

# Prediction of lichen diversity in an UNESCO biosphere reserve – correlation of high resolution remote sensing data with field samples

Lars T. Waser · Meinrad Kuechler · Markus Schwarz ·  
Eva Ivits · Silvia Stofer · Christoph Scheidegger

Received: 31 October 2004 / Revised: 1 February 2006 / Accepted: 1 May 2006 / Published online: 22 December 2006  
© Springer Science + Business Media B.V. 2006

**Abstract** The present study focuses on developing models to predict lichen species richness in a UNESCO Biosphere Reserve of the Swiss Pre-Alps following a gradient of land-use intensity combining remote sensing data and regression models. The predictive power of the models and the obtained  $r$  ranging from 0.5 for lichens on soil to 0.8 for lichens on trees can be regarded as satisfactory to good, respectively. The study revealed that a combination of airborne and spaceborne remote sensing data produced a variety of ecological meaningful variables.

**Keywords** biodiversity · biosphere reserve · hotspot · lichens · linear regression models · CIR orthoimages · remote sensing

## 1 Introduction

The alpine environment and ecosystems are exposed to both natural and anthropogenetic threats [7, 44]. The pressure on alpine environment [6] arising from past development, tourism and possible climate change [12, 14, 23, 45] is far greater than on other types environments [36]. Inventory of biodiversity and monitoring efficacy of measures for its conservation have emerged as important scientific challenges in recent years [19, 28]. Because it is almost impossible to have a complete biodiversity survey at the regional scale [33, 47], methods for extrapolations are needed. These methods provide information that is remotely similar to field samples and which would allow to considerably reduce extensive field surveys [8].

To date, regression analyses in particular have been broadly applied in ecology [14, 22]. Some studies [15, 43] reveal that use of modern regression approaches has proven particularly useful for modeling spatial distribution of species and communities. Thus, in combination with regression analyses, airborne and spaceborne remote sensing data may help in the assessment of biodiversity in a mountainous region. Estimates of species richness within a region can then be used to focus on targets in inventories so that appropriate levels of sampling can be reached in these areas. Calculation of potential biodiversity hotspots might be helpful for conservation efforts particularly in mountainous regions, e.g., for an assessment of the landscape itself and for future protection planning.

This study focused on an assessment of lichen species richness for six test sites within the UNESCO Biosphere Reserve *Entlebuch* in the Swiss Pre-Alps following a gradient of land-use intensity combining remote sensing techniques and regression analyses. This study ties in with

---

L. T. Waser (✉) · M. Kuechler · M. Schwarz ·  
S. Stofer · C. Scheidegger  
Department of Landscape Inventories,  
Swiss Federal Research Institute WSL,  
Zuercherstrasse 111,  
CH-8903 Birmensdorf, Switzerland  
e-mail: lars.waser@wsl.ch

M. Kuechler  
e-mail: meinrad.kuechler@wsl.ch

M. Schwarz  
e-mail: markus.schwarz@wsl.ch

S. Stofer  
e-mail: silvia.stofer@wsl.ch

E. Ivits  
Department of Remote Sensing and LIS, Freiburg University,  
Tennenbacherstr. 4, 70106 Freiburg, Germany

the European Union Project *BioAssess*, which is aimed at quantifying patterns in biodiversity and developing “Biodiversity Assessment Tools” that can be used to rapidly assess biodiversity. For this project, seven biological indicators (soil macrofauna, collembola, ground beetles, plants, butterflies, birds and lichens) as well as remote-sensing-based indicators (non-biological) for a biodiversity assessment were collected in the test sites for eight participating countries (for more details, see the BioAssess homepage <http://www.nbu.ac.uk/bioassess/>).

Lichens are mutualistic symbiotic organisms and consist of two unrelated components: a fungus (the mycobiont) and one or more algae or cyanobacteria (the photobionts). Many species have evolved a requirement for substrates that are themselves by-products of advanced succession in more dominant ecosystems. Lichens are affected by various forms of anthropogenic disturbance such as agriculture, pasture and forest management [39], atmospheric pollution and climate change [29, 31]. These disturbances can be detected by using remote sensing data and ecological modeling. Some studies show the combination of lichens with remote sensing methods: e.g., lichens have been particularly used as an indicator of ecosystem disturbance [32], serve as indicators of forest age [3, 37], or are involved in remotely sensed evaluation of net ecosystem productivity [20]. In other studies, predictivity of ecological values was tested using lichen relevés [30] or lichen diversity has been predicted using stand characteristics in CIR aerial photographs [3].

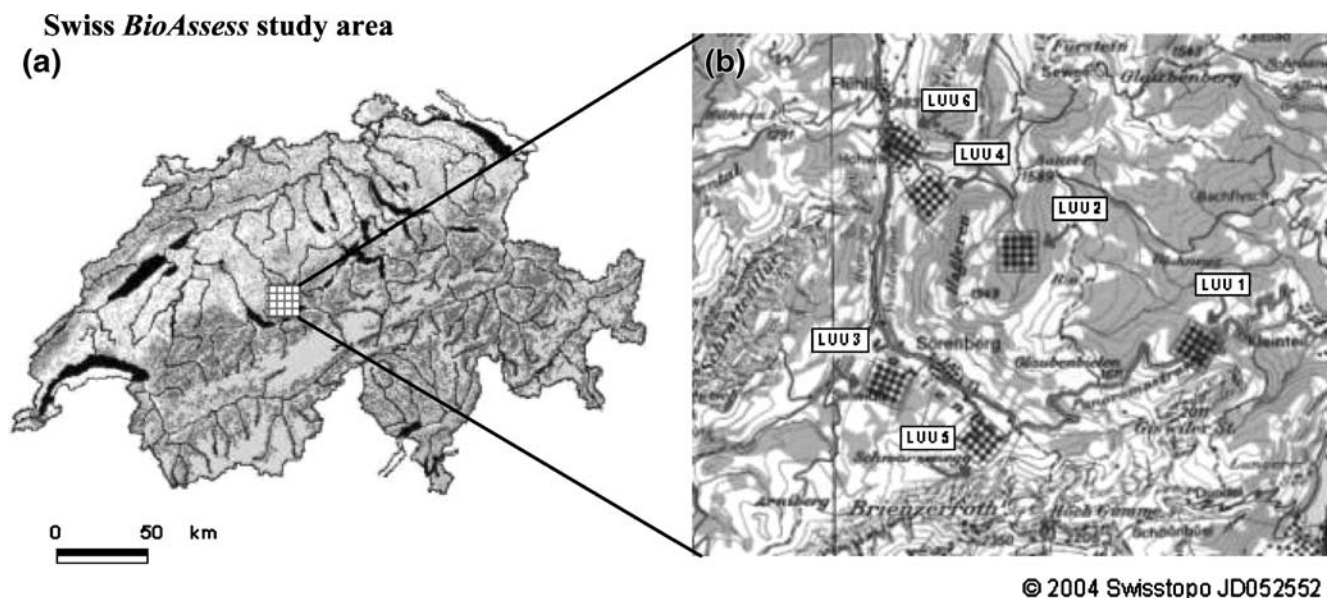
The objectives of this study were to extract and correlate ecological meaningful variables derived from both airborne and spaceborne remote sensing data with lichen species

richness of the field surveys and to build regression models to predict lichen diversity on the investigated test sites.

## 2 Methods

### 2.1 Study area

The study area is located in the northern Pre-Alps of central Switzerland in the region of Entlebuch, which has been accredited as an UNESCO Biosphere Reserve since September 2001 (figure 1a, b). The region is characterized by a complex topography with impenetrable gorges, rocky slopes, karst areas and fluvial deposits, which has resulted in a fragmented landscape. The region covers an area of 395 km<sup>2</sup>. The study area is dominated by strong topography with an altitudinal range between montane (600 m) and alpine zone (2,300 m). The mean annual precipitation in the area is 1,581 mm and the annual temperature ranges from −4 to 15.5°C. The landscape is mainly dominated by fragments of forest, rich and poor pastures and natural grassland, mires as well as rocks and small settlements. The study area consists of six landscape types each covering 1 km<sup>2</sup>. These landscape types are also called land-use units (LUU) and have been defined using the BioAssess gradient of land-use intensity. LUU1 contains more than 50% old growth forest and represents extensive land use. LUU6 on the other end of the gradient contains more than 50% grassland and represents intensive land use. The other LUUs are distributed according to management intensity, which is defined after the percentage of different land-use classes inside the test areas (table 1, figure 2).



