Mapping habitat and biological diversity in the Maasai Mara ecosystem

B. O. Oindo , A. K. Skidmore & P. de Salvo

Taylor & Francis
Taylor & Francis Group

Mapping habitat and biological diversity in the Maasai Mara ecosystem

B. O. OINDO[1]*, A. K. SKIDMORE[2]

[1] National Environment Management Authority (NEMA), P.O. Box 47146 Nairobi, Kenya
[2] International Institute for Aerospace Survey and Earth Sciences (ITC), PO Box 6, 7500 AA Enschede, The Netherlands

and P. DE SALVO

Via S. Quasimodo 59, 00144, Roma, Italy

Abstract. There is need to identify ecosystems that support richer assemblages of biological species in order to preserve habitats and protect the greatest number of species. Remotely sensed data hold tremendous potential for mapping species habitats and indicators of biological diversity, such as species richness. Landscape level habitat analysis using remotely sensed data and Geographical Information Systems (GIS) has the potential to aid in explaining species richness patterns at fine-scale resolutions. We used Landsat Thematic Mapper (TM) image and GIS as well as field data to classify habitat types in the Maasai Mara ecosystem, Kenya. The accuracy of the resulting habitat map was assessed and indices of habitat diversity computed. We then determined the relationship between large mammal species richness and habitat diversity indices, and investigated whether this relationship is sensitive to changes in spatial scale (extent and grain size). Statistical analyses show that species richness is positively correlated with habitat diversity indices and changes of scale in calculations of habitat diversity indices influenced the strength of the correlation. The results demonstrate that mammalian diversity can be predicted from habitat diversity derived from satellite remotely sensed data.

1. Introduction

Biodiversity is commonly used to describe the number, variety and variability of living organisms. As the living world is mostly considered in terms of species, the number of species in a site or habitat (species richness) is commonly used as an indicator of biodiversity (Groombridge 1992). The general perception of conservation biologists and ecologists is that, all things being equal, diverse ecosystems are more important to preserve than low diversity systems (Podolsky 1994). The reason for this is that diverse ecosystems support richer assemblages of biological species than do simple ones (Diamond 1988, Podolsky 1994). However, it must be pointed out that many low diversity ecosystems support rare or endangered species whose protection is also critical (Podolsky *et al.* 1992).

---

*Corresponding author; e-mail: booluoch@Yahoo.com

Mapping species richness and distributions is an important aspect of conservation and land use planning (Spellerberg 1991, Miller and Allen 1994). For example, maps can help identify areas of special biodiversity importance where conservation resources should be focused. Such areas include 'hot spots' of high species richness as well as places where species assemblages of particular interest occur (Cardillo *et al.* 1999). As the current rate of species extinction causes increasing concern (Wilson 1988), land managers and biologists have sought to identify habitats important to the preservation of species diversity (Debinski *et al.* 1999). To identify and conserve areas with high biological importance, remote sensing technology can provide information concerning many variables useful for inventorying, modelling and monitoring species richness (Stoms and Estes 1993). Coupled with GIS, remote sensing can provide information about landscape history, topography, soil, rainfall, temperature and other climatic conditions, as well as about present-day habitat and soil coverage—factors on which the distribution of species depend (Noss 1996). Relationships between species distribution patterns and remotely sensed/GIS data, if known, can be used to predict the distribution of single species or sets of species over large areas (Debinski and Humphrey 1997). However, the utility of remotely sensed data to biological preservation is a function of the extent to which these data correlate with various biological resources (Podolsky *et al.* 1992).

Spatial heterogeneity is one of the most popular hypotheses used to explain patterns of species richness. The heterogeneity hypothesis states that diverse ecosystems support richer assemblages of biological species than simple ecosystems (Diamond 1988). Researchers have examined the spatial configuration of habitat variables to predict species richness (Stoms and Estes 1993). Measures of spatial patterning include a diversity index from information theory based on number and proportions of vegetation or soil types (Miller *et al.* 1989). Species richness has been found to be strongly associated with measures of habitat heterogeneity (Owen 1990).

Ecologists recognize three levels of diversity: *alpha* diversity is a measure of richness within a single homogeneous community; *beta* diversity measures the change in composition along environmental gradients between communities within a landscape; and *gamma* diversity (landscape scale) describes the number of species in a landscape containing more than one community type (Stoms and Estes 1993). Since most management decisions concerning the conservation of species richness are made at landscape scale, it is essential to examine *gamma* diversity (Bohning-Gaese 1997). This study aims to determine the relationship between indices of habitat diversity based on spectral reflectance and large mammal species richness, and to test whether the prediction of species richness by indices of habitat diversity is sensitive to change in spatial scale.

1.1. *Remote sensing of spatial heterogeneity*

Remote sensing is the primary tool for the synoptic analysis of habitats at landscape scale. It allows researchers to address such general questions as (i) what elements are present, (ii) what spatial arrangements these elements have and (iii) what their temporal dynamics are (Quattrochi and Pelletier 1991). Because remote sensing affords the ability to classify habitats based on species composition, structural attributes and phenological differences, and to detect and monitor natural as well as human-induced vegetation dynamics and disturbances, it readily permits the detection and monitoring of spatial heterogeneity (Weishampel *et al.* 1997). Thus, remote sensing has been applied to characterize spatial patterns and processes of vegetation such as the dynamics of biome boundaries (Tucker *et al.* 1985). Landscape ecologists

have also derived or adapted indices of habitat heterogeneity from remotely sensed data products such as land cover or habitat maps (O'Neil *et al.* 1988).

The digital nature of land cover information from satellite imagery enables a potentially large number of landscape metrics to be derived (Haines-Young and Chopping 1996). Jorgensen and Nohr (1996) used Landsat TM derived land cover to compute the Shannon-Wiener and Simpson's indices of diversity for the Ferlo region of Senegal. Areas with the highest landscape diversity were found to support higher numbers of bird species than areas with relatively less landscape diversity. However, some of these indices can be sensitive to spatial resolution and to the number of land cover classes, making generalizations of their relationships to species richness difficult (Stoms and Estes 1993).

2. Methods

2.1. *Study area and animal species data*

Kenya is situated between latitudes $5° 40'$ north and $4° 4'$ south and between longitudes $33° 50'$ and $41° 45'$ east. The study area is the Maasai Mara ecosystem, which is an approximately 7000 km^2 area in Narok district (figure 1). The vegetation varies from grasslands to shrublands to wooded grasslands and riverine forests. The riverine vegetation is spread along major river valleys and covered with discontinuous forests. One of the world famous wildlife sanctuaries, Maasai Mara National reserve, is situated in the Maasai Mara ecosystem. Our study focused on large mammal species because the Maasai Mara ecosystem has a high diversity of mammal species, which attract large numbers of tourists annually.

