Available online at www.sciencedirect.com

SCIENCE DIRECT®

Biological Conservation 118 (2004) 501–511

BIOLOGICAL CONSERVATION

www.elsevier.com/locate/biocon

Development and evaluation of predictive habitat models to assist the conservation planning of a threatened lucanid beetle, *Hoplogonus simsoni*, in north-east Tasmania

Jeff M. Meggs [a,b,*], Sarah A. Munks [a,b], Ross Corkrey [c], Karen Richards [a,b]

[a] *Forest Practices Board, 30 Patrick Street, Hobart, Tasmania, 7000 Australia*
[b] *Department of Zoology, University of Tasmania, Sandy Bay, Tasmania, 7005 Australia*
[c] *Lighthouse Research Station, George Street, Cromarty Ross-shire, IV11 8VJ UK*

Received 25 March 2003; received in revised form 1 October 2003; accepted 1 October 2003

Abstract

The use of predictive habitat distribution models by land managers in the conservation management of threatened species is increasing. Few models, however, are subsequently field-checked and evaluated. This study evaluates the statistical strength and usefulness for conservation purposes of three predictive habitat models developed for a threatened stag beetle, *Hoplogonus simsoni*, found in the wet eucalypt forests and mixed/rainforests of north-east Tasmania. The relationship between various environmental variables for which spatial (GIS) information was available and the density, frequency of occurrence and presence/absence of the species was investigated using generalised linear modelling. Models developed were coupled with the GIS data to develop maps of predicted occurrence within the species' range, grouped into categories of habitat quality. The models found that altitude, aspect, slope, distance to nearest stream and overstorey tree height were significantly associated with the occurrence of the species. Evaluation of the statistical strength of the models with independent data of species' occurrence collected at 95 sites found that the density model performed poorly with little correlation between predicted and observed densities of the species. The frequency of occurrence model, however, showed a moderate ability to predict both species' abundance and presence/absence. The presence/absence model had a similar discriminatory ability in predicting presence or absence of *H. simsoni*, but also showed some potential as an indirect predictor of species' abundance. Assuming a correlation between relative abundance and habitat quality, the frequency of occurrence predictive model appeared to be the better and more direct discriminator of high quality habitat relative to the other models. The value of species' habitat models and the need to evaluate their utility in the development of conservation strategies are discussed.

Keywords: Predictive modelling; Habitat modelling; Model evaluation; GIS; Threatened species; Forestry; Invertebrates

1. Introduction

With the advent of more powerful statistical tools and geographic information systems (GIS) there has been a growing use of predictive habitat distribution models in the management of fauna and flora of conservation significance. Most often, this technique has been used on "charismatic" vertebrates, particularly mammals such as wolves (Mladenoff et al., 1995; Corsi et al., 1999), the grizzly bear (Mace et al., 1999), the black bear (Van Manen and Pelton, 1997; Clevenger et al., 2002), lemurs (Smith et al., 1997), and marsupials (Munks, 1993; Lindenmayer et al., 1995; Pearce and Ferrier, 2001), but also birds (Pearce and Ferrier, 2001), including the golden eagle (Fielding and Haworth, 1995). Similar studies have been conducted on reptiles (Pearce and Ferrier, 2001), and vascular plants (Pearce and Ferrier, 2001), including alpine grasslands (Zimmermann and Kienast, 1999) and rare plants (Elith and Burgman, 2002).

In contrast, there have been very few attempts at creating spatially explicit predictive models for

* Corresponding author. Present address: Forestry Tasmania, 79 Melville Street, Hobart, Tasmania, Australia 7000. Fax: +61-3-6233-8252.
E-mail address: jeffrey.meggs@forestrytas.com.au (J.M. Meggs).

0006-3207/$ - see front matter © 2003 Published by Elsevier Ltd.
doi:10.1016/j.biocon.2003.10.001

invertebrates. Exceptions include a study by Fleishman et al. (2001) that developed statistically significant models predicting species occurrence for 36 of 56 resident butterflies in the central Great Basin of western North America. Rushton et al. (1994) created a relatively coarse-scale predictive distribution model that related the occurrence of carabid beetle species in Britain to various land classification categories. In addition, Ehrlich and Murphy (1987) and Murphy et al. (1990) have incorporated models of habitat quality with spatial and temporal metapopulation dynamics to examine the population viability of threatened butterflies. The poor representation of invertebrates in the literature on predictive distribution modelling may be because of the lack of comprehensive distribution data (including absences) and quantitative habitat data for invertebrates. Environmental variables that may be important to terrestrial invertebrates such as soil, leaf litter and understorey characteristics may not exist in GIS, or may not have adequate surrogates within GIS (York, 1999; although this can also be the case for vertebrates, see Lindenmayer et al., 1999). Many environmental variables such as climate and topography that can influence invertebrate distributions may exist in GIS at a scale too coarse to be of use for organisms that often operate at small spatial scales (Elith, 2000). The paucity of invertebrate studies of this type is also a likely consequence of the general lack of attention invertebrate conservation has received from land managers (Yen et al., 1990).

To conserve threatened species effectively, knowledge is required of their physiological and ecological requirements and responses to disturbance. In the absence of such information, the development of models that predict the extent and distribution of habitats that threatened species utilise can be invaluable to land managers. The generally high stakes involved in threatened species management (i.e., potential species' extinction and/or economically or socially significant land-use changes) dictate that such models should be rigorously evaluated (Fielding and Bell, 1997; Manel et al., 1999; Guisan and Zimmermann, 2000). Some predictive habitat models have been subsequently field-checked and evaluated in terms of their statistical strength and/or their usefulness for conservation purposes (e.g. Rushton et al., 1994; Fielding and Haworth, 1995; Lindenmayer et al., 1995; Mladenoff et al., 1995; Zimmermann and Kienast, 1999; Pearce and Ferrier, 2001; Elith and Burgman, 2002). However, as well as a lack of empirical data, urgency for action is often a driver for the development of predictive distribution models to assist land-use decisions. Hence, model evaluation may not be undertaken (e.g. Munks, 1993; Smith et al., 1997; Mace et al., 1999) or the evaluation involves comparing the predictions against the original data used to develop the model (e.g. Bustamante, 1997; Clevenger et al., 2002), often leading to an overly optimistic assessment of the model's predictive ability.