**Figure 1** (a) View of the study area at UNESCO Biosphere Reserve Entlebuch, central Switzerland. (b) Locations of LUU1 to LUU6, each measuring 1 × 1 km<sup>2</sup>, and the 96 sampling plots that are illustrated as black spots

**Table 1** Description of the six land-use units (LUUs) with characterization of landscape and the land-use criteria (columns 1–3).

Description of the six land-use units (LUUs)						
LUU	Classification	Criteria (% of land use)	Dominant lichen species	No. of frequency units	Substrates	No. of frequency units
1	Old-growth forest	Old-growth forest >50%	<i>Arthonia leucopellaea</i>	230	Trees	105
		Other forests–woodland–shrubland >10% Other land uses?	<i>Graphis scripta</i>	235	Rocks	44
			<i>Lepraria lobificans</i>	320	Soil	12
			<i>Phlyctis argena</i>	329		
2	Managed forest	Managed forest >50%	<i>Chaenotheca chrysocephala</i>	119	Trees	91
		Other forests–woodland–shrubland >10% Other land uses?	<i>Hypogymnia physodes</i>	387	Rocks	49
			<i>Imshaugia aleurites</i>	421	Soil	18
			<i>Parmeliopsis ambigua</i>	262		
			<i>Phlyctis argena</i>	225		
3	Mixed use dominated by forest or woodland	Forest–woodland–shrubland >50%	<i>Chaenotheca chrysocephala</i>	159	Trees	72
		Grassland >10% Crops >10%	<i>Hypogymnia physodes</i>	556	Rocks	40
			<i>Imshaugia aleurites</i>	226	Soil	11
			<i>Parmeliopsis ambigua</i>	172		
			<i>Pseudevernia furfuracea</i>	229		
4	Mixed use not dominated by a single land use	Forest–woodland–shrubland >25%	<i>Lecanora carpinea</i>	152	Trees	93
		Grassland >25% Crops >25%	<i>Lecanora chlarotera</i>	387	Rocks	111
			<i>Phlyctis argena</i>	242	Soil	5
			<i>Lecidella elaeochroma</i>	259		
			<i>Parmelia sulcata</i>	119		
5	Mixed use dominated by pasture	Grassland >50%	<i>Hypogymnia physodes</i>	100	Trees	34
		Crops >10% Forest–woodland–shrubland >10%	<i>Lecanora dispersa</i>	140	Rocks	98
			<i>Placynthium nigrum</i>	129	Soil	12
			<i>Dermatocarpon minutum</i>	136		
			<i>Verrucaria macrostoma</i>	92		
6	Mixed use dominated by arable crops	Crops >50%	<i>Lecanora carpinea</i>	131	Trees	42
		Grassland >10% Forest–woodland–shrubland >10%	<i>Lecanora chlarotera</i>	226	Rocks	69
			<i>Phlyctis argena</i>	175	Soil	8
			<i>Verrucaria muralis</i>	154		
			<i>Lecidella elaeochroma</i>	208		

For each LUU, the five dominant lichen species are listed (column 4) with number of frequency grids where the lichens were collected (column 5). Column 7 lists the species richness (number of different species) of the three substrates (trees, rocks and soil) collected in the field survey for each LUU.

## 2.2 Training and reference data sets

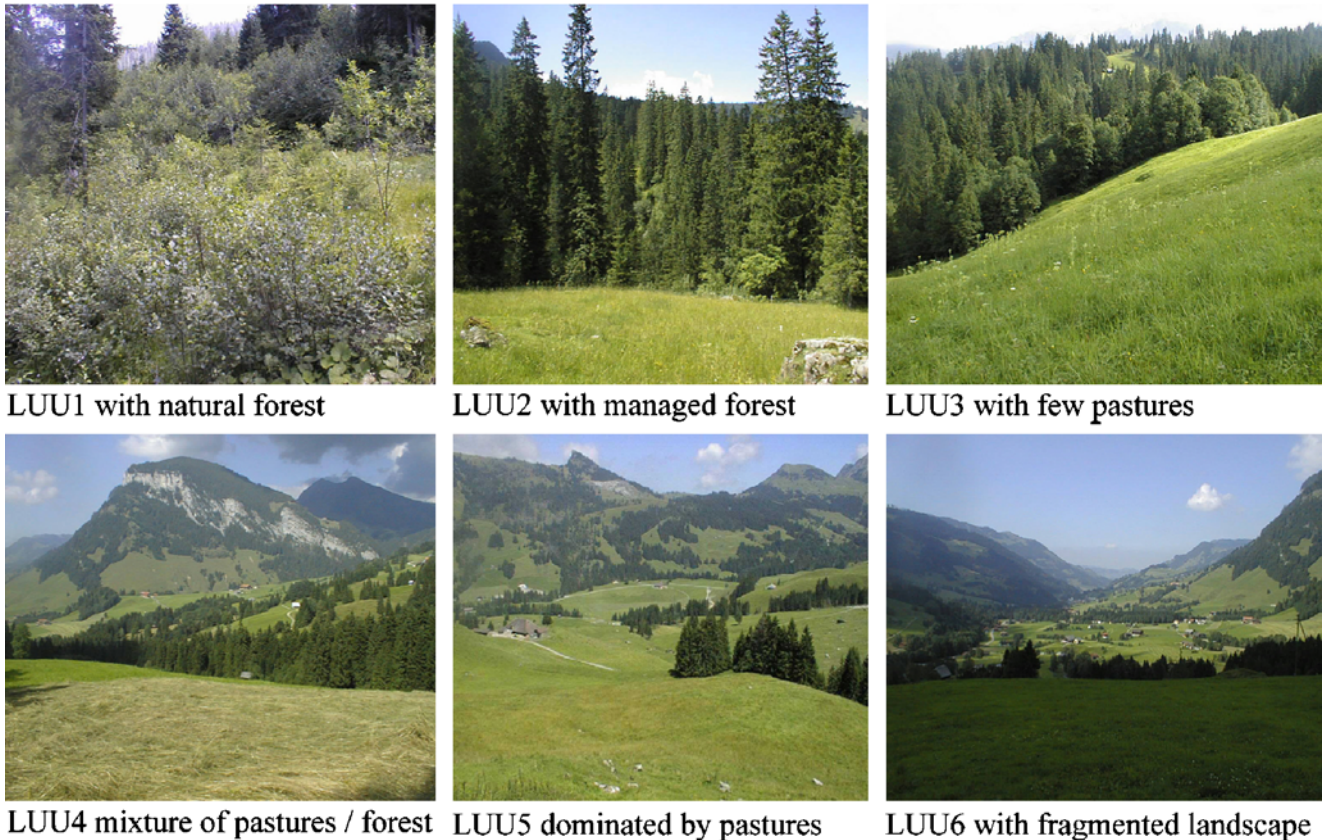
*Field data – lichen relevés* A training data set is required to calibrate the models, whereas reference data are required to validate the quality of the calibrated models. In our case, we used training data of the lichen surveys.

A total of 96 sampling plots (6 × 16) were collected that form a grid of 200 m<sup>2</sup> mesh size (figure 1). All 96 lichen sampling plots were set up by differential GPS measurements with an accuracy of ± 0.5 m. Lichen surveys were

carried out in 2001 and 2002 on the 96 sampling plots (16 per LUU) on a circular area of 1 ha (radius, 56.41 m). Within each sampling plot, 12 collecting sites were randomly selected (figure 3).

At each of the 12 collecting sites, lichen relevés were carried out on three different substrates – trees, rocks and soil – representing all major lichen substrates which could be affected by changes in agricultural and forestry management. All lichens (except if they were smaller than 5 mm) were considered that occurred inside a 50 × 40 cm frequency grid (mesh size, 10 cm) of the relevés on soil or

### **BioAssess land-use gradient**



**Figure 2** Typical landscape of LUU1 to LUU6 following a land-use intensity gradient from low to high intensively use and from closed canopy to open landscape

inside four neighboring frequency ladders (each with an extent of  $50 \times 10$  cm, mesh size 10 cm) of the relevés on trees (see figure 4). The same principle of frequency ladder was also applied for lichens on rocks. Thus, the area investigated in each relevé always remained the same,  $0.2 \text{ m}^2$  [40]. For relevés on trees, the nearest tree within the border of the sampling plot was selected (starting from the center of a collecting site). For relevés on rocks, the nearest saxicolous object within the border of the sampling plot with a size larger than  $50 \times 40$  cm was selected (starting from the center of a collecting site). For relevés on soil, in the center of each collecting site, a frequency grid of  $50 \times 40$  cm, mesh size 10 cm) was placed on the ground.