The large mammal species data (1981–1997) were obtained from the Department of Resource Surveys and Remote Sensing (DRSRS), Ministry of Environment and Natural Resources, Kenya. The systematic reconnaissance flight methodology used by DRSRS for aerial census of animals is well documented (Norton-Griffiths 1978). Statistical analyses to validate DRSRS survey methodology have proved the method and data to be reliable (De Leeuw *et al.* 1998, Ottichilo and Khaemba 2001). Topographic maps of scale 1:250 000 were used for flight planning and all transects conform to the Universal Transverse Mercator (UTM) coordinate system. The aerial surveys were carried out along transects oriented in an east–west direction and spaced at 5 km intervals. The standard flying height and aircraft speed were 120 m and 190 km h^{-1} respectively. Two experienced and well-trained observers occupied the rear seats of a high wing aircraft (Cessna 185 or Partenevia) and counted animals that appeared between two rods attached to the wing struts. The field of vision between these rods was calibrated by flying repeatedly across ground markers of known spacing (Ottichilo and Sinange 1985). The number of animals falling within the survey strips on either side of the aircraft along each 5 km transect segment were counted and recorded into tape recorders by the two rear-seat observers. Groups of animals more than 10 in number were also photographed. After every survey the tape-recorded observations were transcribed to data sheets which, together with processed photographs, were interpreted for animal species using $10 \times$ binocular microscope and overhead projector. Since we carried out the study also at other spatial scales, the processed data at 5 km × 5 km quadrats were converted to 10 km × 10 km, 15 km × 15 km, 20 km × 20 km, 25 km × 25 km and 30 km × 30 km quadrats by spatial aggregation. The analyses focused on large mammal species that are non-migratory. Hence, animal species that were consistently observed over a 17-year period were presumed to be resident to the study area.
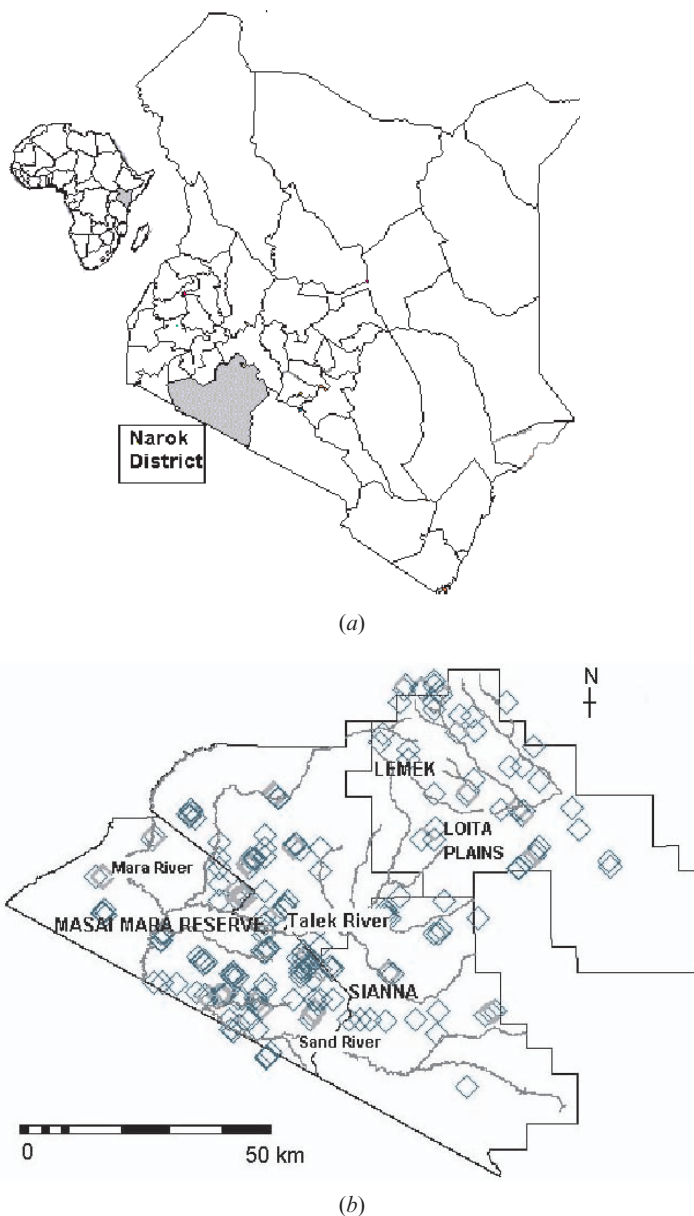
(a)

(b)

Figure 1.   (a) Location of Kenya and Narok district where the study area is situated. (b) The Maasai Mara ecosystem with the major rivers and location of sample sites taken in May–June 1997.

2.2. *GIS and remote sensing analysis*

The methodology for this study was directed toward producing a map of spectrally distinct habitat types as a basis for measuring habitat diversity of the study. Since the study focused on non-migratory large mammal species richness, single date imagery was considered appropriate for measuring habitat diversity. The Landsat 5 TM image (pixel size = 30 m) of the study area for 2 January 1995 was georeferenced to ± 0.5 pixel (15 m) accuracy and resampled to a UTM coordinate system to match

topographic maps of the region. A colour composite image was created to serve as a background image during sampling and subsequent supervised classification. The combination of bands 4, 3 and 2 (red: near infrared band 4; green: red visible band 3; and blue: green visible band 2) at ratio 1:1:1 was the best that allowed the identification of different habitat types in the study area.

An Iterative Self-Organizing Data Analysis (ISODATA) clustering algorithm was applied to the three-band image file to identify spectrally similar pixels. The ISODATA algorithm operates by initially seeding a specified number of cluster centroids in spectral feature space. The euclidean distance between each pixel and each cluster centroid is calculated, and the pixel assigned to a cluster centroid (class). The process of pixel evaluation–centroid recomputation continues iteratively until a threshold percentage (typically 95%) of pixels no longer change cluster centroid assignment. Ten initial clusters were specified for the ISODATA clustering, producing a map of 10 classes, which enabled us to discriminate habitat types present in the study area.

