenly spaced trees) may have had similar observed SCI scores but substantially different reflectance values, thereby causing predicted SCI scores to diverge. Our grid-based, non-presence sample set was another possible source of error. The collection methodology for this dataset was intended to provide an unbiased representation of study area microhabitat availability. Using this set to train SCIs and regression models while excluding bongo samples avoided bias towards bongo-favored habitat configurations. In actuality, three variables (canopy height, stdev of canopy height, and bamboo density; see Table 1) had smaller maxima in the non-presence set than in the bongo set, thus the non-presence plots did not capture the full structural range of sampled habitat. These divergences indicate that some extrapolation occurred between the SCI/model training and testing processes, thereby increasing error potential. However, extrapolation-related error should be small, since only a few observations (canopy height = 4, stdev canopy height = 3, bamboo density = 1) were responsible for these three variables' range discrepancies.

4.3. Broader application

The approach we demonstrated in this study has the potential to be more widely applied. Locally, our model fitting procedure (using the suitably representative non-presence point data to train the RS-SCI models, and excluding the habitat-specific bongo plots) permits us to create an unbiased habitat structure map for the Aberdares. This map allows the number of presence observations to be more rapidly increased, because it permits collecting and incorporating additional bongo sightings without having to record the associated microhabitat

data. Similarly, this map may be used to model the distributions of other resident species.

This approach can also be readily transferred to different areas, since the SCI and image-processing methods have been proven in different ecosystems. The technique used to construct AI_{12} and AI_3 was developed and tested in Australian woodland and sclerophyllous dry forest (McElhinny et al., 2006), while structural complexity measures have been effectively used to model species distributions in different ecosystems, including a bird (Gibson et al., 2004) and several marsupials (Coops & Catling, 2002) found in Australian coastal forests and heathlands. SMA has been applied to forests ranging from the boreal (Peddle et al., 2001) to the Amazonian (Lu et al., 2003), a variety of grassland, shrubland, and woodland types (Asner & Heidebrecht, 2002; Roberts et al., 1993), and even for measuring urban vegetation (Song, 2005). Texture analysis has been used for assessing habitat structure in a similar range of vegetation types (Franklin et al., 2001; Wulder et al., 1998; St-Louis et al., 2006). This methodology should be accessible to most ecologists, as all image-processing techniques described in this study can be implemented using built-in functions found in commercially available IDL/ENVI software (ENVI, 2007). IDL code that allows batch image pre-processing (for surface reflectance conversion and topographic normalization) can be downloaded from sites.google.com/site/mountainbongo/dataandcode.

5. Conclusion

Our results confirm that quantitative SCIs are effective for discerning microhabitat variability and selection, and facilitate improved microhabitat mapping accuracy when combined with advanced RS analysis techniques. Of the SCIs tested, AI_3 most effectively addressed these two aspects, and the manual, additive technique (McElhinny et al., 2006) used to calculate it is preferable to PCA because it allows the user greater control over variable selection.

These results are relevant for both ecological and remote sensing science. This study demonstrated that RS can be successfully applied to the fine spatial scales that are fundamental to habitat selection, and of greatest interest to ecologists (Hilden, 1965; Turner et al., 2003). Ecologists can use these (relatively accessible) methods to improve the quality and increase the scope of spatial distribution models, which can in turn facilitate sounder ecological insight and more informed conservation action. Aside from the gains in mapping accuracy, using the SCI approach further benefits distribution modeling by compressing important microhabitat data into a single, statistically effective measure. Given that vegetation structure is often just one of several factors influencing habitat selection, using an SCI to represent this habitat aspect conserves degrees of freedom that can be used to test other possibly influential variables. This consideration is important for modeling rare species, for which sample sizes are usually small. Lastly, this work suggests that additional investigation into alternate methods for characterizing habitat data may hold promise for further improving RS-based habitat mapping accuracy.