Lichenicolous fungi and non-lichenized fungi, which are often treated by lichenologists (e.g., Arthopyrenia), were not studied. For each lichen species, the number of unit areas ( $10 \times 10$  cm) where the species occurred was counted (a value ranging from 1 to 20 for both the frequency grid and the four frequency ladders). Since delimitation of individuals is often difficult or even not possible in lichens, we used the number of occupied unit areas as abundance

measure. The lichen data are stored in a relational database at the Federal Research Institute WSL at Birmensdorf, Switzerland. Table 1 gives an overview of dominant lichen species and species richness for each LUU.

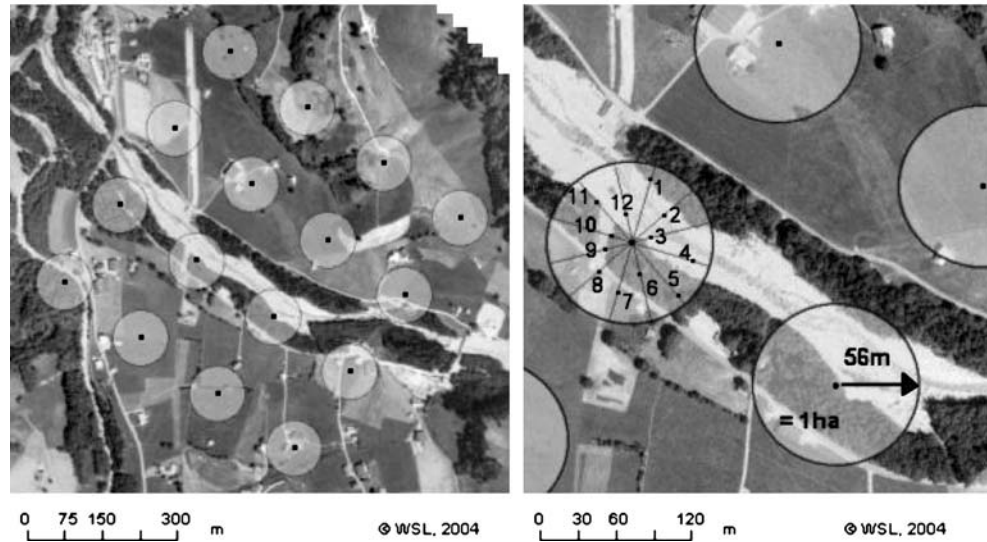
As calibration data, 48 sampling plots were chosen. The remaining 48 sampling plots served as reference data set (see figure 3).

#### 2.3 Model calibration data

Biological/ecological meaningful features as explanatory variables had to be found in order to calibrate a model of predicting lichen species richness. For this purpose, original, derived, and a combination of spectral and spatial information of both airborne and spaceborne remote sensing data were used. In this study, 12 airborne digital Color Infrared (CIR) orthoimages of 1999 and 2001 were used, each covering an area of approx.  $2 \text{ km}^2$ . The scale of 1:5,000 provides a ground resolution of 0.3 m. Each image offers three color bands of numerical information with 256 intensity levels: visible green (500–600 nm), visible red

**Figure 3** An example of the BioAssess sampling design for LUU6 with the 16 sampling plots (a circle with 56.41 m), where the three different lichen relevés were carried out at 1–12 randomly selected collecting sites (within each sampling plot). All model variables were calculated within these circles around each sampling plot

**BioAssess sampling design for LUU6**



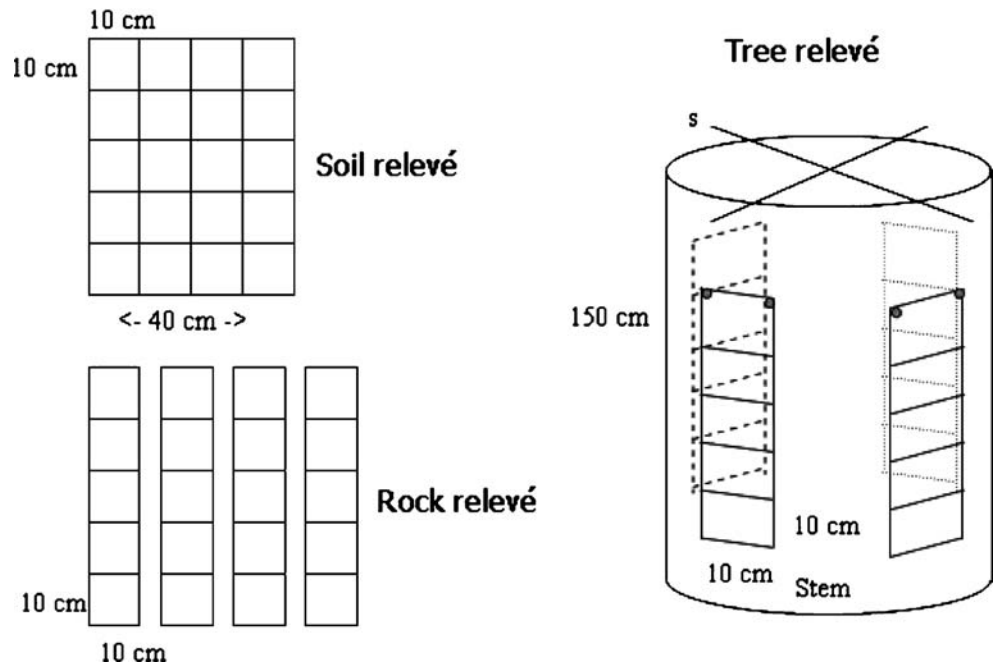
(600–700 nm) and near infrared (750–1,000 nm). The special accuracy of the CIR orthoimages is about 0.3 m. In addition to the original spectral and spatial information, several derivatives of the CIR orthoimages were calculated. A few derivatives were combined with Quickbird satellite data. Here, multispectral Quickbird satellite data of June 2002 with a spatial resolution of 2.8 m was used. Multispectral bands include blue (450–520 nm), green (520–600 nm), red (630–690 nm) and near infrared (760–900 nm). For our approach, we extracted derivatives using

both standard methods as well as methods with additional expert knowledge.

Furthermore, a digital terrain model with a spatial resolution of 25 m (DHM25 ©2003 Swisstopo, DV 455.2) and digital surface models (DSM) were used. All data sets are based on the coordinate system of the Swiss Federal Office of Topography (Wabern, Switzerland). The spatial resolution chosen for all data sets used in this study was 0.5 m, and the spatial accuracy for all data sets lies in a range of 0.5–1 m.

**Figure 4** The collection of lichen on trees, rocks and soil using frequency ladders and grids

**BioAssess lichen sampling design**



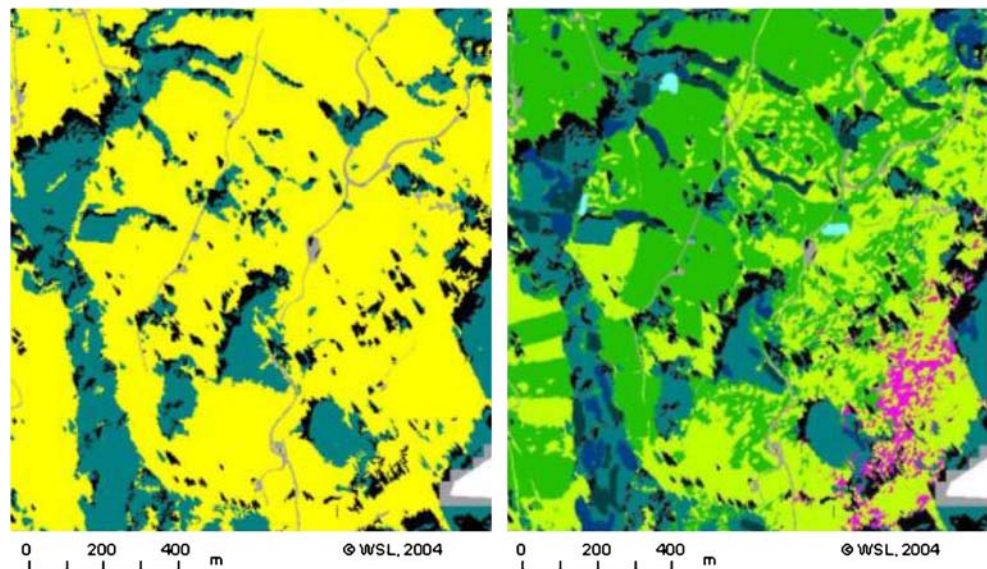
**Table 2** Overview of all 32 calculated variables divided into first and second category of detail and their linked environmental/ecological features.