2.3. *Field vegetation mapping*

The field vegetation mapping was carried out from mid-May to mid-June 1997 at the height of the growing season when the vegetation cover was present and adequate. This enabled us to get a good discrimination of the vegetation structure, especially herbaceous layer, in order to prepare a classification scheme that is hierarchical in nature (table 1). A clustered random sampling design (McIntyre 1978) was used to collect data in the field. Computer generated primary cluster points randomly as well as six secondary sample sites, each within a distance of 500 m from the primary cluster. Sample sites were located in the field using Global Positioning System, topographic map (scale 1:250 000) and the Landsat TM 1995 image. Once the sample site was located (figure 1(b)), a quadrat of 20 m × 50 m was established, the height (in metres) and the total cover percentage of each habitat type was visually estimated.

Since multiple layering of vegetation will often result in total land cover values of well over 100%, the Braun-Blanquet scale was used where the range 0–100% was partitioned into five classes (Kent and Coker 1992). The field data were further processed in a spreadsheet and habitat types were categorized into six distinct classes: cultivated/fallow/bare land; dwarf shrubs/short grassland; forest; tall grassland; shrubland; and woody/shrubby grassland (table 1). The field data were divided into training and testing samples (table 2).

2.4. *Supervised classification and accuracy assessment*

Representative or prototype pixels for each of the habitat types were chosen which form training data. The training pixels for each habitat type lay in a training field, 3 × 3 window (nine pixels). The training data estimated the parameters of the classifier algorithm to be used—these parameters are the properties of the probability model used or equations that define partitions in the multispectral space. Using the trained maximum likelihood classifier, every pixel in the image was classified into one of the desired habitat types. A majority filter was applied on the classified image to smooth classification results. An accuracy assessment was then performed on the filtered habitat map (see figure 5).

Test samples were selected from every cluster sampled in the field for each habitat type class (table 2). They were carefully located on the habitat map using the UTM coordinates and since each test sample lies in the centre of 3 × 3 window, the latter

Table 1. Criteria used to classify habitat types from the field data.

Habitat type	Sample size	Woody species	Non-woody species
Cultivated/fallow/bare land	22	–	wheat farms/fallow
Dwarf shrubs/short grassland	103	15–50% ($\leqslant$0.5 m)	>50% ($\leqslant$0.25 m)
Forest	4	>50% (>6 m)	<15%
Tall grassland	39	<15%	>50% (>0.25 m)
Shrubland	16	>50% ($\geqslant$0.5–4 m)	<15%
Woody/shrubby grassland	38	15–50% ($\geqslant$0.5–$\leqslant$6 m)	>50%

Table 2. Sample sizes for each habitat type used for training classifier and testing the accuracy of the supervised classification.

Habitat type	Training	Testing
Cultivated/fallow/bare land	9	13
Dwarf shrubs/short grassland	50	53
Forest	2	2
Tall grassland	18	21
Shrubland	7	9
Woody/shrubby grassland	18	20

was considered to be a homogeneous test area representing a single habitat type class. Therefore, every classified pixel in the window was checked for correctness by comparing image class and the field habitat type class in the reference data. From the field checked data an error matrix table was constructed (see table 5).

2.5. *Calculation of the TM-based habitat diversity index and species richness*

The large mammal species richness and habitat diversity index were estimated in six quadrats of different sizes, 5 km × 5 km, 10 km × 10 km, 15 km × 15 km, 20 km × 20 km, 25 km × 25 km and 30 km × 30 km. The size of the smallest quadrat (5 km × 5 km) was determined by the aerial sampling method used by DRSRS and described in §2.1. In each quadrat the number of species was counted to give a value for total species richness. Prior to calculation of habitat diversity indices for these six quadrats, two point maps were created for each quadrat. The two point maps were then rasterized to two different pixel sizes (25 m and 75 m) with appropriate point sizes (table 3). For example, point size 200 means that each sample unit in a

Table 3. The quadrat sizes (km) with corresponding distance (m) between quadrats, point sizes and total number of pixels for pixel sizes 25 m and 75 m in a quadrat.

Quadrat	Distance	Point size		Total pixels	
		25 m	75 m	25 m	75 m
5 × 5	5000	200	67	40 000	4489
10 × 10	10 000	400	133	160 000	17 689
15 × 15	15 000	600	200	360 000	40 000
20 × 20	20 000	800	267	640 000	71 289
25 × 25	25 000	1000	333	1 000 000	110 889
30 × 30	30 000	1200	400	1 440 000	160 000

point map will be represented by 200×200 pixels in the output raster map. Thus, all 40 000 pixels had the same identity as the sample unit.

Integrating habitat map (pixel size = 30 m) and quadrat raster maps (pixel size = 25 m or 75 m) requires compatible pixel sizes. By using the Nearest Neighbour resampling method, the habitat map was resampled twice to pixel sizes of 25 m and 75 m and then crossed with each of the six quadrat raster maps of 25 m and 75 m pixel sizes respectively. Table 4 shows an example of the resulting output table for an individual sample unit.

The habitat diversity was calculated for every sample unit based on habitat map of fine spatial scale (pixel size = 25 m) and relatively coarse spatial scale (pixel size = 75 m). The two commonly used diversity indices for quantifying landscape structure (Haines-Young and Chopping 1996); Shannon-Wiener index (H') and Simpson's index (D) were employed:

$$H' = -\Sigma p_i \ln p_i \tag{1}$$

$$D = 1/\Sigma p_i^2 \tag{2}$$

where p_i represents the fractional abundance of each habitat type in a quadrat (i.e. number of pixels of specific habitat type i divided by total number of pixels in a quadrat).

The Shannon-Wiener index is based on information theory that tries to measure the amount of uncertainty (Krebs 1989) in every quadrat. Thus, measuring the amount of uncertainty in a quadrat can provide a measure of diversity. Hence, a quadrat with only one habitat type has no uncertainty (no diversity) in it, $H' = 0$ so the larger the value of H', the greater the uncertainty (diversity); whereas Simpson's index is based on the probability that two habitat types are similar (Simpson 1949).

Regression lines between the dependent variables (species richness) and the independent variables (Shannon and Simpson's habitat diversity indices) were calculated along with 95% confidence interval for different quadrat sizes. The correlation coefficients for the relationships between large mammal species richness and habitat diversity indices calculated at two different pixel sizes were compared at different quadrat sizes. In addition, the relationship between the number of classes in the image classification and habitat diversity indices was calculated.

3. Results

Table 5 shows the error matrix resulting from classifying training set pixels and testing the accuracy. The overall accuracy is computed by dividing the total correct

Table 4. Cross table resulting from crossing a quadrat (15 km × 15 km) raster map and habitat map of pixel size 25 m for a single sample unit. The number of pixels was used to compute habitat diversity using Shannon-Wiener and Simpson's indices.

