

## **Abstract**

Species distribution modeling has received accelerating popularity in ecology and is becoming an increasingly potent technique in this field. However, this method is infrequently applied to investigate small-scale patterns of biodiversity or to model community composition where inferences could be applied to effective and meaningful conservation actions.

In this study, Coleoptera community dissimilarity and diversity patterns in response to ecological gradients were investigated using an extensive dataset obtained from a fragmented New Zealand landscape in combination with GIS and species distribution modeling software. Results were utilized to investigate the potential of modeling the responses of individual species to environmental variables prior to community composition aggregation in order to explain observed patterns of alpha and beta diversity.

Modeled community structure was incongruent with observed assemblages based on the Bray-Curtis index of dissimilarity ( $p < 0.05$ ,  $r = 0.5$ ), exhibiting only a moderated correlation regarding dissimilarity between pristine forest controls and fragmented areas in comparison to the observed data ( $p < 0.05$ ,  $r = 0.7$ ). Models which had occurrence probabilities constrained to a threshold, predicted similar diversity patterns to the observed data in response to environmental gradients which were previously reported as strong drivers in the study area. However, model occurrence probability predictions in their unconstrained entirety showed an opposite trend to the observed data with regards to absolute species richness.

Overall the models had a general tendency to significantly over-predict alpha diversity and under-estimate beta diversity in comparison to the observed data. Possible mechanisms responsible for these results and potential mitigation methods were discussed and evaluated. To summarize, conclusions were made regarding the transferability of community modeling to successful invertebrate conservation, planning and management in the face of environmental perturbations such as habitat fragmentation and climate change.

## **1.0 Introduction**

## *1.1 Forest Fragmentation*

The area and condition of the world's forest and the connectivity and fragmentation of ecosystems are key indicators for the Biological Diversity 2010 targets (Balmford et al., 2005) making the synergy of the two, the fragmentation of forest habitats, an essential ecological study system. The abiotic, the direct biotic and the indirect biotic effects of fragmentation (Murcia, 1995) interact with organisms in the ecosystem and generally have a detrimental effect on biodiversity (Davies et al., 2000; Didham et al., 1998a; Ferraz et al., 2003; Larsen et al., 2005; Murcia, 1995). Forest fragmentation has been shown to significantly affect insect community compositions, with different community assemblages apparent in the interior forest, edge and matrix habitats (Barbosa and Marquet, 2002; Davies et al., 2000; Didham et al., 1998a; Major et al., 2003).

The extent and manifestation of this negative effect of fragmentation is however, dependent on a variety of sensitivity traits exhibited by the species concerned (Gonzalez and Chaneton, 2002; Henle et al., 2004; Tschardt et al., 2002) with large (Larsen et al., 2005), rare (Davies et al., 2000) and higher trophic level insect species (Didham et al., 1998b) more likely to be negatively affected by fragmentation. These sensitivity traits interact with features of the environment such as the absolute size of the fragment and the season (Barbosa and Marquet, 2002) which can magnify or accelerate the impact of fragmentation. These results do however, vary to some extent between taxonomic groups (Watling and Donnelly, 2006) with conclusions drawn from investigations only site specific and not absolute (Murcia, 1995).

Biodiversity is an important factor for ecosystem functioning (Balvanera et al., 2006; Hooper et al. 2005; Loreau et al., 2001) with the loss of biodiversity proposed to have a detrimental effect on ecosystem functioning in forest systems (e.g. Larsen et al., 2005). In order to predict the consequences of habitat loss and fragmentation to ecosystem functioning, the individual responses of species and populations to perturbation need to be identified (Rosenlew and Roslin, 2008). Coleoptera in particular provide vital ecosystem roles in forest habitats especially regarding nutrient recycling (Andresen 2003; Larsen et al., 2005; Rosenlew and Roslin, 2008). The loss of large beetles can result in as much as a 90% decline in the ecosystem functions to which they contribute (Larsen et al., 2005). Beetles exhibit responses to forest edges over large spatial scales (Ewers and Didham, 2008) and a strong synergy between habitat edge and fragment area (Ewers et al., 2007). This means that more investigation into the responses of Coleoptera to habitat fragmentation is vital for

invertebrate conservation actions and management to be implemented especially in the face of environmental change.