Acknowledgements

We thank Assistant Director Joseph Warutere (retired), former Senior Wardens Anne Kahihia and Catherine Wambani, Assistant Warden Gichohi, and the ranger staff of the Aberdares National Park for supporting and facilitating this research, as well as the former and current Directors of Biodiversity, Ecosystems and Research for the Kenya Wildlife Service (KWS), Drs. Richard Bagine and Samuel Kasiki. The Bongo Surveillance Team (Michael Prettejohn, Joseph Kariuki, Matthew Gichuri, Kiragu Gichuri, Boniface Nderitu, Peter Mwangi, Stanley Gichure, Laban Kariuki, and Josphat Korage) were instrumental in collecting field observations. Jessica and Henry Henley, Francis Maina, Charlie Babault, Adrian Babault, and Leo and Bella Niskanen provided logistical support. Jacquelyn Bishop (University of Cape Town), Dorcas Kavembe, Joseph Jung'a, and Charles

Kimwele (the University of Nairobi), and Katherine Dunbar, Patricia Faria, and Michael Bruford (Cardiff University) validated bongo presence observations using mitochondrial DNA analysis of associated dung samples. Additional assistance was provided by Colin Church of the Rhino Ark Foundation, Olivier Hanotte and Dorine Adhoch of the International Livestock Research Institute, and Christian Lambrechts of the United Nations Environmental Programme. This work was supported by NASA Headquarters under the Earth System Science Fellowship Grant NNG05GR43H. Additional funding was provided by the Wildlife Conservation Society (Conservation Research Fellowship), the Explorers Club Washington Group (Exploration and Field Research Grant), and the University of Virginia's Department of Environmental Science. John Porter, David Richardson, Robert Davis, Robert Washington-Allen, David Carr, Henry Wilbur, Richard Estes, and Anna Estes provided valuable advice on presentation and methodology.

References

- Adams, J. B., Smith, M. O., & Gillespie, A. R. (1989). Simple models for complex natural surfaces: A strategy for the hyperspectral era of remote sensing. *Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS 1989) 12th Canadian Symposium on Remote Sensing* (pp. 16–21).
- Asner, G. P., & Heidebrecht, K. B. (2002). Spectral unmixing of vegetation, soil and dry carbon cover in arid regions: comparing multispectral and hyperspectral observations. *International Journal of Remote Sensing*, 23, 3939–3958.
- Baraldi, A., & Parmiggiani, F. (1995). An investigation of the textural characteristics associated with gray level co-occurrence matrix statistical parameters. *IEEE Transactions on Geoscience and Remote Sensing*, 33, 293–304.
- Belsley, D., Kuh, E., & Welsch, R. (1980). *Regression diagnostics*. New York: Wiley.
- Chavez, P. S. (1988). An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. *Remote Sensing of Environment*, 24, 459–479.
- Cohen, W. B., & Spies, T. A. (1992). Estimating structural attributes of Douglas-fir/Western hemlock forest stands from Landsat and SPOT imagery. *Remote Sensing of Environment*, 41, 1–17.
- Coops, N. C., & Catling, P. C. (1997). Utilising airborne multispectral videography to predict habitat complexity in eucalypt forests for wildlife management. *Wildlife Research*, 24, 691–702.
- Coops, N. C., & Catling, P. C. (2002). Prediction of the spatial distribution and relative abundance of ground-dwelling mammals using remote sensing imagery and simulation models. *Landscape Ecology*, 17, 173–188.
- Danson, F. M., & Curran, P. J. (1993). Factors affecting the remotely sensed response of coniferous forest plantations. *Remote Sensing of Environment*, 43, 55–65.
- East, R. (1999). *African Antelope Database 1998*. Gland, Switzerland: IUCN.
- ENVI. (2007). *Environment for visualizing images*. Boulder, CO, USA: ITT Visual Information Solutions.
- Estes, L. D., Okin, G., Mwangi, A., & Shugart, H. (2008). Habitat selection by a rare forest antelope: A multi-scale approach combining field data and imagery from three sensors. *Remote Sensing of Environment*, 112, 2033–2050.
- Fielding, A. H., & Bell, J. F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, 24 (1), 38–49.
- Franklin, S. E., Wulder, M. A., & Gerylo, G. R. (2001). Texture analysis of IKONOS panchromatic data for Douglas-fir forest age class separability in British Columbia. *International Journal of Remote Sensing*, 22(13), 2627–2632.
- Gemmell, F. M. (1995). Effects of forest cover, terrain, and scale on timber volume estimation with Thematic Mapper data in a rocky mountain site. *Remote Sensing of Environment*, 51, 291–305.
- Gibson, L. A., Wilson, B. A., Cahill, D. M., & Hill, J. (2004). Spatial prediction of rufous bristlebird habitat in a coastal heathland: A GIS-based approach. *Journal of Applied Ecology*, 41, 213–223.
- Guisan, A., & Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*, 135, 147–186.
- Hansen, M. J., Franklin, S. E., Woudsma, C., & Peterson, M. (2001). Forest structure classification in the North Columbia mountains using the Landsat TM Tasseled Cap wetness component. *Canadian Journal of Remote Sensing*, 27, 20–32.
- Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25, 1965–1978.
- Hilden, O. (1965). Habitat selection in birds. *Annales Zoologici Fennici*, 2, 53–57.
- Hyde, P., Dubayah, R., Walker, W., Blair, J. B., Hofton, M., & Hunsaker, C. (2006). Mapping forest structure for wildlife habitat analysis using multi-sensor (LiDAR, SAR/InSAR, ETM+, Quickbird) synergy. *Remote Sensing of Environment*, 102, 63–73.
- Jeganathan, P., Green, R. E., Norris, K., Vogiatzakis, I. N., Bartsch, A., Wotton, S. R., et al. (2004). Modelling habitat selection and distribution of the critically endangered Jerdon's courser *Rhinoptilus bitorquatus* in scrub jungle: An application of a new tracking method. *Journal of Applied Ecology*, 41, 224–237.
- Kayitakire, F., Hamel, C., & Defourny, P. (2006). Retrieving forest structure variables based on image texture analysis and IKONOS-2 imagery. *Remote Sensing of Environment*, 102, 390–401.
- Kerr, J. T., & Ostrovsky, M. (2003). From space to species: Ecological applications for remote sensing. *Trends in Ecology & Evolution*, 18, 299–305.