Sample unit	Habitat type	No. of pixels
60095	Woody/shrubby grassland	212 125
60095	Forest	18 860
60095	Dwarf shrubs/short grassland	69 191
60095	Tall grassland	53 095
60095	Shrubland	2872
60095	Cultivated/fallow/bare lands	3857
Total		360 000

sum of the major diagonal by the total number of pixels in the error matrix (Jensen 1996). The producer accuracy indicates the probability of a reference pixel being correctly classified and it is calculated by dividing diagonal value for each category by its column total. Conversely, user accuracy is the probability that a pixel classified on the map actually represents that category on the ground and it is calculated by dividing diagonal value for each category by its row total (Story and Congalton 1986).

KAPPA analysis is a discrete multivariate technique of use in accuracy assessment (Congalton and Mead 1983). It yields a KHAT statistic (an estimate of KAPPA) that is a measure of agreement between image data and the reference data or accuracy (Congalton 1991). The KHAT statistic is computed as:

$$K_{hat} = N \sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r} (x_{ri} \times x_{ci})/N^2 - \sum_{i=1}^{r} (x_{ri} \times x_{ci}) \tag{3}$$

where r is the number of rows in the matrix, x_{ii} is the number of observations in row i and column i, and x_{ri} and x_{ci} are the marginal totals for row i and column i, respectively, and N is the total number of observations. The overall classification accuracy is 89%, while the KHAT accuracy is 87%. The results are different because the two measures incorporated different information. The overall accuracy only incorporated the major diagonal and excluded the omission and commission errors. By contrast, KHAT accuracy computation incorporated the off-diagonal elements as a product of the row and column marginals.

Figure 2 shows the relationship between number of habitat type classes in the image classification and Shannon habitat diversity measure derived from habitat map of two different pixel sizes, 25 m and 75 m. The Shannon habitat diversity index based on habitat map of pixel size 25 m has a wider range of values (figure 2(*a*)) than that derived from habitat map of pixel size 75 m. As the number of classes in the image classification increases, the value of the habitat diversity index increases. This indicates that the habitat diversity index is sensitive to the presence of rare habitat types, which is consistent with the findings of Haines-Young and Chopping (1996).

The habitat diversity indices (Shannon-Wiener and Simpson) based on habitat

Table 5. Error matrix of the classification map derived from Landsat TM data of the Maasai Mara ecosystem. Diagonal elements represent correctly classified pixels and are measured by overall accuracy (OA). All non-diagonal elements represent errors of omission and commission, which are measured by producer accuracy (PA) and user accuracy (UA) respectively.

Classification	Cultl	Dshg	For	Gra	Shgr	Shbl	Total	UA (%)
Cultl	100	36	0	0	0	0	136	74
Dshg	8	254	0	9	6	0	277	91
For	0	0	30	0	1	0	31	96
Gra	0	6	9	292	1	0	308	94
Shgr	0	0	4	1	110	11	126	87
Shbl	0	0	5	13	0	61	79	77
Total	108	296	48	315	118	72	957	
PA (%)	93	85	62	93	93	85	OA	89

Habitat types classified: cultivated/fallow/bare land (Cultl); dwarf shrubs/short grassland (Dshg); forest (For); tall grassland (Gra); woody/shrubby grassland (Shgr); shrubland (Shbl).
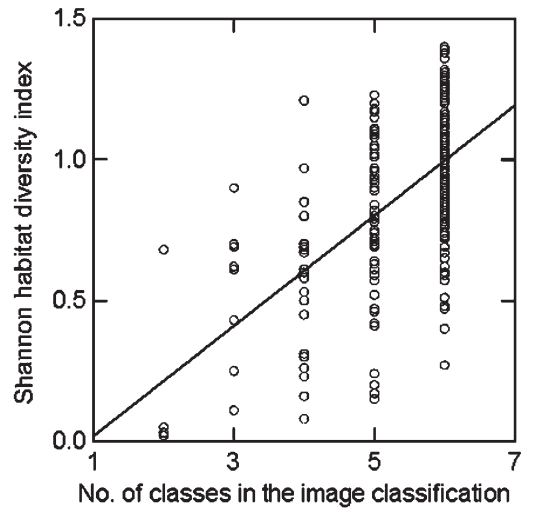
(*a*)
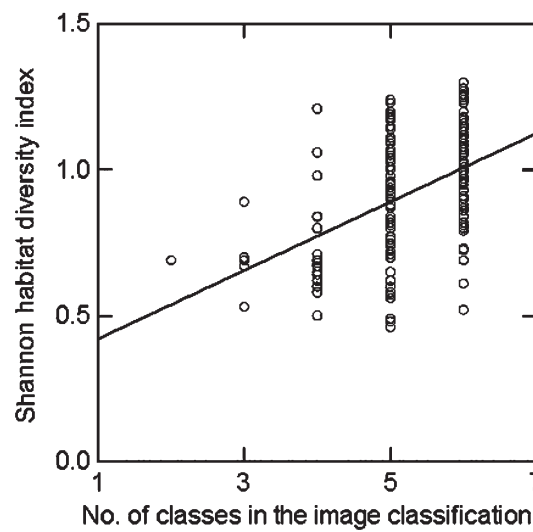
(*b*)

Figure 2. The influence of the number of habitat type classes in the image classification on Shannon habitat diversity measure based on (*a*) habitat map of 25 m pixel size ($r^2 = 0.365$), and (*b*) habitat map of 75 m pixel size ($r^2 = 0.237$) at quadrat size, 5 km × 5 km. The correlation coefficients are significant at $p < 0.05$. The straight lines were fitted because residuals tend not to vary in a systematic fashion between positive and negative.

map of pixel size 25 m have significant correlation with mammalian species richness at all spatial scales (table 6). On the other hand, habitat diversity indices based on habitat map of pixel size 75 m do not have significant correlation at quadrat sizes, 25 km × 25 km and 30 km × 30 km (table 6). Generally, Shannon-Wiener habitat diversity index has a stronger correlation with species richness than Simpson's habitat diversity index. At quadrat sizes 5 km × 5 km and 10 km × 10 km, habitat diversity

Table 6. Coefficient of correlation (r^2) for the relationships between large mammal species richness and habitat diversity indices (Shannon-Wiener and Simpson) based on habitat map at 25 m and 75 m pixel sizes at different quadrat sizes (km).