### *1.2 Species distribution modeling*

Despite the accelerating demand to identify and conserve vulnerable populations, a major obstruction to conservation planning is the deficiency of absolute information regarding the distribution of biodiversity (Ferrier, 2002). This problem is particularly acute where areas of conservation value may be overlooked if survey data is bias or incomplete (Bojorquez-Tapia et al., 1995). This challenge can be addressed through species distribution modeling by combining observed occurrence location data with environmental variables to predict where a species may potentially be located (Soberon, 2005). The use of this technique has received accelerating popularity in ecology and is becoming increasingly potent in this field (Guisan and Thuiller, 2005; Guisan and Zimmermann, 2000). This technique has high utility for modeling the distribution of individual species (Buermann et al., 2008; Ferrier and Guisan, 2006), as well as for identifying biodiversity hotspots (Bojorquez-Tapia et al., 1995; Funk and Richardson, 2002) at a regional (Bojorquez-Tapia et al., 1995; Buermann et al., 2008) and continental scale (Buermann et al., 2008; Randin et al., 2006). Species distribution modeling is becoming a more frequently used tool in conservation planning (Wintle et al., 2005), as it can be used to: (1) predict biodiversity patterns where survey data is sparse (Ferrier and Guisan, 2006); (2) improve the sampling efficiency of rare species (Guisan et al., 2005; Singh et al., 2009); and (3) define which variables are affecting a species distribution before management actions are assigned (Rushton et al., 2005).

Despite the high utility and applicability of species distribution modeling, there are however, limitations to this technique (Guisan and Thuiller, 2005). There have been reports that some modeling techniques are not always accurate or appropriate (Beale et al., 2008), such that different techniques produce highly variable spatial patterns even when using the same underlying parameters (Thuiller, 2004). Furthermore, modeled distribution is often a coarse oversimplification of a species realized niche (Guisan and Thuiller, 2005) meaning there is weak spatial transferability of models (Randin et al., 2006). Consequently, unless the errors and constraints of species distribution modeling are made explicit, model results are not easily and directly applicable to real conservation management options (Pearce and Ferrier, 2000). Therefore, more critical evaluations regarding the inaccuracies and limitations of species distribution modeling are required in order to direct conservation actions (Loiselle et al, 2003; Wintle et al., 2005).

### *1.3 Modeling biological communities*

Many studies have aimed to model the distribution and abundance of individual species of conservation concern (e.g. Guisan et al., 2005; Loiselle et al., 2003; Singh et al., 2009), however, it has been suggested that it may be more beneficial to focus on the distribution of biodiversity as a whole at the community level as opposed to individual species (Ferrier, 2002; Ferrier and Guisan, 2006). This approach would use biological data and knowledge to its best purpose (Ferrier, 2002), and can include biodiversity measures such as local species richness (Bojorquez-Tapia et al., 1995) or differentiation diversity (Ferrier, 2002). Regions of high biodiversity are often data poor at a local scale where regional conservation is conducted (Ferrier, 2002). This means that community level modeling, especially if communities contain rare or sparsely recorded species may hold more utility for meaningful conservation actions to be implemented (Wintle et al., 2005). It has been suggested that initially modeling species individually and then aggregating the varied responses into an overall collective community model may be beneficial and appropriate when modeling community change in response to perturbation. This is because the idiosyncratic responses of individual species can be incorporated into the final community distribution model (Ferrier and Guisan, 2006; Guisan and Zimmermann, 2000).