For many of the invertebrates listed in the schedules of the Tasmanian *Threatened Species Protection Act* 1995 comprehensive distribution and habitat utilisation data are not available (Taylor and Bryant, 1997; Munks and Taylor, 2000). An exception is Simsons stag beetle, *Hoplogonus simsoni* Parry (Coleoptera: Lucanidae), which is listed as vulnerable due to its restricted distribution, generally low population densities and the potential adverse impacts of forestry practices within its range (Meggs et al., 2003). The species is patchily distributed throughout its 250 km^2 range, centred on the Blue Tier in north-east Tasmania, and potential habitat for the species (i.e., wet eucalypt forest and mixed/rainforest) encompasses 18,200 ha of its range (Meggs et al., 2003). Meggs et al. (2003) identified optimal habitat for the species as wet eucalypt forest below 300 m altitude, with a slope less than 5°, a deep leaf-litter layer, and a forest structure with a well-developed tall shrub layer. They suggested that these characteristics relate to the beetle's requirement for a relatively cool, moist, stable microclimate and the absence of disturbance for at least 50 years, but possibly longer. They also found that potential habitat of *H. simsoni* was poorly reserved across its range and a high percentage had been identified by the forest industry as having potential for conversion to pine plantation, a practice that results in the local extinction of the species (Meggs et al., 2003). It was recommended that the conservation requirements of the species would best be served by the reservation of areas containing high-density populations, limitation of the area of potential habitat that may be converted to plantation, and the retention of contiguous links of undisturbed forest throughout its range (Meggs et al., 2003).

The study by Meggs et al. (2003) provides a basis for predicting the spatial distribution of habitats important to *H. simsoni* and the areas where the conservation of the species may conflict most strongly with planned forestry activities. In the present study the aim was to develop habitat models from the abundance and habitat variable data collected by Meggs et al. (2003) that could be coupled with GIS data to create predictive distribution and abundance maps of *H. simsoni*. The GIS-generated predictions of presence and abundance were then field-checked and the models evaluated in terms of their statistical strength and their utility as a conservation planning tool for the management of *H. simsoni*.

2. Methods

2.1. Study area, animal survey and habitat variables for model development

The species' presence, abundance and habitat variable data used for the development of the GIS-based predictive models were the same as collected by Meggs et al. (2003). Details of the study area, the stratification methods for the field survey, and the collection methods for each of the habitat variables and the beetle are described in Meggs et al. (2003). Hence, only a brief outline of the original sampling methodology is provided here. All field work was conducted between November 1996 and May 1997 in an area encompassing the known range of *H. simsoni* in north-east Tasmania (Fig. 1). Surveying was stratified according to the five broad forest types occurring within the study area: mature wet eucalypt forest, mixed forest and rainforest, mature dry eucalypt forest, regenerated wet eucalypt forest after clearfelling, and plantation. At least five different geographic locations within each forest type were selected within or immediately adjacent to the range of the study species *H. simsoni* (total of 42 locations). Locations were selected to cover the widest geographic range that would ensure sampling of all combinations of environments within the potential range of the species. At each location for a particular forest type, six sites were selected (total of 252 sites) covering the range of topography (i.e., gully/flat, mid-slope, and ridge-top), different aspects, slopes, and proximity to streams, present within a location. Where these attributes were relatively consistent within a location, sites were located greater than 100 m from one another. At a site, six 1 m^2 plots were placed haphazardly within a 10 m radius circle, thus ensuring all potential microhabitats were sampled. The plots were systematically searched by hand for live *Hoplogonus* specimens and body parts of dead ones. Identifiable body parts included male heads, female heads with thorax attached, and the thorax and abdomen of both sexes, which have distinctive humeral spines (Bartolozzi, 1996).

Two measures of beetle abundance for each site were estimated from the data: beetle density (no. of individuals/m^2 – calculated from the minimum number of *H. simsoni* known to have been alive from dead parts and live individuals in each plot); and the frequency of occurrence of beetles at each site (calculated as the proportion of plots in which beetles were found). Habitat variables recorded at each site were chosen for their anticipated value as predictors of beetle distribution and abundance, and for the ease with which they could be collected. The following habitat variables were assessed at each site: floristic composition and structure, altitude (m), distance to nearest stream (<30 m; 30–100 m; >100 m), leaf litter depth (<1 cm; 1–3 cm; >3 cm), leaf litter cover (% ground cover), rock cover (very low; low; medium; high), dead wood cover (% ground cover of logs >10 cm mid-diameter), moss cover (% ground cover including on rocks and logs), average aspect (N, S, E, W, none), average slope (degrees), distance to nearest road (m), weeds (present/absent), and soil characteristics. Spatial information for only a sub-set of these habitat variables was available in Forestry Tasmania's GIS. Therefore the following habitat variables were used to develop spatially explicit habitat models: Overstorey tree height (m); Distance to stream (< or >100 m);

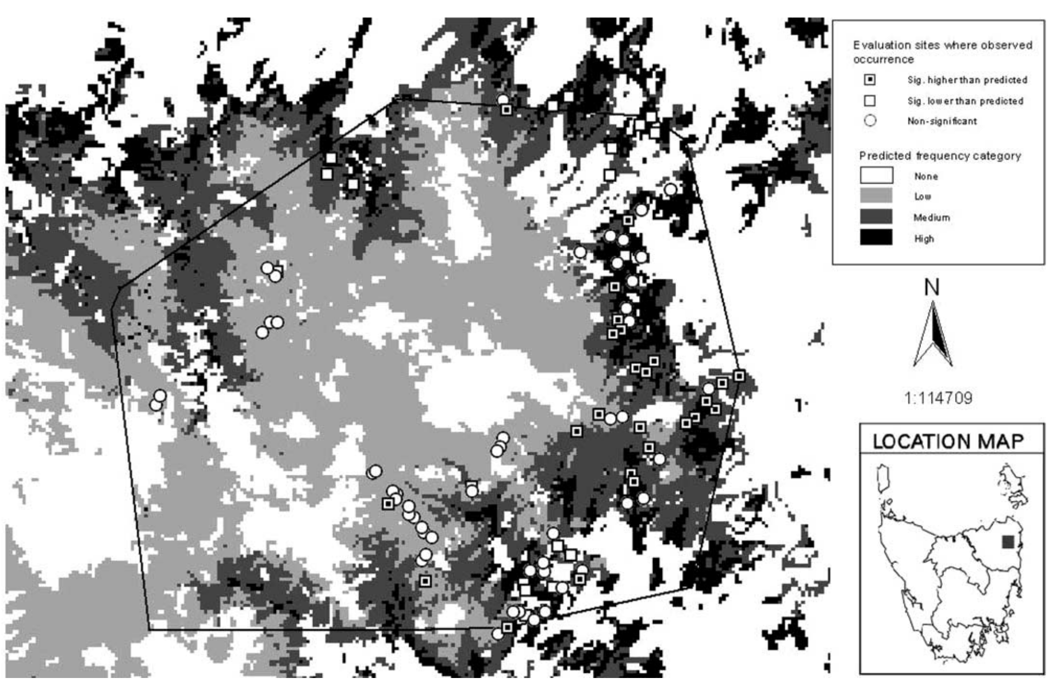

Fig. 1. The predicted frequency of occurrence of *H. simsoni* within and adjacent to its 250 km^2 range (solid line) and the location of the evaluation sites, indicating sites where observed values significantly differed from predicted frequencies at $p < 0.05$.