Quadrat	Shannon		Simpson		
	25 m	75 m	25 m	75 m	*n*
5 × 5	0.119	0.104	0.090	0.087	182
10 × 10	0.554	0.552	0.494	0.492	43
15 × 15	0.854	0.315	0.831	0.258	21
20 × 20	0.862	0.618	0.819	0.415	13
25 × 25	0.855	ns	0.810	ns	8
30 × 30	0.833	ns	0.775	ns	7

ns = not significant at $p < 0.05$; n = sample size.

indices based on habitat map of pixel sizes 25 m and 75 m have more or less similar correlation with species richness. However, at quadrat sizes 15 km × 15 km and 20 km × 20 km habitat diversity indices based on habitat map of pixel size 25 m have a stronger correlation with species richness than habitat diversity indices derived from habitat map of pixel size 75 m (table 6).

The coefficient of correlation is so low for the base quadrat size (5 km × 5 km) because of the presence of few outlier quadrats, which have a relatively high number of species but low habitat diversity (figure 3(*a*)). However, the strength of correlation increases as quadrat size increases (table 6) because the outlier quadrats (5 km × 5 km) are eliminated with increasing spatial aggregation. The highest correlation between species richness and habitat diversity indices based on habitat map of pixel sizes 25 m and 75 m were obtained at intermediate quadrat size (20 km × 20 km) with r^2 values of 0.862 and 0.618 respectively (table 6).

Figure 3 shows that mammalian species richness increases with increase in habitat diversity at 5 km × 5 km and 20 km × 20 km quadrats which account for 12% and 86% respectively of the observed variation of species richness. When the coefficient of correlation between species richness and habitat diversity basemaps (pixel size = 25), with different resolutions (i.e. side length of squares used for habitat diversity calculations) are plotted, the coefficient of correlation increases exponentially with increase in side length of squares and levels off at a side length of 15 000 m (360 000 pixels) (figure 4).

Figure 6(*a*) shows that the highest habitat diversity is associated with two ecological units of the Maasai Mara ecosystem, namely, Mara (including Mara Reserve) and Sianna (figure 5). On the other hand, low values of habitat diversity are found in the northern part of the ecosystem (i.e. Loita region), associated with areas covered mainly with cultivated/fallow/bare land and dwarf shrubs/short grassland (figure 5). Figure 6(*b*) shows that large mammal species richness is relatively high in the southern parts of the Maasai Mara ecosystem, which have diverse habitats (figure 5).

4. Discussion

4.1. *Relationship between habitat diversity and species richness*

An interesting question is: which features in the Maasai Mara ecosystem give high habitat diversity? The major rivers, namely, Mara, Talek and Sand (figure 1(*b*)) and their tributaries, drain Mara (including Mara reserve) and Sianna, hence there
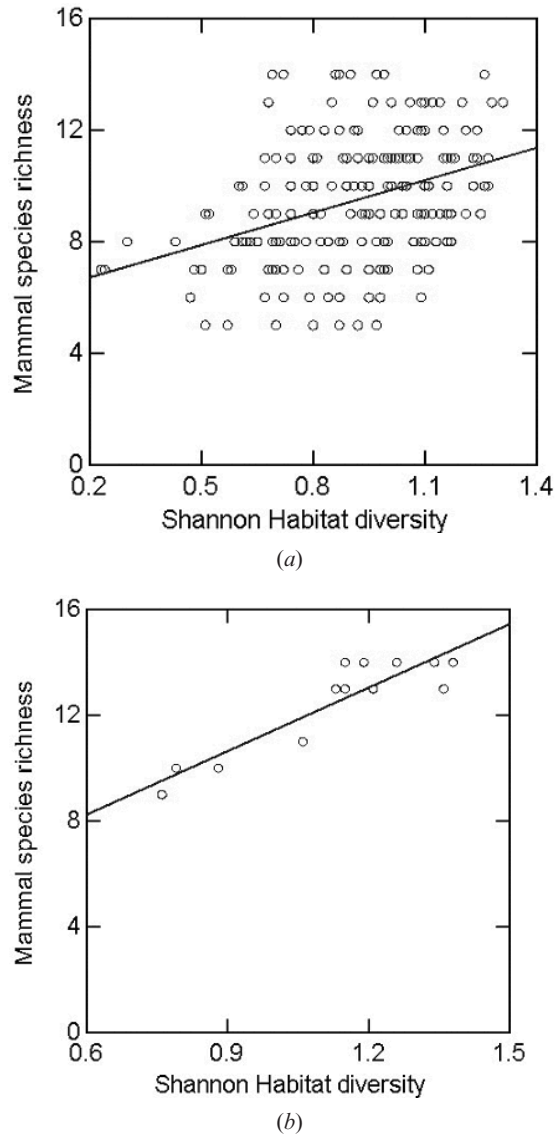
(a)

(b)

Figure 3.   Scatterplot of large mammal species richness (*S*) vs Shannon habitat diversity index (*H'*) based on habitat map of pixel size 25 m calculated at quadrat sizes (*a*) 5 km × 5 km (*S* = 5.968 + 3.865*H'*), and (*b*) 20 km × 20 km (*S* = 3.095 + 8.375*H'*). The least-squares fit for the relations in both (*a*) and (*b*) were fitted with straight lines because residuals tend not to vary in a systematic fashion between positive and negative.

is a variety of habitat types such as riverine forests, tall grassland, dwarf shrubs/ short grassland, shrubland and woody/shrubby grassland (figure 5). In addition, there are no significant human activities that may cause large-scale habitat destruction. Consequently, Mara and Sianna regions support greater numbers of large mammal species (compare figures 5 and 6(*b*)). By contrast, the Loita region has a low number of species probably due to extensive habitat destruction caused by mainly large-scale wheat farming. As a consequence, the most common habitat types
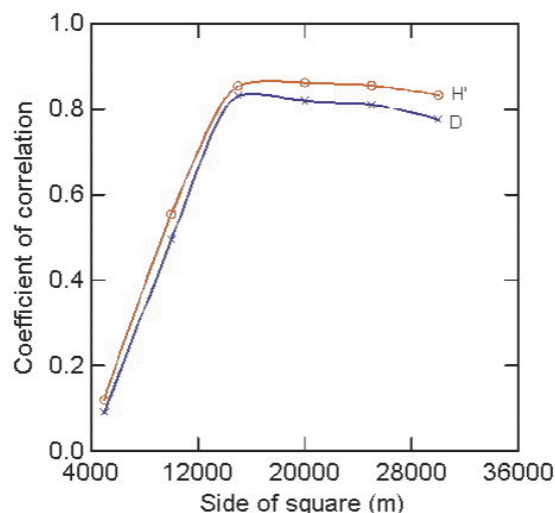
Figure 4. The change in coefficient of correlation (r^2) between the number of species and different side lengths in squares used for calculation of habitat diversity index based on Shannon-Wiener (H') and Simpson's indices (D).
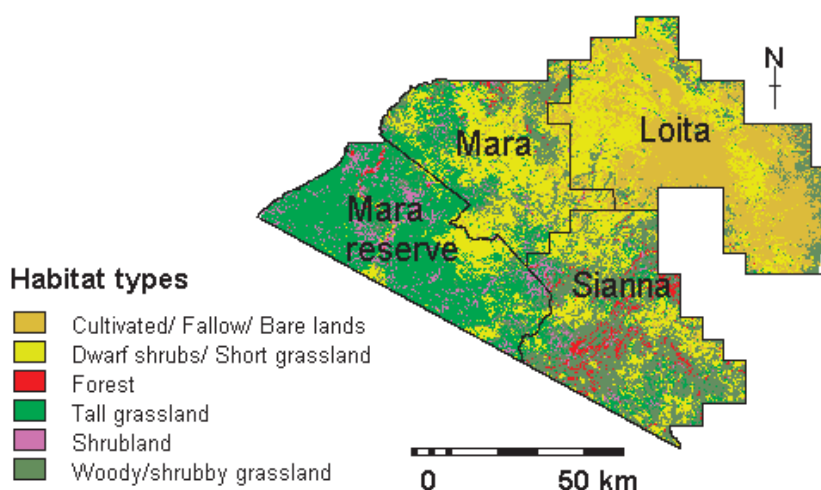