Overview of all 32 explanatory variables

Name	Linked environmental/ecological features	Comments	Variable
First category			ID
Mean, majority, minority, and sum of red, green, NIR	Spectral reflection, absorption and transmission of the vegetation cover	Original channels of CIR orthoimage	1–3
Ratio1	Spectral reflection, absorption and transmission of the vegetation cover	Channel green/Channel (red + NIR)	4
Ratio2	Spectral reflection, absorption and transmission of the vegetation cover	Channel red/Channel (green + NIR)	5
Ratio3	Spatial heterogeneity of vegetation cover	Channel NIR/Channel (red + green)	6
Variance red, green, NIR	Spatial heterogeneity of vegetation cover	Returns variance in a moving window	7–9
Skewness	Spatial heterogeneity of vegetation cover	Returns skewness in a moving window	10–12
Contrast red, green, NIR	Chlorophyll absorption	Returns contrast in a moving window	13–15
Vegetation Index	Leaf area index, chlorophyll content,	NIR – red	16
NDVI	Aboveground phytomass, moisture	NIR – red/NIR + red	17
Second category			
Fraction of land cover (3 classes)	Fragmentation of vegetation and landscape	Forest, non-forest, non-vegetation	18–20
Fraction of land cover (12 classes)	Fragmentation of vegetation and landscape	Coniferous, mixed and deciduous forest, mown grass, unmown grass, rock and gravel, bare soil, sealed surface, single trees and hedges, shadows, wetlands, water bodies	21–32

The third column provides additionally information on each variable.

**Figure 5** Difference between simple and detailed classification for LUU 4. The simple classification (left map) only distinguishes non-forest (yellow), forest (green), shadows (black) and sealed surface (gray). In the more detailed classification (right map), non-forest areas are subdivided into wetlands (light blue), unmown (light green) and mown (darker green) grassland, rock&gravel (pink), and bare soil (beige). Additionally, forest areas are subdivided into deciduous forest (dark green), coniferous forest (green) and mixed forest (blue green)

**Simple and detailed classification for LUU4**

To assess and categorize the contribution of ecological meaningful variables to the model, we compared the two categories. First category variables provide information on spatial heterogeneity, spectral reflection, absorption and transmission, chlorophyll content and aboveground phyto-mass of vegetation cover. This implies simple image processing methods (standard methods) of the CIR orthoimages, and can be performed without additional expert knowledge, e.g., of biologists. In addition to the three original channels (red, green, NIR), several new variables were generated using both spatial and spectral information within a moving window of different sizes. Table 2 lists all variables applied in this study and their linked environmental/ecological features.

Second category variables are based on first category variables and were built using Quickbird data, expert knowledge and field experiences [30]. To meet these requirements, new image processing techniques were applied to produce homogenous objects and well-defined object edges [5]. Beyond the pure spectral information, these image objects are characterized by a number of additional features, such as texture and neighborhood information, which cannot be optimally exploited using standard pixel-based approaches [17]. For the extraction of second category variables, two land cover classifications were performed using both Quickbird data and CIR orthoimages: (1) a simple classification only distinguishing between forest, non-forest and non-vegetation; and (2) a more detailed classification distinguishing 12 land cover classes, representing the three lichen substrates of the field survey: 1) coniferous forest, 2) mixed forest, 3) deciduous forest, 4) mown grassland, 5) unmown grassland, 6) rock&gravel, 7) bare soil, 8) sealed surface, 9) single trees and hedges, 10) shadows, 11) wetlands, 12) water bodies. For this classification, an object-oriented approach was applied. These two classifications are shown for LUU4 in figure 5. Comparisons with reference maps (photogrammetric image interpretation of 2002) revealed an overall accuracy of 0.89 for the simple classification and 0.87 for the more detailed classification.

Table 2 gives an overview of first and second category variables. To summarize, we produced a total of 32 explanatory variables for the model. Of the total, 17 were allocated to first category variables, mainly based on simple reflection values of the three channels of the CIR orthoimages as well as on spatial information. The remaining 15 were allocated to the second category variables.

Finally, in accordance with the lichen relevés that are representative for a 56-m (radius) circle, for each variable, the sum of values was calculated within a 56-m circle for each of the 96 sampling plots. This was performed by using a moving window approach – in our case, a moving circle (see figure 6).

## 2.4 Statistical model

A model used for biodiversity assessment should not only be precise but also ecologically sensible, meaningful and interpretable [4]. Therefore, the choice of an adequate model should be carefully made, considering possible advantages and disadvantages. An important statistical development in the last 30 years has been the advance in regression analysis provided by various linear models [48].

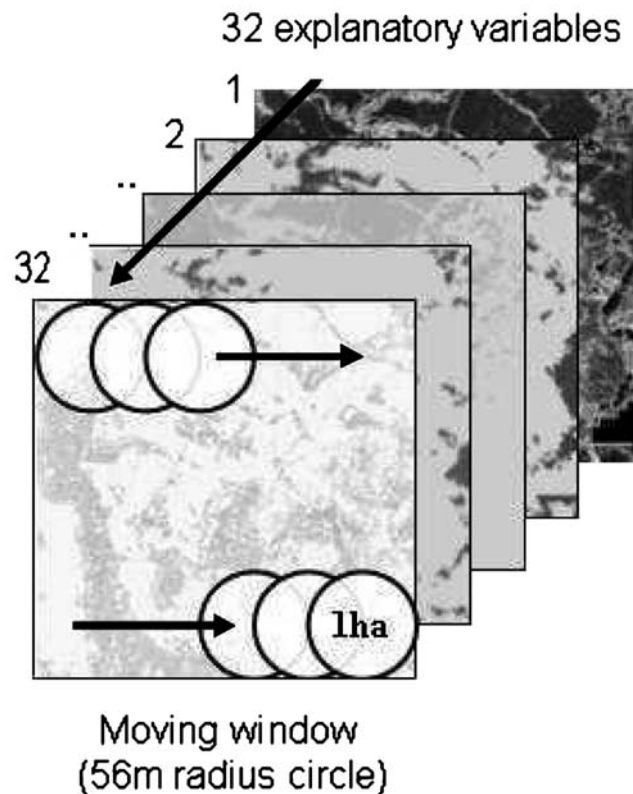
A linear model specifies the relationship between a dependent (or response) variable  $Y$ , and a set of explanatory variables  $X_i$ , so that

$$Y \sim b_0 + b_1X_1 + b_2X_2 + \dots + b_kX_k \tag{1}$$

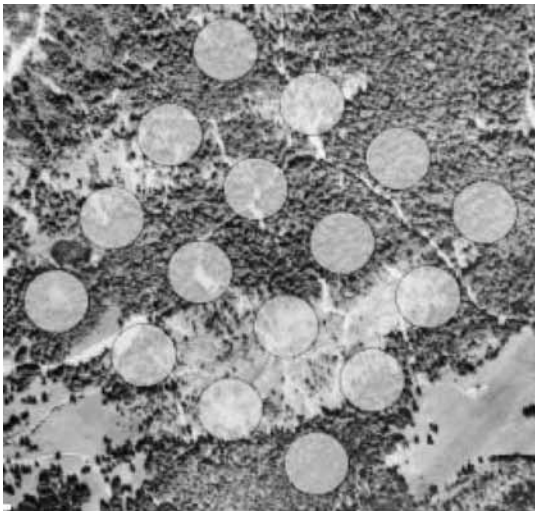
where  $b_0$  is the regression coefficient for the intercept and  $b_i$  values are the regression coefficients for the explanatory variables 1 through  $k$ , computed from calibration data. Linear models can also include quadratic terms or higher order.