### *1.4 Modeling with presence-only data*

There are many well established techniques for modeling the distribution of species which require data on the locations of occurrences along with reported absences in order to accurately distinguish between suitable and unsuitable conditions for the particular species or population of interest (Elith et al., 2006). Examples include; General Linear Modeling (GLM), General Additive Modeling (GAM) and Multivariate Adaptive Regression Splines (MARS). However, there has been growing inquiry and support for the use of presence-only data in distributional modeling due to its high prevalence (Funk and Richardson, 2002; Hirzel et al., 2006) and growing accessibility (Graham et al., 2004), especially as only well explored areas have true absence data (Soberon, 2005). Techniques requiring presence-only data such as the Bioclimatic prediction system (BIOCCLIM), Maxent (Maximum Entropy) and Genetic Algorithm for Rule-set Prediction (GARP), are becoming more commonly employed (evaluated by Elith et al., 2006). The use of these techniques especially Maxent have only recently been applied to ecology (Phillips et al., 2004). These presence-only techniques

have been shown to be reliable (Hirzel et al., 2006), highly accurate (Elith et al., 2006) and useful, especially with sparse records (Pearce and Boyce, 2006), with similar results reported to models that use both presence and absence data (Wintle et al., 2005).

### *1.5 Aims of the project*

In light of the potential application of species distribution modeling and the general acceptance of their validity, the proceeding investigation uses a large dataset on beetle communities in a fragmented New Zealand landscape to address the following questions:

Does modeling the individual responses of species to environmental variables, prior to community composition aggregation explain observed community-level diversity patterns?

Could presence-only species distribution modeling be utilized to predict observed patterns of alpha diversity in response to ecological gradients?

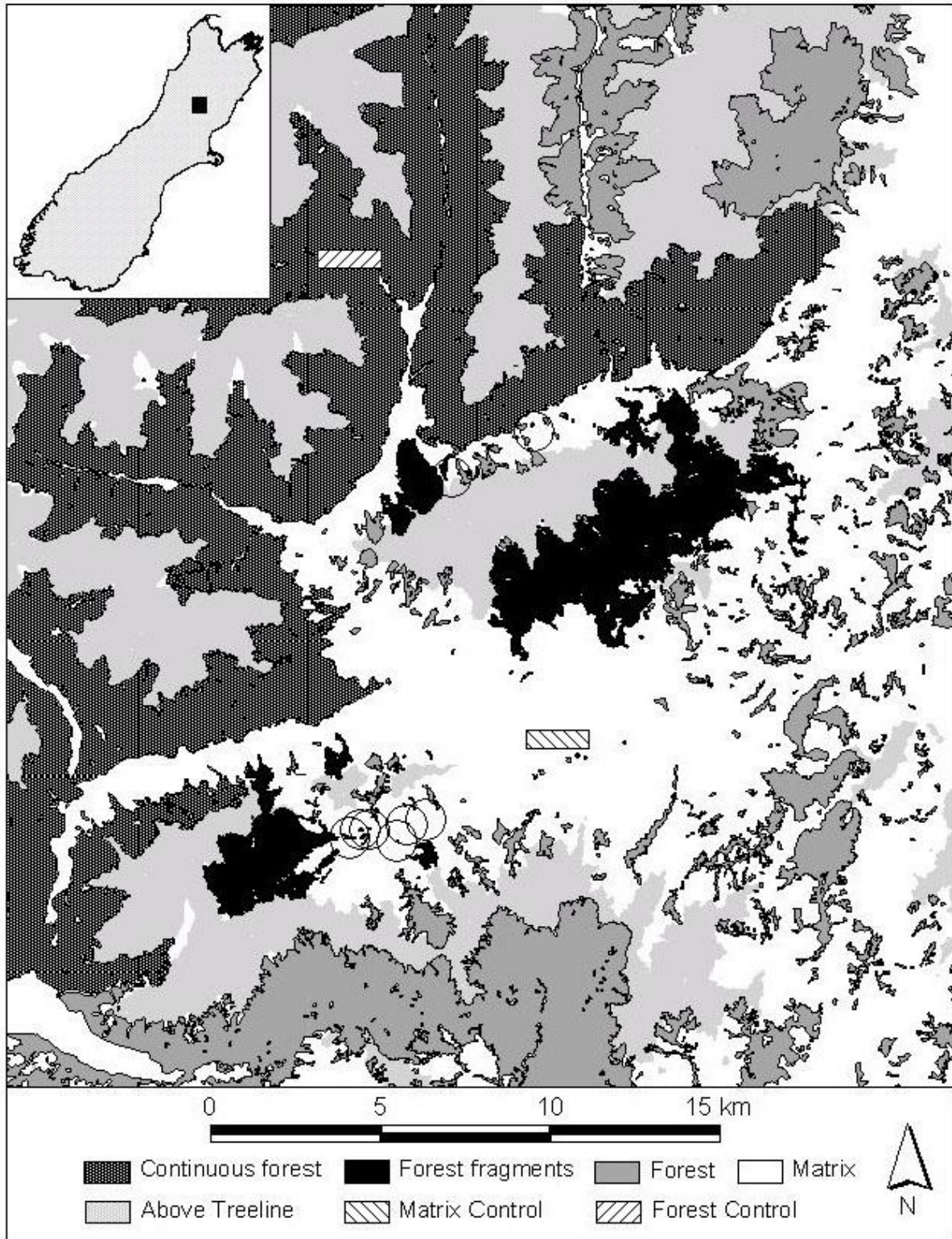
Are these techniques transferable to meaningful invertebrate conservation planning, actions and management in the face of environmental perturbations such as habitat fragmentation and climate change?