Figure 5. Habitat map based on Landsat TM of January 1995—six categories of habitat types important for mammal species were classified.

are dwarf shrubs/short grassland and cultivated/fallow/bare land (figure 5). The results (table 6) show that large mammal species richness is positively correlated with habitat diversity, confirming that highly diverse habitats are endowed with more species (figure 3). So, why does higher habitat diversity increase the species richness of mammals? One possibility is that a particular species tends to occur only in certain habitats and not others. Thus, as one proceeds along a habitat gradient, one accumulates more and more species, and the accumulated number of species increases with the diversity of habitats encountered (Diamond 1988).
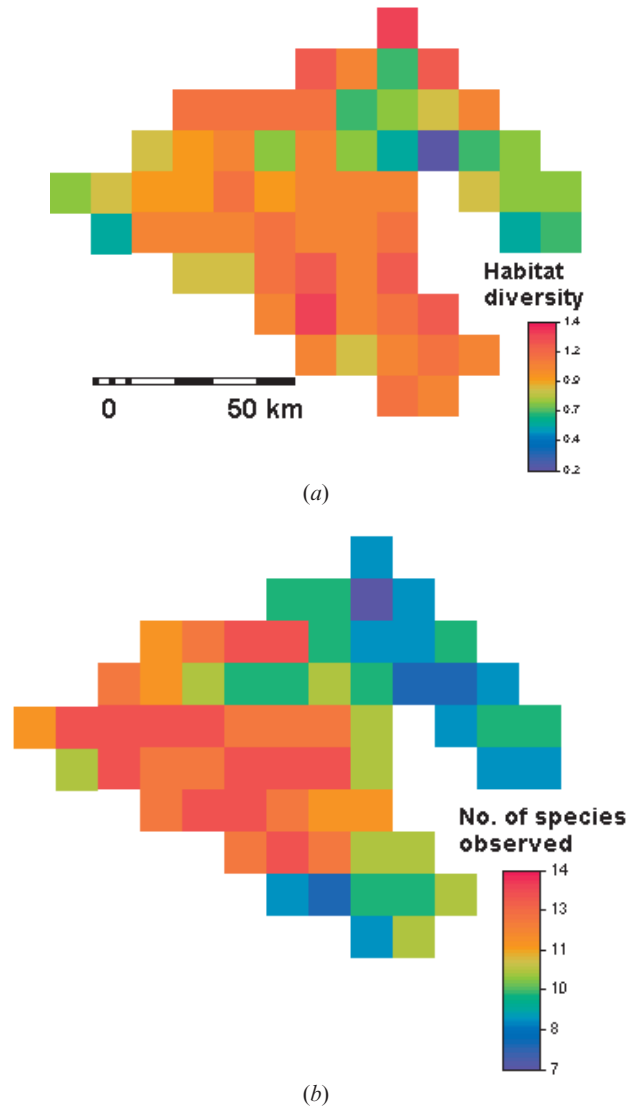
(*a*)

(*b*)

Figure 6.   (*a*) Habitat diversity map derived by computing Shannon-Wiener index at 10 km × 10 km quadrat of the habitat map (pixel size = 25). (*b*) Spatial distribution of the number of large mammal species observed at 10 km × 10 km quadrat size.

4.2. *Habitat diversity–species richness relationship and scale*

The main prerequisite for using satellite images for mapping of biodiversity is that habitat features important to higher animals, in this case large mammal species, can be detected on the images. This is mainly a question about scale of study—extent as well as grain (Allen and Starr 1982). Extent refers to the size of the study area investigated, while grain is the resolution of the remote sensor—radiometric, spatial, spectral and temporal. As extent increases, the level of detail (grain) that can be maintained, given constraints on time, effort and money, will decrease, and vice versa. The amount of information that can be retrieved on numbers of habitat types depends on these factors (Nagendra 2001). In view of these, the scale of the satellite

image is important when analysing habitat diversity in relation to species richness (Jorgensen and Nohr 1996).

Our results (table 6) show that the changes of scale in estimation of habitat diversity influence the correlation with large mammal species richness. The latter has a stronger correlation with habitat diversity indices derived from habitat map of fine spatial resolution (25 m pixel size) at quadrat sizes of 15, 20, 25 and 30 km resolutions than relatively coarse resolution (75 m pixel size). Apparently, the choice of raster map pixel size (grain) has important implications for diversity indices that depend on the number of habitat types (figure 2). The loss of information associated with relatively coarser resolution (75 m pixel size) particularly at quadrat sizes of 15, 20, 25 and 30 km resolutions leads to the disappearance of habitat types represented by the least extensive and least compact patches (Haines-Young and Chopping 1996). Consequently, Shannon-Wiener and Simpson's habitat diversity indices calculated from reduced number of habitat types have a narrow range of values (figure 2(b)) that weakens the strength of correlation with species richness.