Altitude (in 10 m intervals); Aspect (divided into N, S, E, W, and none); and Slope (degrees).

2.2. Model development

The relationship between the presence and abundance of *H. simsoni* and the measured habitat variables for which GIS data were available was examined using generalised linear modelling (GLM) (McCullagh and Nelder, 1989). Only wet eucalypt forest and mixed/rainforest were modelled because Meggs et al. (2003) found that they constituted suitable habitat for *H. simsoni*. These forest types were modelled separately because they were considered distinct ecologically, with the majority of rainforest limited to higher altitudes within the study area. Meggs et al. (2003) found that the abundance and presence of beetles differed between the two, with significantly higher numbers occurring in wet eucalypt forest. Some habitat variables needed logarithm transformation to decrease excessive skewness. This was done in order to make non-linear relationships linear and hence simplify the models. These variables are indicated in the results by the prefix 'L'. Models were constructed using the two measures of beetle abundance derived from the animal survey data as described by Meggs et al. (2003): beetle density and the frequency of occurrence of beetles at each site. A Poisson distribution and log-link function was assumed for the measure of beetle density and a binomial distribution and logit-link function was assumed for the measure of frequency of occurrence. The probability of beetle presence/absence was also modelled (assuming a binomial distribution and logit-link function). A stepwise-forward fitting process was used in which variables were added to the model if the subsequent deviance change was sufficiently large. A large deviance change was chosen so that all variables that were finally included in the model were highly significant. Hence, the criterion for entry of a habitat variable to the model was a level of statistical significance of at least 0.01. This had the effect of simplifying the model. Only two-way interactions were fitted, again to keep the model relatively simple.

2.3. Predictive map development

The models for each measure of beetle occurrence (density, frequency of occurrence and presence/absence) were coupled with GIS data for each 100 m (1 ha) grid square within an area encompassing the range of *H. simsoni* to produce three maps of the predicted distribution and/or abundance of *H. simsoni*. A one-hectare grid square scale was selected to match the resolution of the digital elevation model (DEM) used. GIS data for the significant habitat variables were extracted for the two broad forest types that constitute potential habitat for the species, wet eucalypt forest and mixed/rainforest. These forest types were identified from photographic interpretation (PI) codes (Stone, 1998) that categorise patches of forest according to the height and density of the tree layer, and broad type of understorey. Altitude, aspect and slope data for each 100 m grid cell were extracted from a 100 m DEM. Overstorey tree height was obtained from the height-class of the matching PI-code. Distance from the nearest stream was obtained by buffering a drainage coverage, and coding areas within 100 m of a stream or lake. The data that were initially grid were based on a cell size of 25 m. Where both vector and grid coverages required preparation, this was performed in their original state and then converted into the final grid requirements of 100 m cell size.

To simplify interpretation of the final maps derived from the three models, predictions for each grid cell were grouped into categories representing high, medium, low or no occurrence of the beetle. The groupings were based on the relatively low frequency of cells predicted as high occurrence, and the high frequency of cells predicted as low occurrence. For the density map these categories were $>3/\text{m}^2$ (high), 1–3/m^2 (medium), $<1/\text{m}^2$ (low), and Absent (or not predicted). The categories for the frequency of occurrence map were 0.5–1.0 (high), 0.2–0.5 (medium), <0.2 (low) and absent. The predicted probability of beetle presence was grouped into the following categories: 90–100% probability (high), 70–90% (medium), and 50–70% (low). Where the predicted probability of presence was less than 50% the beetle was considered to be absent. This is a generally recognised cut-off for predictions of species' presence/absence (e.g. Fielding and Haworth, 1995; Manel et al., 1999; Fleishman et al., 2001; Clevenger et al., 2002).

2.4. Field methods for the collection of model evaluation data

Preliminary analysis of the three models, comparing the predicted values for each grid cell with the original observed field data, indicated that the predictions of the frequency of occurrence model best reflected the original abundance data ($r = 0.41$; $p < 0.05$; $n = 153$). Therefore, this model was chosen as the basis for the stratification of the field survey to evaluate the models' predictive ability. At least 30 new survey sites within each of the high, medium and low categories of frequency of occurrence (excluding predictions of absence) were sampled, 95 sites in total (Fig. 1). Sites were selected to cover as wide a geographic range as possible within each predicted category. Where this was not logistically possible, sites were located at least 100 m from one another. All sites were located at least 30 m from roads, paddocks or any disturbed habitat. The same animal sampling method as employed in the original survey (Meggs et al., 2003) was used.

2.5. Model evaluation analyses

A variety of analyses were used to compare predicted and observed occurrences of *H. simsoni* at the evaluation sites. Chi-square analysis in GENSTAT (Payne et al., 1993) was used to compare predicted and observed density and frequency of occurrence of beetles. Another measure of a GLM's predictive ability is the Wilcoxon–Mann–Whitney two-sample rank test (Harrell, 2000). This test measures the concordance between predictions and actual presence/absence data. Hence, it was used to compare the predicted and observed probabilities of beetle presence, and the frequency of occurrence (converted to presence/absence data). The Wilcoxon–Mann–Whitney test is equivalent to the area under the "receiver operating characteristic" or ROC curve (Hanley and McNeil, 1982). A test statistic of 0.5 indicates random predictions, whilst a value of 1.0 indicates perfect prediction. A test statistic greater than 0.75 demonstrates model utility. Predicted and observed values for all three measures of beetle occurrence were also compared using simple linear correlation (*Pearson r*).

3. Results

3.1. Predicted distribution of H. simsoni and the relationship between its occurrence and measured habitat variables

Fig. 1 illustrates the map derived from the relationship between the frequency of occurrence of the beetle and habitat variables measured. This map and those developed from the other measures of beetle occurrence indicated that *H. simsoni* has a "ring" or "doughnut"-shaped distribution, with the large "hole" in the centre corresponding to high altitude areas of the Blue Tier (Fig. 1). High-density populations are predicted in a 5–10 km-wide band in the eastern part of the species' range, with relatively isolated populations in the west and north of its range.