- Kingdon, J. (1982). *East African mammals: An atlas of evolution in Africa*. New York: Academic Press.
- Kleinbaum, D., Kupper, L., & Muller, K. (1988). *Applied regression analysis and other multivariable methods*. Boston: PWS-Kent Publishing Co..
- Kumar, L., Skidmore, A., & Knowles, E. (1997). Modelling topographic variation in solar radiation in a GIS environment. *International Journal of Geographical Information Science*, 11(5), 475–497.
- Lu, D. S., Moran, E., & Batistella, M. (2003). Linear mixture model applied to Amazonian vegetation classification. *Remote Sensing of Environment*, 87, 456–469.
- Manis, G., Lowry, J., & Ramsey, R. D. (2001). Preclassification: An ecologically predictive landform model. *Gap Analysis Bulletin*, 10, 11–12.
- McElhinny, C., Gibbons, P., & Brack, C. (2006). An objective and quantitative methodology for constructing an index of stand structural complexity. *Forest Ecology and Management*, 235, 54–71.
- McElhinny, C., Gibbons, P., Brack, C., & Bauhus, J. (2005). Forest and woodland stand structural complexity: Its definition and measurement. *Forest Ecology and Management*, 218, 1–24.
- Neumann, M., & Starlinger, F. (2001). The significance of different indices for stand structure and diversity in forests. *Forest Ecology and Management*, 145, 91–106.
- Newsome, A. E., & Catling, P. C. (1979). Habitat preferences of mammals inhabiting heathlands of warm temperate coastal, montane and alpine regions of southeastern Australia. In R. L. Specht (Ed.), *Ecosystems of the world: Heathlands and related shrublands of the world, Vol. 9A*. (pp. 301–316) Amsterdam: Elsevier.
- Osborne, P., Alonso, J., & Bryant, R. (2001). Modelling landscape-scale habitat use using GIS and remote sensing: a case study with great bustards. *Journal of Applied Ecology*, 38, 458–471.
- Parker, A. (1982). The topographic relative moisture index: an approach to soil-moisture assessment in mountain terrain. *Physical Geography*, 3, 160–168.
- Pearce, J., & Ferrier, S. (2000). Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling*, 133, 225–245.
- Peddle, D. R., Brunke, S. P., & Hall, F. G. (2001). A comparison of spectral mixture analysis and ten vegetation indices for estimating boreal forest biophysical information from airborne data. *Canadian Journal of Remote Sensing*, 27, 627–635.
- Pommerening, A. (2002). Approaches to quantifying forest structures. *Forestry*, 75, 305–324.
- Riaño, D., Chuvieco, E., Salas, J., & Aguado, I. (2003). Assessment of different topographic corrections in Landsat-TM data for mapping vegetation types. *IEEE Transactions on Geoscience and Remote Sensing*, 41, 1056–1061.
- Roberts, D. A., Smith, M. O., & Adams, J. B. (1993). Green vegetation, nonphotosynthetic vegetation, and soils in AVIRIS data. *Remote Sensing of Environment*, 44, 255–269.
- Rushton, S. P., Ormerod, S. J., & Kerby, G. (2004). New paradigms for modelling species distributions? *Journal of Applied Ecology*, 41, 193–200.