The strengths of association between species richness and habitat diversity indices derived from habitat maps of pixel sizes 25 m and 75 m increase as quadrat size increases and peak at intermediate scale (20 km × 20 km) then decline. Presumably, as quadrat size increases the species richness is increased by the presence of uncommon habitat types such as riverine forests and shrubland with unique set of species. However, as the quadrat size increases further, the number of habitat types mapped within quadrats tends to decrease with corresponding decline in number of species predicted. Because small patches are no longer mapped due to generalization, some habitat types and their associated species are no longer predicted to occur in a given quadrat (Stoms 1992). Figure 4 illustrates the correlation between habitat diversity calculated with different sizes of squares and number of species observed. The curves level off at side length of square 15 000 m (15 km × 15 km)—this may be recommended as the minimal area for assessing the relationship between Landsat TM image (25 m pixel size) derived habitat diversity and large mammal species richness in the study area.

4.3. *Habitat map classification and accuracy assessment*

In ecological systems, local heterogeneity can play a comparably central role in regulating stability and diversity (Weishampel *et al.* 1997). Heterogeneity among patches is believed to generally increase species coexistence (Czárán and Bartha 1992). Differences among climates, soils, hydrological conditions and disturbance histories are thought to produce repeatable differences among vegetation patches (McIntosh 1985). This view provides a rationale for vegetation (habitat) classification schemes and for the generation of thematic maps showing the spatial disposition of areas of different habitat types (Weishampel *et al.* 1997). To generate signatures that accurately represent habitat types during supervised classification of the Landsat TM image, training samples were repeatedly selected, and the signatures generated from samples were evaluated or manipulated by merging or deleting. The spectral separability between the datasets was studied and the training sets that were overlapping were either merged or deleted.

A classification is not complete until it has been assessed and that is when the decisions made based on that information have any validity. It is obvious that in order to adequately to assess the accuracy of the remotely sensed classification, accurate ground or reference data must be collected. In addition, selection of the proper sampling scheme is absolutely critical to generating an error matrix that is representative of the entire classified image (Congalton 1991). During the field

surveys the choice of sampling scheme was influenced by cost and time effectiveness, good sample distribution throughout the study area and accessibility of the terrain. The clustered random sampling was employed which involved clustering of mapping units which ensured survey efficiency. The clusters were distributed throughout the accessible parts of the study area (figure 1(*b*)) so that all habitat types were sampled. Multiple samples for all habitat types were taken to ensure evenness of the spread of observations over the whole of the Maasai Mara ecosystem. Moreover, the sample parameters were accurately estimated. Due to inaccessibility fewer samples were taken in categories such as forest and shrubland (table 1). Although Congalton (1991) suggested that a minimum of 50 samples for each habitat type category in the error matrix should be collected, we failed to reach the minimum samples for all categories, except dwarf shrubs/short grassland (table 1), because of inaccessibility, time and money constraints. Since sufficient effort should also be given to the classification scheme to be used (Congalton 1991), we used a classification scheme that is hierarchical in nature and included every habitat type in the study area (table 1).

The overall classification accuracy is 89%, which is above the 85% proposed by Anderson *et al.* (1976). However, such a non-site-specific accuracy assessment yields very high accuracy but misleading results when all the errors balance out in a region (Jensen 1996). To perform classification accuracy assessment correctly, it is necessary to compare two sources of information: the remote sensing derived classification map and reference test information. Therefore, an error matrix (table 5) is a standard method to represent accuracy because the accuracy of each category is clearly described, along with both the errors of inclusion (commission errors) and errors of exclusion (omission errors). If we were primarily interested in the ability to classify just cultivated/fallow/bare land, the producer's accuracy of this category was 93%, which is quite good. However, only 74% (user's accuracy) of the areas called cultivated/fallow/bare land are actually cultivated/fallow/bare land. A careful evaluation of the error matrix reveals that there was confusion when discriminating cultivated/fallow/bare land from dwarf shrubs/short grassland. Therefore, although the producer of this map can claim that 93% of the time an area that was cultivated/fallow/bare land was identified as such, a user of this map will find that only 74% of the time an area is visited in the field using the map will it actually be cultivated/fallow/bare land. On the other hand, the producer's accuracy of forest is only 62% because there was confusion when discriminating forest from tall grassland, woody/shrubby grassland and shrubland as shown in the error matrix (table 5). Even though the producer of this map can claim that only 62% of the time an area that was forest was identified as such, a user of this map will find that 96% of the time an area is visited in the field using the map will it actually be forest.

5.   Conclusion

The study has provided an assessment of high resolution Landsat TM image for the purpose of predicting and mapping large mammal species richness. The results underscore the importance of spatial scales (extent and grain size) in understanding the relationship between species richness and habitat diversity indices derived using spectral reflectance. Changes of spatial scales in calculations of habitat diversity indices influence the strength of correlation with species richness. However, the influence does not change the direction of the correlation between habitat diversity and species richness.

The study has practical implications for the use of habitat diversity indices derived

from remotely sensed data for predicting large mammal species richness. The diversity indices (Shannon-Wiener and Simpson's) are sensitive both to the number of habitat types and to grain size, so care must be exercised when defining the nature of the landscape mosaic which forms the object of study (Haines-Young and Chopping 1996). Moreover, the strength of association between species richness and habitat diversity indices varies at different spatial scales. Hence, before starting to manage for high levels of species richness it is necessary to determine at which spatial scale species richness and habitat diversity should be protected. The study further reveals that the association between species richness and habitat diversity indices is strongest at a relatively intermediate scale (20 km × 20 km). Thus, protecting the diversity of habitats at this spatial scale may ensure high levels of species richness in the study area.

Acknowledgments

References

ALLEN, T. F. H., and STARR, T. B., 1992, *Hierarchy: Perspectives for Ecological Complexity* (Chicago: University of Chicago Press).

ANDERSON, J. R., HARDY, E. E., ROACH, J. T., and WITMER, R. E., 1976, A land use and land cover classification system for use with remote sensor data. Geological survey professional paper 964, USA Government Printing Office, Washington, DC, USA.

BOHNING-GAESE, K., 1997, Determinants of avian species richness at different spatial scales. *Journal of Biogeography*, **24**, 49–60.

CARDILLO, M., MACDONALD, D. W., and RUSHTON, S. P., 1999, Predicting mammal species richness and distributions: testing the effectiveness of satellite-derived land cover data. *Landscape Ecology*, **14**, 423–435.

CONGALTON, R. G., 1991, A review of assessing the accuracy of classification of remotely sensed data. *Remote Sensing of Environment*, **37**, 35–46.

CONGALTON, R. G., and MEAD, R. A., 1983, A quantitative method to test for consistency and correctness in photointerpretation. *Photogrammetric Engineering and Remote Sensing*, **49**, 69–74.