Linear least-square regression can be generalized by transforming the dependent variable [25]. Generalized

## Moving window approach

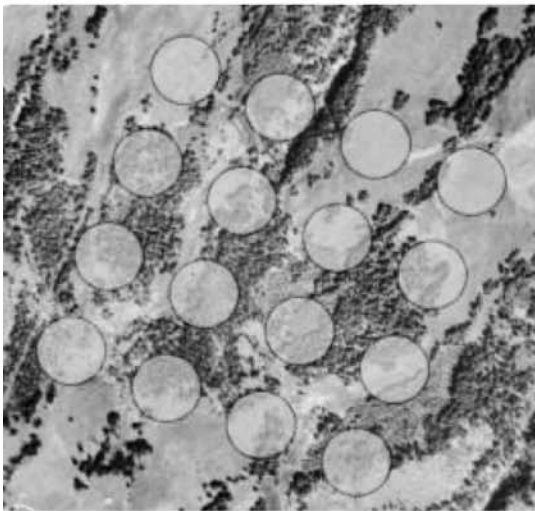
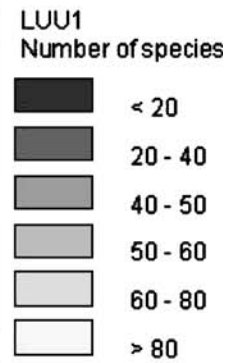
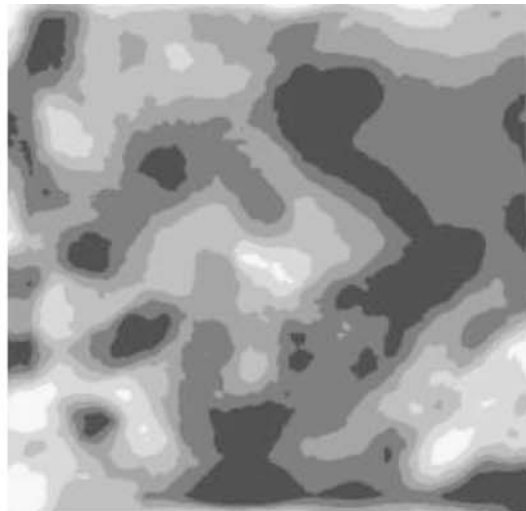


**Figure 6** Illustration of the moving window approach within the 56-m (radius) circle as applied for all 1–32 explanatory variables. The four models were applied to calculate lichen diversity hotspots for the entire extent of the LUUs



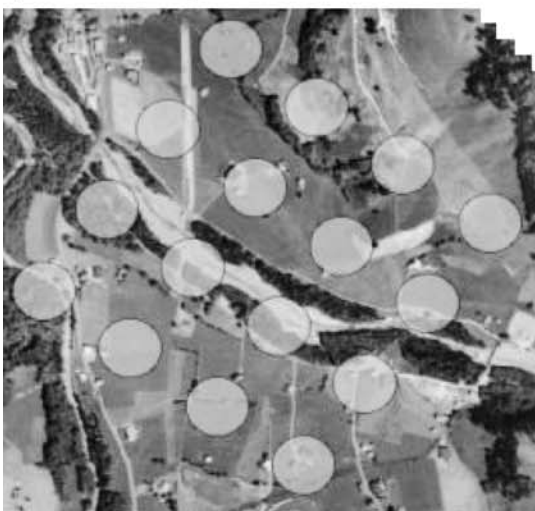
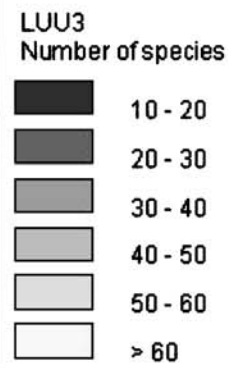
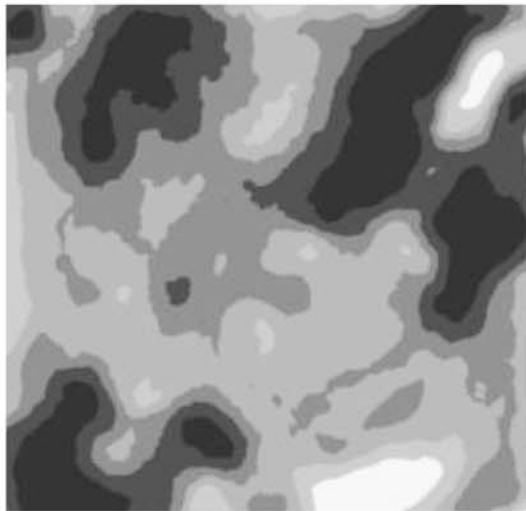
0 75 150 300 m

© WSL, 2004



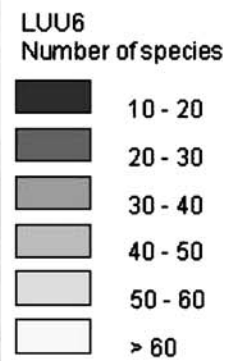
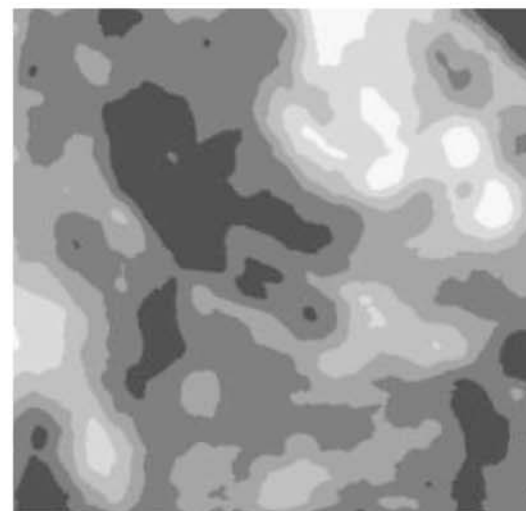
0 75 150 300 m

© WSL, 2004



0 75 150 300 m

© WSL, 2004





linear models (GLM) comprise a number of model families, e.g., binomial, Poisson, etc. [9, 13, 15, 16]. However, assuming a specific theoretical distribution for the data used in this study seems to be difficult. A first test using a Poisson distribution produced poor results and was not chosen because of the different lichen sampling design applied in this study. Different collecting procedures (i.e., different ways to the next tree and rock patch) rules out the model of the data as a Poisson process. Therefore, the simplest “first aid” transformation (square-root transformation) was used that allows coping with count data [46]. For each of the four field data sets (total species richness, species richness for lichens on trees, on rocks and on soil), stepwise dropping of our 32 explanatory variables was performed – allowing both backward and forward selection to build the models. We assumed that the relatively high number of explanatory variables, often intercorrelated, would be adequately handled by this stepwise methodology. Among the variables remaining in the final models, first category variables are used as single and as quadratic terms, whereas second category variables were square-root transformed. The analysis was done in S-PLUS [24]. The complete final models and their explanatory variables are listed below:

- Richness\_total ~ variance\_nir + variance\_nir<sup>2</sup> + ratio2 + ratio2<sup>2</sup> + sqrt(coniferous forest) + sqrt(unmown grass)
- Richness\_tress ~ variance\_nir + variance\_nir<sup>2</sup>
- Richness\_rocks ~ variance\_nir + variance\_nir<sup>2</sup> + skewness + skewness<sup>2</sup> + sqrt(unmown grass)
- Richness\_soil ~ ratio1 + ratio1<sup>2</sup> + skewness + skewness<sup>2</sup> + sqrt(rocks and gravel)

The 96 sampling plots are divided into a calibration data set of 48 randomly sampled relevés and a reference data set consisting of the remaining 48. With this calibration data set, the model was calculated and prediction values were calculated for the 48 sampling plots of the reference data. This was carried out 200 times. The means of the 200 runs are shown in table 3.

### 2.5 Model performance

Several statistic measures were applied to evaluate the predicted species richness against the measured species richness of the sampling plots. Correlation of the fitted values with the calibration data values was chosen as a measure for the model quality (*r* model in table 3). The predictive power of the model is estimated by the

correlation of predicted data values with the reference data values (*r* reference in table 3).

The accuracy of the model can be estimated by analyzing the absolute differences between each fitted value and its correspondent real value (i.e., the residual errors) of the calibration data. The accuracy of the prediction is measured by the differences between each predicted value and its correspondent real value of reference data set.

In the present study, the 95% quantile of the absolute errors, the bias (difference between the mean values and the mean fitted values), median and mean of absolute errors MAE (predicted species richness compared to reference species richness) and the *G* value are applied as accuracy measures. The *G* value (*G*) is a measure of accuracy in the case of a quantitative response and gives an indication of how effective a prediction might be, relative to that which could have been derived from using the sample mean alone. It is commonly used in ecological modeling [1, 11]. *G* is given by equation (3):

$$G = 1 - \frac{\sum_{i=1}^n |Z_{(xi)} - Z_{(Xi)}|^2}{\sum_{i=1}^n |Z_{(xi)} - \bar{Z}|^2} \tag{3}$$

where *Z*(*Xi*) is the measured value at a sampling plot *i*, *Z*(*xi*) is the estimated value and  $\bar{Z}$  is the overall mean of the measured sampling plots. A value of 1 indicates a perfect prediction, while a value of 0 describes no significant agreement, and negative values indicate that the predictions are less reliable than if one had used the sample mean instead [41].