Altitude was a significant predictor of the density, frequency of occurrence and presence/absence of *H. simsoni*, with slope and aspect also found to be significantly influencing the density and frequency of occurrence of *H. simsoni*. The height of the overstorey tree layer and proximity to the nearest stream were also significant predictors of the density of *H. simsoni*.

3.1.1. Relative density model

The model constructed for the density of *H. simsoni* per site for wet eucalypt forest consisted of the following set of variables that together best explained the data collected:

$$\begin{aligned}\text{Beetle density} &= \text{constant} + \text{Altitude} + \text{Aspect} + \text{LSlope}\\ &\quad + \text{Distance from a stream(RipD)}\\ &\quad + \text{Overstorey tree height(OSTH)}\\ &\quad + \text{Alt} \times \text{Aspect} + \text{LSlope} \times \text{Aspect}\\ &\quad + \text{OSTH} \times \text{RipD}.\end{aligned}$$

The density of *H. simsoni* decreased with increasing altitude at all aspects, but particularly quickly at sites with southerly and northerly aspects (Table 1). It also decreased with increasing slope at northerly and easterly aspects, but there was no significant trend with slope at other aspects. The few sites at which the beetle was found in extraordinarily high numbers were flat and hence had no particular aspect. However, these sites did not appear to exert a strong influence on the overall trends. For sites less than 100 m from streams, beetle density increased significantly with increasing overstorey tree height (Table 1). There was no significant trend at distances greater than 100 m from streams.

At mixed forest/rainforest sites the significant habitat variables that were influencing the density of *H. simsoni* were:

$$\text{Beetle density} = \text{Constant} + \text{Altitude} + \text{Aspect}.$$

In the mixed forest and rainforest sampled, beetle density again decreased with increasing altitude at all aspects (Table 1). In addition, beetles consistently occurred at higher densities on slopes with southerly and northerly aspects.

3.1.2. Frequency of occurrence model

The habitat variables that had a significant influence on the frequency of occurrence of *H. simsoni* in wet eucalypt forest were:

$$\text{Frequency of occurrence} = \text{Constant} + \text{Altitude} + \text{Aspect} + \text{LSlope}.$$

The proportion of plots with *H. simsoni* within a site decreased with increasing altitude at all aspects (Table 2). There was an apparent trend whereby beetles occurred with a greater frequency at sites with southerly and easterly aspects, but this difference was not statistically significant. There was also a significant trend for the frequency of occurrence of beetles to decrease with increasing slope at all aspects and altitudes.

In the mixed forest and rainforest surveyed only altitude appeared to be influencing the frequency of occurrence of *H. simsoni*. The model took the form:

$$\text{Frequency of occurrence} = 0.512 - 0.004877 \times \text{Altitude}.$$

Again, beetles were found with less frequency as altitude increased ($t = -5.46$). Sites above 500 m where

Table 1
Estimates of the regression coefficients for the GLM that best explained the density of *H. simsoni* in the wet eucalypt forest and mixed/rainforest surveyed

Habitat variable	Estimate	SE	*t* (*)
Wet eucalypt forest			
Constant	−0.05400	0.64500	−0.08
Altitude	−0.00165	0.00139	−1.18
Aspect E	0.29100	0.49800	0.58
Aspect S*	1.45400	0.49700	2.93
Aspect N*	3.52900	0.81600	4.32
Aspect none	0.45400	0.56300	0.81
Stream distance > 100 m	5.38700	0.54900	9.81
Altitude × Stream distance > 100 m*	−0.01091	0.00123	−8.88
Altitude × E*	0.00482	0.00148	3.25
Altitude × S	−0.00297	0.00163	−1.82
Altitude × N*	−0.00474	0.00196	−2.41
Altitude × Aspect none	0.00117	0.00190	0.62
LSlope	0.12100	0.10800	1.12
LSlope × E*	−0.63500	0.14300	−4.45
LSlope × S	−0.22800	0.13700	−1.66
LSlope × N*	−1.48400	0.23900	−6.22
LSlope × None	–	–	–
Overstorey tree height*	0.06120	0.01130	5.42
Overstorey tree height × Stream distance > 100 m*	−0.05760	0.01220	−4.72
Mixed/rainforest			
Constant	3.432000	0.412000	8.32
Altitude*	−0.004886	0.000655	−7.46
Aspect W*	−0.878000	0.366000	−2.40
Aspect E*	−1.367000	0.376000	−3.63
Aspect N	−0.042000	0.312000	−0.14
Aspect none	−1.627000	0.753000	−2.16

Significant variables are indicated with an asterisk (SE = standard error; *t* = the *t*-statistic, used as a rough guide to test whether each of the factor levels differ from the first level. In this case, any *t*-statistic greater than 2.26 is approximately significant at the 0.01 level).

Table 2
Estimates of the regression coefficients for the GLM that best explained the frequency of occurrence of *H. simsoni* in the wet eucalypt forest surveyed

Habitat variable	Estimate	SE	*t* (*)
Constant	3.80200	0.53200	7.15
Altitude*	−0.01048	0.00107	−9.80
Aspect E*	0.61500	0.27500	2.23
Aspect S	0.55400	0.25900	2.14
Aspect N*	−1.15500	0.35000	−3.30
Aspect none	−0.33600	0.57800	−0.58
LSlope*	−0.60000	0.15600	−3.85

Significant variables are indicated with an asterisk (SE = standard error; *t* = the *t*-statistic, used as a rough guide to test whether each of the factor levels differ from the first level. In this case, any *t*-statistic greater than 2.26 is approximately significant at the 0.01 level).

the beetle was absent appear to have strongly influenced this trend.

3.1.3. Presence/absence model

Altitude was the only habitat variable found to influence significantly the probability of finding *H. simsoni* (i.e., the probability of the species' presence) in both wet eucalypt forest and mixed/rainforest. In both forest types the chance of finding the beetle greatly decreased with increasing altitude. The models constructed were

$$\text{Prob. beetle presence(wet eucalypt forest)} = 3.064 - 0.00827 \times \text{Altitude},$$

$$\text{Prob. beetle presence(mixed/rainforest)} = 3.620 - 0.00765 \times \text{Altitude}.$$

There was a very high likelihood of finding *H. simsoni* in wet eucalypt forest below 300 m altitude within its range. Sixty-two percent of sites in which the species was found occurred below this altitude compared to only 19% of absent sites. No wet eucalypt forest above 500 m was sampled.

3.2. Evaluation of the habitat models

3.2.1. Relative density model

The relative density model proved to be a poor predictor of the abundance of *H. simsoni*. Comparison of the predicted and observed densities of the evaluation sites revealed a Chi-square deviance of almost 30 times greater than would be expected for an acceptable model fit ($\chi^2_{94} = 2.842$). This indicates that the model is over-dispersed, with too much variation at a number of scales for the model to counter random site effects adequately. There was little correlation between predicted and observed densities ($r = 0.04$). In general, the model appeared to be significantly underestimating relative densities of *H. simsoni* (Fig. 2a).