CZÁRÁN, T., and BARTHA, S., 1992, Spatiotemporal dynamic models of plant populations and communities. *Trends in Ecology and Evolution*, **7**, 38–42.

DEBINSKI, D. M., and HUMPHREY, P. S., 1997, An integrated approach to biological diversity assessment. *Natural Areas Journal*, **17**, 355–365.

DEBINSKI, D. M., KINDSCHER, K., and JAKUBAUSKAS, M. E., 1999, A remote sensing and GIS-based model of habitats and biodiversity in the Greater Yellowstone Ecosystem. *International Journal of Remote Sensing*, **20**, 3281–3291.

DE LEEUW, J., PRINS, H. H. T., NJUGUNA, E. C., SAID, M. Y., and DE BY, R., 1998, Interpretation of DRSRS animal counts (1977–1997) in the rangeland districts of Kenya. Ministry of Planning and National Development, Nairobi, Kenya.

DIAMOND, J., 1988, Factors controlling species diversity: overview and synthesis. *Annals of the Missouri Botanical Garden*, **75**, 117–129.

GROOMBRIDGE, B., 1992, Status of the Earth's Living Resources. Global Biodiversity (London: World Conservation Monitoring Centre).

HAINES-YOUNG, R., and CHOPPING, M., 1996, Quantifying landscape structure: a review of landscape indices and their application to forested landscapes. *Progress in Physical Geography*, **20**, 418–445.

JENSEN, J. R., 1996, *Introductory Digital Image Processing: A Remote Sensing Perspective*, 2nd edn (New Jersey: Prentice Hall).

JORGENSEN, A. F., and NOHR, H., 1996, The use of satellite images for mapping of landscape and biological diversity in Sahel. *International Journal of Remote Sensing*, **17**, 91–109.

KENT, M., and COKER, P., 1992, *Vegetation Description and Analysis: A Practical Approach* (New York: John Wiley & Sons).

KREBS, C. J., 1989, *Ecological Methodology* (New York: Harper Collins).

McINTOSH, R. P., 1985, *The Background of Ecology: Concept and Theory* (Cambridge: Cambridge University Press).

McINTYRE, G. A., 1978, Statistical aspects of vegetation sampling. In *Measurements of Grassland Vegetation and Animal Production*, edited by L.'t Mannetje (England: Commonwealth Agricultural Bureau), pp. 8–21.

MILLER, R. I., and ALLEN, J. H., 1994, Mapping the elements of biodiversity: the rare species of Madagascar. In *Mapping the Diversity of Nature*, edited by R. I. Miller (London: Chapman & Hall), pp. 37–50.

MILLER, R. I., STUART, S. N., and HOWELL, K. M., 1989, A methodology for analyzing rare species distribution patterns utilizing GIS technology: the rare birds of Tanzania. *Landscape Ecology*, **2**, 173–189.

NAGENDRA, H., 2001, Using remote sensing to assess biodiversity. *International Journal of Remote Sensing*, **22**, 2377–2400.

NORTON-GRIFFITHS, M., 1978, *Counting Animals*. Handbook No. 1, 2nd edn (Nairobi: Africa Wildlife Leadership Foundation).

NOSS, R. F., 1996, Ecosystems as conservation targets. *Trends in Ecology and Evolution*, **11**, 351.

O'NEIL, R. V., KRUMMEL, J. R., GARDENER, R. H., SUGIHARA, G., JACKSON, B., DEANGELIS, D. L., MILNE, B. T., TURNER, M. G., ZYGMUNT, B., CHRISTENSEN, S. W., DALE, V. H., and GRAHAM, R. L., 1988, Indices of landscape pattern. *Landscape Ecology*, **1**, 153–162.

OTTICHILO, W. K., and KHAEMBA, W., 2001, Validation of observer and aircraft calibration for aerial animal surveys: a case of the Department of Resource Surveys and Remote Sensing (DRSRS), Kenya. *African Journal of Ecology*, **39**, 45–50.

OTTICHILO, W. K., and SINANGE, R. K., 1985, Differences in the visual and photographic measurements in the estimation of strip widths for aerial censuses of animal populations. DRSRS, Ministry of Planning and National Development, Nairobi, Kenya.

OWEN, J. G., 1990, Patterns of mammalian species richness in relation to temperature, productivity, and variance in elevation. *Journal of Mammalogy*, **71**, 1–13.

PODOLSKY, R., 1994, Ecological hot spots. A method for estimating biodiversity directly from digital Earth imagery. *Earth Observation Magazine*, June, 30–36.

PODOLSKY, R., FREILICH, J., and KNEHR, R., 1992, Predicting plant species richness from remotely sensed data in a high desert ecosystem. *Technical Papers of the 1992 Annual Meeting of ASPRS/ACSM/RT, Washington, DC, 3–8 August 1992. Vol. 4, Remote sensing and data acquisition* (Bethesda, MD: American Society for Photogrammetry and Remote Sensing), pp. 57–63.

QUATTROCHI, D. A., and PELLETIER, R. E., 1991, Remote sensing for analysis of landscapes: an introduction. In *Quantitative Methods in Landscape Ecology*, edited by M. G. Turner and R. H. Gardner (New York: Springer-Verlag), pp. 51–76.

SIMPSON, E. H., 1949, Measurement of diversity. *Nature*, **163**, 688.

SPELLERBERG, I. F., 1991, *Monitoring Ecological Change* (Cambridge: Cambridge University Press).

STOMS, D. M., 1992, Effects of habitat map generalization in biodiversity assessment. *Photogrammetric Engineering and Remote Sensing*, **58**, 1587–1591.

STOMS, D. M., and ESTES, J. E., 1993, A remote sensing research agenda for mapping and monitoring biodiversity. *International Journal of Remote Sensing*, **14**, 1839–1860.

STORY, M., and CONGALTON, R., 1986, Accuracy assessment: a user's perspective. *Photogrammetric Engineering and Remote Sensing*, **52**, 397–399.

TUCKER, C. J., TOWSHEND, J. R. G., and GOFF, T. E., 1985, African land-cover classification using satellite data. *Science*, **227**, 369–375.

WEISHAMPEL, J. F., KNOX, R. G., RANSON, K. J., WILLIAMS, D. L., and SMITH, J. A., 1997, Integrating remotely sensed spatial heterogeneity with a three-dimensional forest succession model. In *The Use of Remote Sensing in the Modeling of Forest Productivity*, edited by H. L. Gholz, K. Nakane and H. Shimoda (Dordrecht: Kluwer Academic Publishers), pp. 109–134.

WILSON, E. O., 1988, *Biodiversity* (Washington, DC, USA: National Academy Press).