### 2.6 Application of model

To extrapolate the predicted species richness of the sampling plots to the entire area of the six LUUs, the model had to be applied accordingly. For each pixel of the six test sites, lichen species richness was calculated implementing the explanatory variables for the final models in a moving window approach (in our case, a moving circle). The sum of values within our 56-m (radius) circle was calculated for each pixel of the selected explanatory variable (see figure 6) using Geographic Information System (GIS) operations. The four corresponding model equations (with their coefficients) as given in the section **Statistical model** serve as a basis for pixel-wise calculation of species richness for all lichens, lichens on trees, on rocks and on soil. As results, six maps of predicted number of

◀ **Figure 7** Maps of predicted species richness for all lichens for LUU1–LUU6 with their corresponding CIR orthoimages and the 1-ha circles of sampling plots the models were calibrated with. On the right side, light values indicate highest species richness (=hotspots), whereas dark values represent low species richness

lichen species for each pixel in the entire six LUUs were obtained (figure 7).

### 3 Results

The best results of the models and the combination of explanatory variables retained in each model are given in table 3. The quality of the models ( $r$  model) ranges between 0.59 for lichens on soil and 0.79 for lichens on trees. Highest predictive power, with a correlation coefficient ( $r$  reference) ranging between 0.5 and 0.8, and  $G$  ranging between 0.65 and 0.38, is obtained using both first and second category variables, with the exception of the species richness for lichens on trees. In general, species richness is slightly underestimated for sampling plots with high species richness and overestimated for sampling plots with low species richness. The *bias* for all four models ranges from +0.51 to +3.19. A total of 32 variables correlated with the number of lichen species, but only seven were used for the final models. Best results for the model for all lichens are obtained using the explanatory variables *variance\_nir* (var-

iance of NIR channel), *ratio2* (ratio Channel red/Channel (green + NIR)), *coniferous\_forest* (fraction of coniferous forest) and *unmown\_grass* (fraction of unmown grassland), resulting from the detailed land cover classification. The model produces an  $r$  model of 0.68, an  $r$  reference of 0.59 and a  $G$  of 0.53.

Concerning the model for lichens on trees, best results are obtained by the single use of explanatory variables *variance\_nir*;  $r$  model equals 0.79 and  $r$  reference reaches 0.8. The associated  $G$  equals 0.65.

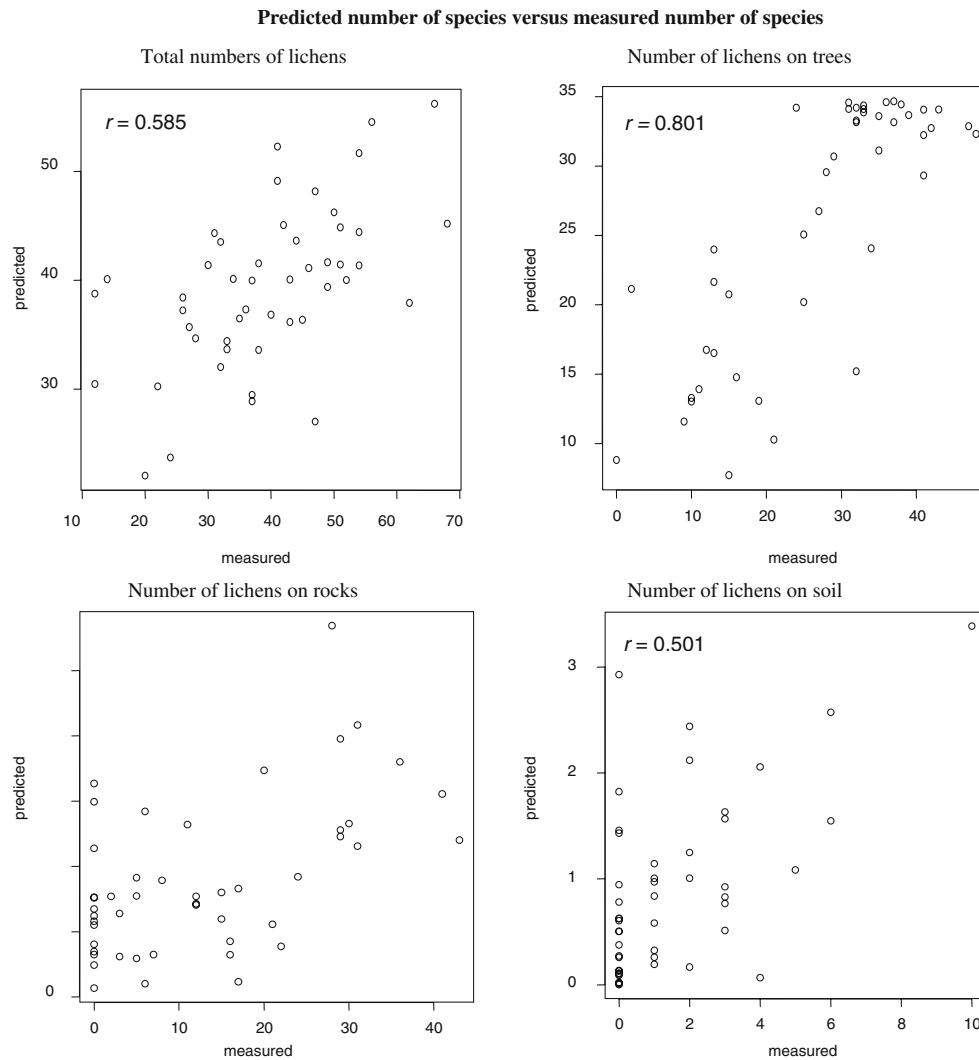
Similar correlation coefficients for both model and reference are obtained for the models lichens on rocks and on soil. Best results are obtained for the model lichens on rocks using the explanatory variables *ratio2*, *skewness* and *unmown\_grass*. This model generates an  $r$  model of 0.61, whereas  $r$  reference is 0.56 and  $G$  reaches 0.38.

Finally, for the model for lichens on soil, best results are obtained using the variables *ratio1* (ratio Channel green/Channel (red + NIR)), *skewness* and *rock&gravel* (fraction of rock and gravel); the latter again resulting from the detailed land cover classification. The  $r$  model is of 0.59, while  $r$  reference reaches 0.5. The associated  $G$  equals 0.42.

**Table 3** Means of 200 runs for validation of the four calibrated linear regression models of the species richness for lichens total, on trees, on rocks and on soil.

Models	Species richness			
	Lichen total	Lichens on trees	Lichens on rocks	Lichens on soil
Goodness-of-fit ( $r$ model)	0.682	0.791	0.611	0.594
Predictive power ( $r$ reference)	0.585	0.801	0.560	0.501
Median (model)	7.456	5.278	6.423	0.598
Median (reference)	8.651	5.652	7.634	0.698
MAE (model)	8.456	6.313	8.689	0.997
95% Quantile of error (model)	20.123	15.120	19.175	3.613
MAE (reference)	9.898	7.001	9.344	1.256
95% Quantile of error (reference)	25.270	15.456	24.232	4.023
Bias (model)	+1.012	+1.501	+3.190	+0.510
Bias (reference)	+0.870	+1.532	+2.703	+0.535
$G$	0.534	0.654	0.382	0.421
Model parameters				
First category explanatory variables	Variance_NIR	Variance_NIR	–	–
	Variance_NIR <sup>2</sup>	Variance_NIR <sup>2</sup>	–	–
	–	–	–	Ratio1
	–	–	–	Ratio1 <sup>2</sup>
	Ratio2	–	Ratio2	–
	Ratio2 <sup>2</sup>	–	Ratio2 <sup>2</sup>	–
	–	–	Skewness	Skewness
Second category explanatory variables	–	–	Skewness <sup>2</sup>	Skewness <sup>2</sup>
	Sqrt(coniferous_forest)	–	–	–
	Sqrt(unmown_grass)	–	Sqrt(unmown_grass)	–
			Sqrt(rock&gravel)	

Only explanatory variables as used for the final models are listed.



**Figure 8** Predicted number of species versus measured number of species for all lichens and lichens on the substrates trees, rocks and soil. The predictive power of the models (means of 200 runs) is given by the correlation coefficient  $r$  reference

As an overview, figure 8 illustrates the means of the 200 runs of the predicted species richness *versus* measured species richness for all lichens and lichens on the substrates trees, rocks and soil. Figure 7 shows the maps of the predicted number of species for all lichens in LUU1, LUU3 and LUU6 following a land-use intensity gradient from low to high intensively use. Areas with low numbers of species are mapped in black, whereas areas with high numbers of species are white.