3.2.2. Frequency of occurrence model

In terms of predicting the species presence or absence the frequency of occurrence model was found to have a moderate discriminatory ability with a Wilcoxon–Mann–Whitney statistic of 0.77. A direct comparison of predicted and observed frequencies of occurrence showed a significant correlation of marginal to moderate strength, with a correlation coefficient of 0.37 ($p < 0.05$) (Fig. 2b). Further analysis revealed a Chi-square deviance of five times greater than would be expected for an acceptable model fit ($\chi^2_{94} = 512$); indicating that this model was also over-dispersed.

Examination of the spatial distribution of the residuals of the Chi-square analysis (Fig. 1) indicated that the model was significantly over-estimating frequency in the southern and northern margins of the species range, and under-estimating in the largest area of quality habitat in the east of its range.

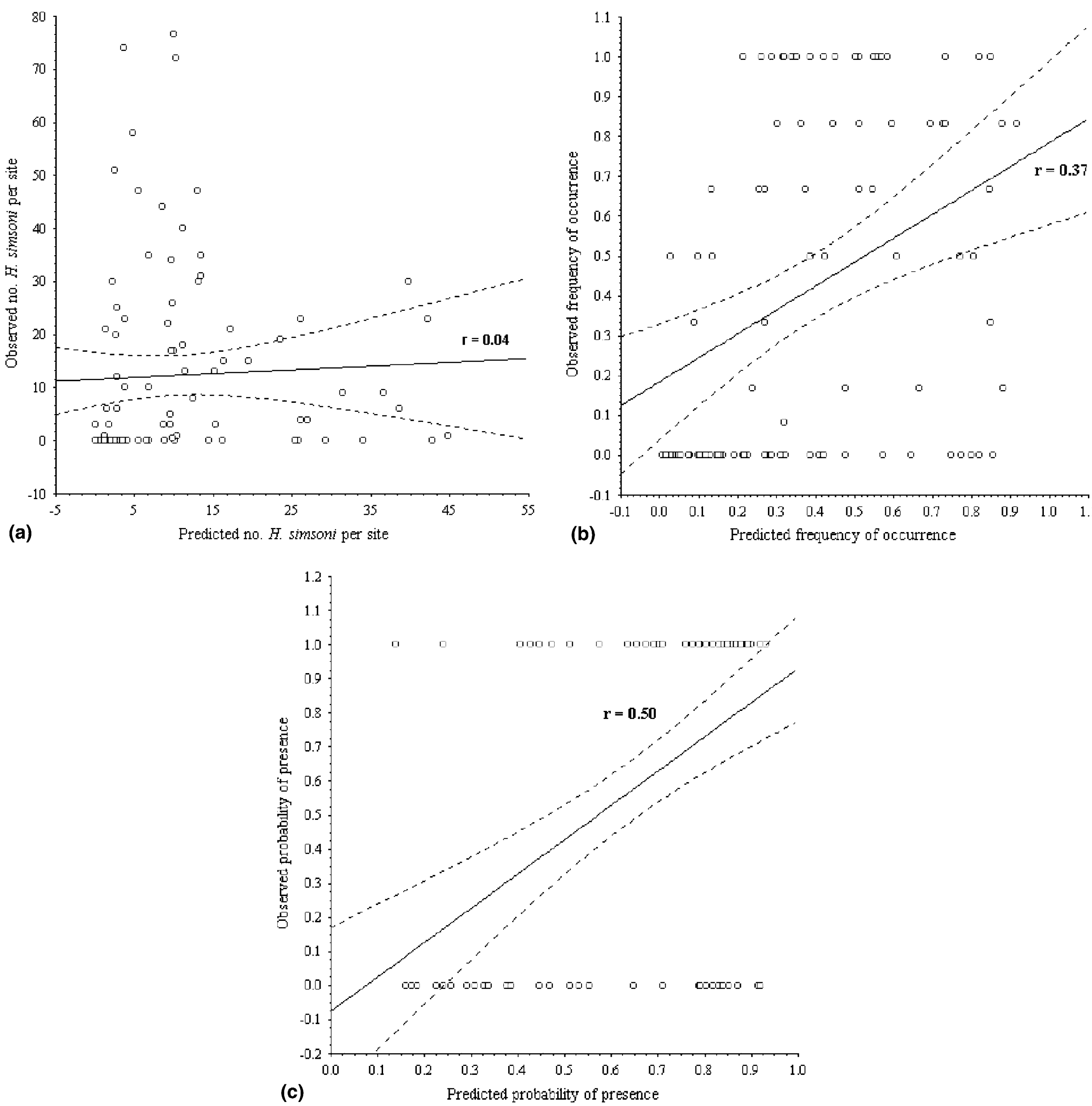

Fig. 2. Correlation between the predicted and observed values for (a) the density, (b) frequency of occurrence, and (c) probability of presence of *H. simsoni* at each of the 95 evaluation sites (dotted lines indicate 95% confidence intervals).

3.2.3. Presence/absence model

The presence/absence model showed a moderate discriminatory ability in predicting the probability of beetle presence with a Wilcoxon–Mann–Whitney statistic of 0.76, and a significant correlation coefficient of 0.50 ($p < 0.05$) (Fig. 2c). A comparison of the predicted probability of presence with observed density ($r = 0.42$; $p < 0.05$) and frequency of occurrence ($r = 0.52$; $p < 0.05$) indicated that this model had potential as an indirect predictor of species' abundance.

4. Discussion

It is critical for threatened species management that land managers have access to information on the extent and spatial distribution of habitats important to a species in order to assess the consequences of various land management options. Unfortunately, for most invertebrates, including threatened species, this information is not available. Exceptions include invertebrates with extremely restricted ranges such as the Eltham copper butterfly *Paralucia pyrodiscus lucida* (Vaughan, 1988), or species whose distribution is tied to a particular food plant such as the Ptunarra brown butterfly *Oreixenica ptunarra* (Neyland, 1992) and Checkerspot butterflies *Euphydryas* spp. (Ehrlich and Murphy, 1987). However, for more broadly ranging or ecologically opaque invertebrate species, techniques that couple statistical models with GIS may be powerful tools in enabling extrapolation of data of species' distribution and abundance beyond the sites sampled (Nicholls, 1989; Neave and Norton, 1991). Given the high stakes often involved in threatened species management, the use of an evaluation dataset independent of the model provides the most rigorous test of a model's predictive ability (Chatfield, 1995; Fielding and Bell, 1997; Guisan and Zimmermann, 2000). To our knowledge this is the first study to develop predictive distribution and abundance models for a threatened invertebrate that have then been field-checked and evaluated.