- Sabol, D. E., Gillespie, A. R., Adams, J. B., Smith, M. O., & Tucker, C. J. (2002). Structural stage in Pacific Northwest forests estimated using simple mixing models of multispectral images. *Remote Sensing of Environment*, 80, 1–16.
- Schadt, S., Revilla, E., Wiegand, T., Knauer, F., Kaczensky, P., Breitenmoser, U., et al. (2002). Assessing the suitability of central European landscapes for the reintroduction of Eurasian lynx. *Journal of Applied Ecology*, 39, 189–203.
- Schmitt, K. (1991). The vegetation of the Aberdare National Park, Kenya. *Hochgebirgsforschung, Band 7, 1–250* Innsbruck: Universitätsverlag Wagner.
- Soenen, S. A., Peddle, D. R., & Coburn, C. A. (2005). SCS + C: A modified sun-canopy-sensor topographic correction in forested terrain. *IEEE Transactions on Geoscience and Remote Sensing*, 43, 2148–2159.
- Sokal, R. R., & Rohlf, F. J. (1995). *Biometry*. San Francisco: W. H. Freeman.
- Song, C. (2005). Spectral mixture analysis for subpixel vegetation fractions in the urban environment: How to incorporate endmember variability? *Remote Sensing of Environment*, 95, 248–263.
- Stickler, C. M., & Southworth, J. (2008). Application of multi-scale spatial and spectral analysis for predicting primate occurrence and habitat associations in Kibale National Park, Uganda. *Remote Sensing of Environment*, 112, 2170–2186.
- St-Louis, V., Pidgeon, A. M., Radeloff, V. C., Hawbaker, T. J., & Clayton, M. K. (2006). High-resolution image texture as a predictor of bird species richness. *Remote Sensing of Environment*, 105, 299–312.
- Theau, J., Peddle, D. R., & Duguay, C. R. (2005). Mapping lichen in a caribou habitat of Northern Quebec, Canada, using an enhancement-classification method and spectral mixture analysis. *Remote Sensing of Environment*, 94, 232–243.
- Tomppo, E., Olsson, H., Stahl, G., Nilsson, M., Hagner, O., & Katila, M. (2008). Combining national forest inventory field plots and remote sensing data for forest databases. *Remote Sensing of Environment*, 112, 1982–1999.
- Turner, W., Spector, S., Gardiner, N., Fladeland, M., Sterling, E., & Steininger, M. (2003). Remote sensing for biodiversity science and conservation. *Trends in Ecology & Evolution*, 18, 306–314.
- Vaughan, I. P., & Ormerod, S. J. (2003). Improving the quality of distribution models for conservation by addressing shortcomings in the field collection of training data. *Conservation Biology*, 17, 1601–1611.
- Walsh, P. D., & White, L. J. T. (1999). What it will take to monitor forest elephant populations. *Conservation Biology*, 13, 1194–1202.
- Wu, J., Wang, D., & Bauer, M. E. (2005). Image-based atmospheric correction of QuickBird imagery of Minnesota cropland. *Remote Sensing of Environment*, 99, 315–325.
- Wulder, M. A., Ellsworth, F. L., Franklin, S. E., & Lavigne, M. B. (1998). Aerial image texture information in the estimation of Northern deciduous and mixed wood forest leaf area index (LAI). *Remote Sensing of Environment*, 64, 64–76.