#### 4 Discussion

With this study, we can confirm that the application of homogenous and reproducible land cover information derived from both space- and airborne remotely sensed data as a basis for the model is adequate, and therefore is in

accordance with other studies in this field [38, 42]. Interpretation of CIR aerial photographs could be a useful method to find certain groups of lichens [3]. The accuracies ( $r$  reference) obtained for both model lichens on trees (0.80) and for all lichens (0.59) can be regarded as good for the application purposes by lichenologists. These accuracies are in accordance with similar experiences of other studies [18, 47]. Also,  $G$  values of 0.65 for the model lichen on trees and 0.53 for all lichens indicate a good prediction [11]. On the other hand, models for lichens on rocks ( $r$  reference = 0.56) and for lichens on soil (0.50) produced partly satisfactory results and should be further improved (see below). Both underestimation and overestimation occur in all four models. The crucial question is how can we improve our models for lichen species richness?

In this study, we were confronted with several problems concerning ecological modeling. A model used for biodi-

versity assessment should also be general, which means it should be applicable in other regions or different times [21]. Furthermore, the lack of validation and uncertainty assessment of models remains a serious issue in ecological modeling [10]. Finally, a model used for biodiversity assessment should not only be precise, but should also be ecologically sensible, meaningful and interpretable [4].

Meeting all the suggested requirements turns out to be nearly impossible in our case. For example, the particular model developed here was applied only for six test sites, each with an extent of  $1 \times 1$  km. Thus, the resulting variables of the presented linear models may be used for calculating species richness in neighboring regions of the UNESCO Biosphere Reserve Entlebuch or the Northern Pre-Alps with similar vegetation cover and landscape structures. Applying the model to other regions (e.g., Central or Southern Alps) is a well-known problem [18]. The model would first need to be adapted and validated before being applied elsewhere in the Swiss Alps, but our basic approach could be the same. For a further improvement of model accuracies, especially for the models lichens on soil and all lichens, we recommend the following: (1) further analyses of the distribution of the lichen data and the sampling design; (2) further testing of other model versions; (3) implementing additional sampling plots in all LUU (the number of 96 sampling plots for this study could be regarded as relatively poor); and (4) extraction of further calibration data. Points (1) and (2) are related to a crucial problem that this study had to deal with – the choice of the “right” model because the sampling design does not suggest a Poisson distribution. Different collecting procedures (i.e., different ways to the next tree and rock patch) rules out the model of the data as a Poisson process. After various test runs, a simple square-root transformation was chosen and turned out to be adequate for our modeling purposes. The root-square transformation of both the dependent and the explanatory variables lets certain characteristics of the data express themselves more or less strongly. Concerning points (3) and (4), we plan to improve our models considering these recommendations in the near future and to test the modified models in other alpine regions. Another very flexible, powerful approach to habitat modeling is the use of HyperNiche software [2, 27]. This also implies a method that automatically incorporates interactions into multiplicative kernel smoothers. Furthermore, it is especially useful for models where the predictors are combined multiplicatively rather than additively; we plan to test this software for future model development and refinement [26, 34, 35].

The distinction between first and second category variables was an important step for the development of the final models and allowed us to assess their contribution to the corresponding model. Furthermore, it helped us to

drop the variables that contribute less to the model. All 32 variables are linked to a specific biological/ecological meaningful feature as shown in table 2.

In particular, the use of simple spectral and textural information values of the CIR orthoimages, which is linked to spectral reflection and spatial heterogeneity of the vegetation cover, respectively, produced the best results. In fact, the implementation of additional second category variables improved model accuracy again – with the exception of lichens on trees. For this model, best accuracy ( $r$  model, 0.79;  $G$ , 0.65) was produced with the single use of variance\_nir and its quadratic term, whereas the implementation of additional explanatory variables slightly deteriorated the model’s accuracy. In this case, the number of species is directly related to a high heterogeneous vegetation cover such as forest borders and forest itself.

For the other models, the implementation of root-squared transformation of these second category variables helped to significantly improve model accuracies again. The subdivision of the classes forest into coniferous, mixed and deciduous forest was a good decision. The three second category explanatory variables coniferous\_forest, unmown\_grass and rock&gravel turned out to be the most contributive variables for all three models. In particular, the occurrence of unmown\_grass was positively correlated with the species richness. The omission of unmown\_grass for the model with all lichens and for lichens on rocks lowered accuracies by a factor of about 20%. This indicates that the influence of unmown grassland on lichen diversity should not be underestimated. Furthermore, habitats for lichens on rocks are located in rather heterogeneous landscapes. A reason for this may be that habitats for lichens on rocks are located in highly fragmented landscapes (spatial heterogeneity of vegetation cover), where fragments of unmown area are a dominating land cover type. Interestingly, no other second category variable in our analysis was really significant. Even the implementation of variables deciduous\_forest and mixed\_forest did not have significant influences on the models. According to the lichenologists, one would have expected the explanatory variables sealed\_surfaces (streets, buildings) and mixed\_forest to have a certain influence on the model for species richness.

The 12 land-cover types extracted for this study are based on what was supposed to be detectable in CIR orthoimages and Quickbird data, and what was regarded to be of importance for the lichen diversity. A comparison of this classification with visually interpreted aerial images revealed high accuracy and agreement. The main advantage of the application of an object-oriented image classification method is that it allowed us to define land cover types according to the needs of lichen experts. The implementation of a forest-stand height classification using digital surface models derived from the CIR orthoimages and a

digital terrain model further improved the classification result. A combination of CIR orthoimages with the multispectral Quickbird image considerably improved the distinction between forest and non-forest areas. Thus, the land cover classification applied in this study in combination with image segmentation methods was an important step in the development of the models. The main disadvantage was the relatively high complexity and required amount of time of object-oriented image classification methods.

## 5 Conclusion

In this interdisciplinary study, remote sensing techniques were implemented in an ecological modeling approach according to the needs of lichen experts. Four different models for an assessment of lichen species richness on three different substrates were examined – trees, rocks and soil – for six test sites within the UNESCO Biosphere Reserve Entlebuch in Switzerland. Species richness was modeled as functions of 32 remotely sensed explanatory variables.

There are five points to remember about the benefits of this study. First, linear regression models can be used to predict lichen diversity, but strongly depend on the sampling design of the lichen relevés. Thus, the distribution of the lichen data should be analyzed further. Second, possible hotspots were calculated and may help in reducing field surveys and could be useful for possible conservation efforts in similar regions. The resulting explanatory variables of the presented linear models may be used for calculating species richness in neighboring regions with similar landscape structures. Third, we can confirm that the application of homogenous and reproducible land cover information derived from high-resolution remote sensing data as a basis for the model is very adequate. This means that not-so-well-known areas can still serve as a basis for building the methods. Fourth, explanatory variables can be rapidly derived from high-resolution remote sensing data and distinguishing between first and second categories proved to be a good method for the development of the models. Fifth, the detailed classifications applied in this study produced up-to-date maps for public authorities and nature conservation agencies.

Nevertheless, this method cannot replace lichen surveys altogether; however, it can be used to target focused lichen forays in the future. Finally, it should be noted that this method cannot produce any information on lichen species abundance, dynamics or viabilities; it only indicates the potential presence or absence of species.

**Acknowledgments** The study was carried out in the framework of the Biodiversity Assessment Tools Project (EVK2-CT1999-00041) in

the Energy, Environment and Sustainable Development Programme of the European Union. We thank Christian Ginzler, Peter Longatti, Klaus Ecker, Ariel Bergamini and Niklaus Zimmermann for all their valuable help. We also thank Christine Keller, Urs Groner and Michael Dietrich for lichen sampling and determination. A special thanks also goes to Allan Watt, the BioAssess project coordinator.