The three predictive models showed variable levels of performance. The density model proved to be a poor predictor of species abundance with little correlation between predicted and observed abundances. It was significantly over-dispersed; a problem consistently found with rigorous tests of GLMs (Pearce and Ferrier, 2001), particularly when GLMs are used to model the abundance of rare or cryptic species, resulting in zero-inflated count data (Welsh et al., 1996). The frequency of occurrence model was also over-dispersed, with too much variation at a number of scales for the model to counter random site effects adequately. The lack of fit of the abundance models may be a consequence of the indirect relationship between the species and the variables modelled and/or the non-inclusion of biologically important variables. Meggs et al. (2003) found that leaf litter depth and a well-developed tall shrub layer significantly influenced the abundance of *H. simsoni*. Spatial data for both of these variables did not exist in the GIS used, nor were adequate surrogate variables available in the GIS. The lack of fit of the models may also be due to the patchiness of the original survey (Meggs et al., 2003), and/or the naturally patchy distribution and abundance of the species (Meggs et al., 2003). The inclusion of spatial variables in the models, such as the grid coordinates of the sampling sites, may have provided some insight into the relationship between spatial biases and the models' predictive ability.

Nevertheless, a significant correlation of marginal to moderate strength was found in a comparison of predicted and observed frequencies of occurrence. This model also proved to have a moderate ability to predict the presence/absence of *H. simsoni*. A simple comparison of the accuracy of the model in predicting categories of species frequency of occurrence revealed that the model was stronger at predicting areas of high abundance. The model was found to predict high abundance accurately for 59% of the evaluation sites predicted as high (Fig. 3). When predictions of high and medium abundance categories were pooled this rose to 68% accuracy. The model performed poorly in predicting low abundance, with the beetle not found at over 80% of sites predicted as low. This may suggest that a minimum threshold of frequency of occurrence should have been applied or that the sampling intensity of the evaluation sites was insufficiently intensive to assess very low predictions of occurrence.

The presence/absence model also showed a moderate ability to predict accurately the occurrence of *H. simsoni*. Interestingly, it also exhibited potential as an indirect predictor of species abundance. Pearce and Ferrier (2001) have suggested that more effort should be directed to the collection and modelling of presence/absence data rather than abundance data. They pro-

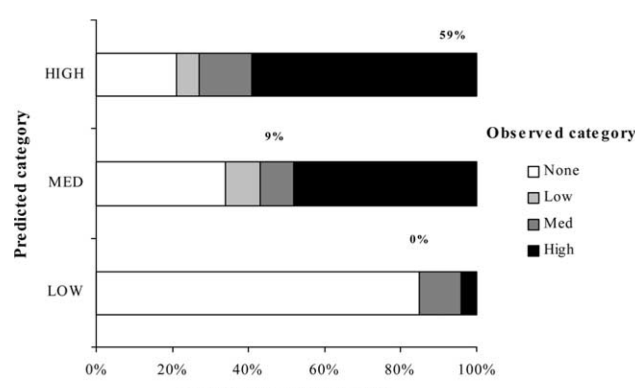

Fig. 3. Comparison of predicted and observed categories of frequency of occurrence of *H. simsoni*. The figures on the bars indicate the percentage of sites where the predicted and observed category matched.

duced reasonably accurate models for 12 of the 44 species of small reptiles, arboreal marsupials, vascular plants and diurnal birds evaluated and found that for all 12 of the species, predictions from direct abundance models performed no better as a relative measure of abundance than predictions from presence/absence models. The results of this study would appear to support this finding but whether this is the case for other invertebrates requires further investigation.

It is essential that the evaluation of the predictive performance of a species model be done within the context of the original aims of the study (Rykiel, 1996; Fielding and Bell, 1997). Hence, it has been recommended that a range of criteria be used to evaluate the performance of a model (Manel et al., 1999; Guisan and Zimmermann, 2000). In this study, we aimed to predict the distribution of habitats important for a threatened species for which we have limited ecological information. Of particular importance was to produce a model to predict the spatial distribution and extent of high-abundance populations of *H. simsoni* and/or optimal habitat, as there appeared to be a high correlation between the occurrence of high-density populations and forest with a high potential for plantation establishment (Meggs et al., 2003). Whilst the presence/absence model exhibited some potential as an indirect predictor of the abundance of *H. simsoni*, the results of this study suggest that the frequency of occurrence predictive model should be used by land-managers in the development of any conservation management plan for the species. The frequency of occurrence model contains a direct, albeit conservative, measure of species abundance and has shown a moderate ability to predict species abundance. In addition, it was a relatively accurate predictor of areas of high beetle abundance.

This study represents a reasonably successful transfer of conservation biology techniques developed predominantly for vertebrates to an invertebrate of conservation significance. This was unexpected since many predictive distribution and abundance models developed for vertebrates have been plagued by problems of scale (e.g. Fielding and Haworth, 1995; Lindenmayer et al., 1995; Cork and Catling, 1996), and invertebrates generally operate at a small spatial scale. *H. simsoni*, for example, is believed to have a dispersal ability of only 100–200 m in an individual's lifetime (G. Bornemissza and P. McQuillan, pers. comms.). These problems of scale may be overcome by designing a sampling regime at a scale appropriate to the organism under investigation (Horne and Schneider, 1995) and also at a spatial scale at which decisions are to be implemented (Elith, 2000). The sampling design used to collect the data from which the predictive models for *H. simsoni* were built incorporated stratification on the basis of both environmental variability and spatial scale (Meggs et al., 2003). Hence, replicate forest types were sampled over as wide a geographic area as logistically possible, and clusters of sites were sampled within each replicate forest type. This allowed measurement of the species' abundance and habitat variability at both a landscape scale (relevant to the scale at which forest management is conducted) and a local scale (relevant to the ecology of the species).