## References

- Agterberg, F. P. (1984). Trend surface analysis. In G. L. Gaile & C. J. Willmott (Eds.), *Spatial statistic and models*, (pp. 147–171). Dordrecht: Reidel.
- Antoine, M. E., & McCune, B. (2004). Contrasting fundamental and realized ecological niches with epiphytic lichen transplants in an old-growth Pseudotsuga forest. *Bryologist*, *107*, 163–173.
- Ask, P., & Nilsson, S. G. (2004). Stand characteristics in color-infrared aerial photographs as indicators of epiphytic lichens. *Biodiversity and Conservation*, *13*, 529–542.
- Austin, M. P., & Gaywood, M. J. (1994). Current problems of environmental gradients and species response curves in relation to continuum theory. *Journal of Vegetation Science*, *5*, 473–482.
- Baatz, M., & Schäpe, A. (2002). Object-oriented and multi-scale image analysis in semantic networks. In J. Strobl, T. Blaschke, & G. Griesebner (Eds.), *Angewandte Geographische Informationsverarbeitung XII* (pp. 12–23). Wichmann-Verlag: Heidelberg.
- Barry, R. G. (1994). Past and potential future changes in mountain environments. In M. Beniston (Ed.), *Mountain environments in changing climates* (pp. 3–33). London: Routledge.
- CIPRA, 2. (2001). *Alpenreport – Daten, Fakten, Probleme, Lösungsansätze* (434). Bern: Verlag Paul Haupt.
- Cousins, S. A. O., & Ihse, M. (1998). A methodological study for biotope and landscape mapping based on CIR aerial photographs. *Landscape and Urban Planning*, *41*, 183–192.
- Dobson, A. J. (2002). *An introduction to generalized linear models*. Boca Raton: Chapman and Hall/CRC.
- 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*, 38–49.
- Gotway, J. C., Ferguson, R. B., Hergert, G. W., & Peterson, T. A. (1996). Comparison of kriging and inverse-distance methods for mapping soil parameters. *Soil Science Society of America Journal*, *60*, 1237–1247.
- Grabherr, G., Gottfried, M., & Pauli, P. (1994) Climate effects on mountain plants. *Nature*, *369*, 448.
- Green, P. J., & Silverman, B. W. (1994). *Nonparametric regression and generalized linear models: A roughness penalty approach* (p. 182). London: Chapman and Hall.
- Guisan, A., Edwards, T. C., & Hastie, T. (2002). Generalized linear and generalized additive models in studies of species distributions: Setting the scene. *Ecological Modelling*, *157*, 89–100.
- Guisan, A., & Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*, *135*, 147–186.
- Hastie, T. J., & Tibshirani, R. J. (1990). *Generalized additive models*. London: Chapman & Hall.
- Hoffmann, A., & Van der Vegt, J. W. (2001). New sensor systems and new classification methods: Laser- and digital camera data meet object-oriented strategies, 6, Proceedings of GeobIT/GIS.
- Iverson, L. R., & Prasad, A. (1998). Estimating regional plant biodiversity with GIS planning. *Diversity and Distribution*, *4*, 49–61.
- Jørgensen, S. E. (1997). Ecological modeling by 'ecological modeling'. *Ecological Modelling*, *100*, 5–10.
- Kushida, K., Kim, Y., Tanaka, N., & Fukuda, M. (2004). Remote Sensing of net ecosystem productivity based on component

- spectrum and soil respiration observation in a boreal forest, interior Alaska. *Journal of Geophysical Research* (Atmosphere), *109*, D6, DO6209.
21. Leathwick, J. R., Whitehead, D., & McLeod, M. (1996). Predicting changes in the composition of New Zealand's indigenous forests in response to global warming: A modeling approach. *Environmental Software*, *11*, 81–90.
  22. Lehmann, A., Overton, J. M., & Austin, M. P. (2002). Regression models for spatial prediction: Their role for biodiversity and conservation. *Biodiversity and Conservation*, *11*, 2085–2092.
  23. Maisch, M. (2001). The long-term signal of climate change in the Swiss Alps: Glacier retreat since the end of the Little Ice Age and future ice decay scenarios. *Geografia Fisica e Dinamica Quaternaria*, *23*, 139–151.
  24. Math Soft (1999). *S-Plus 2000 professional release 1*. Seattle, WA: Math Soft Inc.
  25. McCullagh, P., & Nelder, J. A. (1989). *Generalized linear models*, (511). London: Chapman and Hall.
  26. McCune, B., Berryman, S. D., Cissel, J. H., & Gitelman, A. I. (2003). Use of a smoother to forecast occurrence of epiphytic lichens under alternative forest management plans. *Ecological Applications*, *13*, 1110–1123.
  27. McCune, B., & Mefford, M. J. (2004). *Nonparametric multiplicative habitat modeling, version 1*. Gleneden Beach, OR, USA: MjM Software.
  28. Nagendra, H., & Gadgil, M. (1999). Satellite imagery as a tool for monitoring species diversity: An assessment. *Journal of Applied Ecology*, *36*, 388–397.
  29. Nash, T. (1996). *Lichen biology*. Cambridge: Cambridge University Press.
  30. Nimis, P. L., & Martellos, S. (2001). Testing the predictivity of ecological indicator values – a comparison of real and virtual relevés of lichen vegetation. *Plant Ecology*, *157*, 165–172.
  31. Nimis, P. L., Scheidegger, C., & Wolseley, P. A. (2002). In: *Monitoring with Lichens – Monitoring Lichens*. (408) The Netherlands, Dordrecht, Boston, London: Kluwer Academic Publishers.
  32. Nordberg, M. L., & Allard, A. (2002). A remote sensing methodology for monitoring lichen cover. *Canadian Journal of Remote Sensing*, *28*, 262–274.
  33. Palmer, M. W. (1995). How should one count species? *Natural Areas Journal*, *15*, 124–135.
  34. Peterson, A. T., & Robins, C. R. (2003). Using ecological-niche modeling to predict barred owl invasions with implications for spotted owl conservation. *Conservation Biology*, *17*, 1161–1165.
  35. Peterson, A. T., Stockwell, D. R. B., & Kluza, D. A. (2002). In J. M. Scott, P. J. Heglund, & M. L. Morrison (Eds.), *Predicting species occurrences: Issues of Scale and Accuracy* (pp. 617–623). Washington, DC: Island Press.
  36. Price, M. F., & Haslett, J. R. (1995). Climate change and mountain ecosystems, mountains at risk. In N. J. R. Allan (Ed.), *Current issues in environmental studies* (pp. 73–97). New Delhi: Manohar.
  37. Rose, F. (1976). In D. H. Brown, D. L. Hawksworth, & R. H. Bailey (Eds.), *Lichenology: Progress and problems* (pp. 279–307). London: Academic Press.
  38. Roy, P. S., & Tomar, S. (2000). Biodiversity characterization at landscape level using geospatial modeling technique. *Biological Conservation*, *95*, 95–109.
  39. Scheidegger, C., & Goward, T. (2002). In P. L. Nimis, C. Scheidegger, & P. A. Wolseley (Eds.), *Monitoring with lichens – monitoring lichens. IV. Earth and Environmental Science* (pp. 163–181). Dordrecht, The Netherlands: Kluwer Academic Publishers.
  40. Scheidegger, C., Groner, U., Keller, C., & Stofer, S. (2002). In P. L. Nimis, C. Scheidegger, & P. A. Wolseley (Eds.), *Monitoring with Lichens – Monitoring Lichens, IV. Earth and Environmental Science* (pp. 359–365). Dordrecht, The Netherlands: Kluwer Academic Publishers.
  41. Schloeder, C. A., Zimmermann, N. E., & Jacobs, M. J. (2001). Comparison of methods for interpolating soil properties using limited data. *Soil Science Society of America Journal*, *65*, 470–479.
  42. Scott, J. M., & Csuti, B. (1997). Gap Analysis for biodiversity survey and maintenance. In M. L. Reaka-Kudla, D. E. Wilson, & E. O. Wilson (Eds.), *Biodiversity II: Understanding and protecting our biological resources* (pp. 321–340). Washington, DC: Joseph Henry Press.
  43. Scott, J. M., Heglund, P. J., Samson, F., Haufler, J., Morrison, M., Raphael M., et al. (2002). *Predicted species occurrences: Issues of accuracy and scale*. Covelo, CA: Island Press.
  44. Tasser, E., & Tappeiner, U. (2002). The impact of land-use changes in time and space on vegetation distribution in mountain areas. *Applied Vegetation Science*, *5*, 173–184.
  45. Theurillat, J. P., & Guisan, A. (2001). Impact of climate change on vegetation in the European Alps. *Climatic Change*, *50*, 77–109.
  46. van Tongeren, O. R. F. (2002). In R. H. G. Jongman, C. J. F. Ter Braak, & O. F. R. Van Tongeren (Eds.), *Data Analysis in Community and Landscape Ecology*, (pp. 174–212). Cambridge: Cambridge University Press.
  47. Wohlgemuth, T. (1998). Modeling floristic species richness on a regional scale: A case study in Switzerland. *Biodiversity and Conservation*, *7*, 159–177.
  48. Yee, T. W., & Mackenzie, M. (2002). Vector generalized additive models in plant ecology. *Ecological Modelling*, *157*, 141–156.