Ecological explanations for the success of variables such as altitude, aspect and slope in predicting the distribution and abundance of *H. simsoni* are uncertain. Meggs et al. (2003) suggested that the characteristics of optimal habitat for the beetle related to a requirement of the species for a cool, moist stable microclimate and the absence of disturbance for some time. Altitude and aspect are likely to play a significant role in the local distribution of microclimates. Fleishman et al. (2001) found altitude to be the most common predictor in 20 of the 36 significant species distribution models developed for butterflies in the Great Basin of western North America. The importance of gentle slopes to the beetle may relate to the influence of this variable on the development and maintenance of a deep leaf litter layer and particular soil characteristics. The fine-scale habitat modelling for *H. simsoni* conducted by Meggs et al. (2003) indicated that the abundance of the species was greater in forest with a deep leaf litter layer. Although based on limited information, Meggs et al. (2003) also suggested that a deep, well-drained soil may be important for the species. These soil characteristics were present at sites with a flat topography and were associated with high beetle densities (J. Meggs, pers. obs.). It has been argued that a successful model used to predict the distribution of a species can serve as a valuable planning tool regardless of whether or not ecological explanations are apparent (Fleishman et al., 2001). However, studies have shown that improved modelling success is associated with the use of variables with a known direct or causal effect on a species distribution (Austin and Myers, 1995). Establishing causal relationships between the presence and abundance of *H. simsoni* and particular habitat variables or groups of habitat variables remains to be done.

The intensification of forestry activities in 'off-reserve' forested areas recently increased following Tasmania's *Regional Forest Agreement* (Commonwealth of Australia and State of Tasmania, 1997), resulting in an upsurge in clearing of the native forests in the north-east of the State for conversion to plantations (Munks and McArthur, 2001; Lindenmayer and Franklin, 2002). Meggs et al. (2003) found that the conversion of potential wet forest habitat to plantation results in the local elimination of *H. simsoni*. They also suggested that whilst clearfell, burn and sow silvicultural regimes are likely to have a significant negative impact on populations of the beetle, recovery of beetle populations may occur within a standard 80–90 year harvest rotation, with recolonisation of disturbed habitat through immigration from adjacent undisturbed forest. This information on potential threats and conservation requirements, together with knowledge of the fine-scale habitat requirements of the species (Meggs et al., 2003), can be coupled with the landscape-scale map of predicted habitat derived from the frequency of occurrence model to inform the development of a conservation management strategy for *H. simsoni*.

5. Conclusions

Model accuracy and model utility are not the same thing. It is important to remember the original goals of a study in choosing model evaluation methods. In this study, the aim was to identify the extent and spatial arrangement of habitats important to this species and the areas in which conservation management of the species may conflict with planned forestry activities. Hence, a range of criteria were used to assess the models' predictive ability. Statistically, the frequency of occurrence model performed best as a simple presence/absence predictor, but as such has much less utility for conservation management than as a predictor of abundance (or habitat quality). It was clear from the original sampling (Meggs et al., 2003), that there were definite "hotspots" of abundance – areas where there was a magnitude greater abundance of *H. simsoni*, and the evaluation analysis has shown that the model performs best in predicting the occurrence of such areas. It is for this reason that we conclude that whilst the frequency of occurrence model had only exhibited a moderate level of statistical strength, it has a high level of utility as a landscape-scale conservation planning tool for *H. simsoni*. The usefulness of a model in planning the conservation of a species and its habitat should not necessarily be discounted if the model is found to be inaccurate according to some criteria. In the absence of information on many aspects of the biology and ecology of *H. simsoni*, it is proposed that land managers can confidently utilise the predictive habitat map derived from the frequency of occurrence model in this study as one tool to inform the development of conservation measures throughout the species' range.

Acknowledgements

This work was funded by the W.V. Scott Bequest Fund and Forestry Tasmania, with additional in-kind support from the Forest Practices Board and Eastern Tiers and Bass Districts. Field assistance was provided by Nathan Duhig and Beris Hansberry. Thanks to Steve Candy and Julian Fox for advice on the model evaluation analyses. The GIS analyses were conducted by Luke Ellis and Ian Harvey. Alastair Richardson, Peter McQuillan, Mick Brown, Graham Wilkinson, and two anonymous referees provided useful feedback on earlier drafts of this manuscript.

References

Austin, M.P., Myers, J.A., 1995. Modelling of landscape patterns and processes using biological data. In: Subproject 4: Real data case study. CSIRO Division of Wildlife & Ecology, Canberra.

Bartolozzi, L., 1996. Description of a new species of *Hoplogonus* Parry, 1875 from Tasmania (Coleoptera Lucanidae). Redia 79, 91–95.

Bustamante, J., 1997. Predictive models for lesser kestrel *Falco naumanni* distribution, abundance and extinction in southern Spain. Biol. Cons. 80, 153–160.

Chatfield, C., 1995. Model uncertainty, data mining and statistical inference. J. Roy. Stat. Soc. Ser. A 158, 419–466.

Clevenger, A.P., Wierzchowski, J., Chruszcz, B., Gunson, K., 2002. GIS-generated, expert-based models for identifying wildlife habitat linkages and planning mitigation passages. Conserv. Biol. 16 (2), 503–514.

Commonwealth of Australia, State of Tasmania, 1997. Tasmanian Regional Forest Agreement. Commonwealth of Australia & State of Tasmania, Hobart.

Cork, S.J., Catling, P.C., 1996. Modelling distributions of arboreal and ground-dwelling mammals in relation to climate, nutrients, plant chemical defences and vegetation structure in the eucalypt forests of southeastern Australia. For. Ecol. Mgmt. 85, 163–175.

Corsi, F., Dupre, E., Boitani, L., 1999. A large-scale model of wolf distribution in Italy for conservation planning. Conserv. Biol. 13 (1), 150–159.

Ehrlich, P.R., Murphy, D.D., 1987. Conservation lessons from long-term studies of checkerspot butterflies. Conserv. Biol. 1 (2), 122–131.

Elith, J., 2000. Quantitative methods for modeling species habitat: comparative performance and an application to Australian plants. In: Ferson, S., Burgman, M.A. (Eds.), Quantitative Methods in Conservation Biology. Springer-Verlag, New York, pp. 39–58.

Elith, J., Burgman, M.A., 2002. Predictions and their validation: rare plants in the Central Highlands, Victoria, Australia. In: Scott, J.M., Heglund, P.J., Morrison, M.L., Raphael, M.G., Wall, W.A., Samson, F. (Eds.), Predicting species occurrences: issues of accuracy and scale. Island Press, Covelo, CA, pp. 303–314.

Fielding, A.H., Bell, J.F., 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environ. Conserv. 24 (1), 38–49.

Fielding, A.H., Haworth, P.F., 1995. Testing the generality of bird-habitat models. Conserv. Biol. 9 (6), 1466–1481.

Fleishman, E., Mac Nally, R., Fay, J.P., Murphy, D.D., 2001. Modeling and predicting species occurrence using broad-scale environmental variables: an example with butterflies of the Great Basin. Conserv. Biol. 15 (6), 1674–1685.

Guisan, A., Zimmermann, N.E., 2000. Predictive habitat distribution models in ecology. Ecol. Modell. 135, 147–186.

Hanley, J.A., McNeil, B.J., 1982. The meaning and use of the area under the receiver operating characteristic (ROC) curve. Radiology 143, 29–36.

Harrell, F.E., 2000. Regression Modelling Strategies with Applications to Survival Analysis and Logistic Regression. University of Virginia.

Horne, J.K., Schneider, D.C., 1995. Spatial variance in ecology. Oikos 74, 18–26.

Lindenmayer, D.B., Franklin, J.F., 2002. Conserving Forest Biodiversity: A Comprehensive Multiscaled Approach. Island Press, Sydney.

Lindenmayer, D.B., Ritman, K., Cunningham, R.B., Smith, J.D.B., Horvath, D., 1995. A method for predicting the spatial distribution of arboreal marsupials. Wildl. Res. 22, 445–456.

Lindenmayer, D.B., Cunningham, R.B., McCarthy, M.A., 1999. The conservation of arboreal marsupials in the montane ash forests of the central highlands of Victoria, south-eastern Australia. VIII. Landscape analysis of the occurrence of arboreal marsupials. Biol. Cons. 89, 83–92.

Mace, R.D., Waller, J.S., Manley, T.L., Ake, K., Wittinger, W.T., 1999. Landscape evaluation of Grizzly bear habitat in Western Montana. Conserv. Biol. 13 (2), 367–377.

Manel, S., Dias, J., Ormerod, S.J., 1999. Comparing discriminant analysis, neural networks and logistic regression for predicting species distributions: a case study with a Himalayan river bird. Ecol. Modell. 120, 337–347.

McCullagh, P., Nelder, J.A., 1989. Generalised Linear Models, Second ed. Chapman & Hall, London.

Meggs, J.M., Munks, S.A., Corkrey, R., 2003. The distribution and habitat characteristics of a threatened lucanid beetle, *Hoplogonus simsoni*, in north-east Tasmania. Pac. Cons. Biol. 9 (3).

Mladenoff, D.L., Sickley, T.A., Haight, R.G., Wydeven, A.P., 1995. A regional landscape analysis and prediction of favourable Gray Wolf habitat in the northern Great Lakes region. Conserv. Biol. 9 (2), 279–294.

Munks, S.A., 1993. Arboreal mammals in the Prairie-Torrens Creek Alluvials Province of the Desert Uplands: distribution and habitat preference. Qld. Dept. Environment & Heritage, Townsville.

Munks, S.A., McArthur, C., 2001. Plantation design and fauna conservation in Tasmania. Part 1. Workshop abstracts. Tasforests 12, 161–172.

Munks, S.A., Taylor, R.J., 2000. Conserving threatened fauna in production forests: the Tasmanian process. In: Craig, J.L., Mitchell, N., Saunders, D.A. (Eds.), Nature Conservation 5: Nature Conservation in Production Environments: Managing the Matrix. Surrey, Beatty & Sons, Chipping Norton, pp. 670–674.

Murphy, D.D., Freas, K.E., Weiss, S.B., 1990. An environment-metapopulation approach to population viability analysis for a threatened invertebrate. Conserv. Biol. 4 (1), 41–51.

Neave, H.M., Norton, T.W., 1991. Integrated management of forest wildlife: comments on new ways to research habitat. In: Lunney, D. (Ed.), Conservation of Australia's Forest Fauna. Royal Zoological Society of NSW, Mosman, pp. 229–236.

Neyland, M.G., 1992. The ptunarra brown butterfly Oreixenica ptunarra. Conservation Research Statement. Dept. Parks, Wildlife and Heritage, Tasmania, Hobart.

Nicholls, A.O., 1989. How to make biological surveys go further with generalised linear models. Biol. Cons. 50, 51–75.

Payne, R.W., Arnold, G.M., Morgan, G.W. 1993. Genstat 5 Procedure Library Manual, Release 3[1]. Lawes Agricultural Trust, Rothamsted Experimental Station.

Pearce, J., Ferrier, S., 2001. The practical value of modelling relative abundance of species for regional conservation planning: a case study. Biol. Cons. 98, 33–43.

Rushton, S.P., Wadsworth, R.A., Cherrill, A.J., Eyre, M.D., Luff, M.L., 1994. Modelling the consequences of land use change on the distribution of Carabidae. In: Desender, K., Dufrene, M., Loreau, M., Luff, M.L., Maelfait, J.-P. (Eds.), Carabid Beetles: Ecology and Evolution. Kluwer Academic, Dordrecht, pp. 353–360.

Rykiel Jr., E.J, 1996. Testing ecological models: the meaning of validation. Ecol. Modell. 90, 229–244.

Smith, A.P., Horning, N., Moore, D., 1997. Regional biodiversity planning and lemur conservation with GIS in western Madagascar. Conserv. Biol. 11 (2), 498–512.

Stone, M.G., 1998. Forest-type mapping by photo-interpretation: a multi-purpose base for Tasmania's forest management. Tasforests 10, 15–32.

Taylor, R.J., Bryant, S.L., 1997. Compilation of a list of threatened invertebrates: the Tasmanian experience. Mem. Mus. Vic. 56 (2), 605–609.

Van Manen, F.T., Pelton, M.R., 1997. A GIS model to predict black bear habitat use. J. Forestry 95 (8), 6–12.

Vaughan, P.J., 1988. Management Plan for the Eltham Copper Butterfly (*Paralucia pyrodiscus lucida* Crosby) (Lepidoptera: Lycaenidae). National Parks and Wildlife Division, Dept. Conservation, Forests and Lands, Victoria, Melbourne.

Welsh, A.H., Cunningham, R.B., Donnelly, C.F., Lindenmayer, D.B., 1996. Modelling the abundance of rare species: statistical models for counts with extra zeros. Ecol. Modell. 88, 297–308.

Yen, A.L., New, T.R., Van Praagh, B.D., Vaughan, P.J., 1990. Invertebrate conservation: three case studies in south-eastern Australia. In: Clarke, T.W., Seebeck, J.H. (Eds.), Management and Conservation of Small Populations. Chicago Zoological Society, Illinois, pp. 207–224.

York, A., 1999. Ecologically sustainable management: the utility of habitat surrogates for assessing terrestrial invertebrate diversity in temperate forests. In: Ponder, W., Lunney, D. (Eds.), The Other 99%. The Conservation and Biodiversity of Invertebrates. Transactions of the Royal Zoological Society of New South Wales, Mosman, pp. 34–39.

Zimmermann, N.E., Kienast, F., 1999. Predictive mapping of alpine grasslands in Switzerland: species versus community approach. J. Veg. Sci. 10, 469–482.