## **2.0 Methods**

## 2.1 Biological data

The dataset used for this analysis consisted of an extensive collection of Coleoptera obtained in 2000-2001 for the Hope River Forest Fragmentation Project (Fig 1) South Island, New Zealand, covering 1300km<sup>2</sup> of fragmented beech (*Nothofagus* spp.) forest (Ewers et al., 2007). Invertebrate traps dug into the ground acting as both a pitfall and a low level flight interception trap were situated at 14 (Table 1) forest fragment locations set at a maximum of 11 distances from forest edges into interiors (0,-2,-4,-8,-16,-32,-64,-128,-256,-512 and -1024 meters) and from forest edges into the pasture (0,+2,+4,+8,+16,+32,+64,+128,+256,+256 and +1024 meters). The larger distances were not sampled if the fragments were not large enough to sample to these distances (Table 1). Along with this, 3 control sites (Table 1) were located deep in the forest (-2048 meters), in the surrounding pasture (+2048 meters) and along the forest edge. In total 233 trap locations were sampled, henceforth referred to as trap sites. Fragment area was transformed to log 10 prior to analysis (Table 1) (Ewers, et al., 2007).

Coleoptera specimens were morphologically identified to recognizable taxonomic units, totaling 893 different species for analysis and transformed to flux density of number of beetles per m<sup>2</sup> per day (Ewers et al., 2007). This data set is henceforth referred to as the Full data set.



**Fig 1;** map of 1300km<sup>2</sup> study area, the Hope River fragmentation project. Open circles indicate trap sites smaller than 3 hectares (from Appendix A, Ewers et al., 2007)

Fragment	Distance from edge (m)	Fragment Area (ha)	Fragment Area (log10)
Carlyle	+/- 1024	3485.84	6.03
Kakapo	+/- 1024	1060.87	3.54
Meat Safe	+/- 512	372.74	3.03
Front Dismal	+/- 256	70.68	2.57
Bush Hut	+/- 128	44.29	1.86
Boil	+/- 64	11.25	1.66
Prairie	+/- 64	2.88	1.09
Twin	+/- 32	0.92	0.59
Donut	+/- 16	0.69	0.28
Gully	+/- 8	0.08	0.23
High Ridge	+/- 8	0.06	0.03
Windy Point	+/- 8	0.04	0.02
Umbrella	+/- 4	0.02	0.01
Solo	+/- 4	0.01	0.00
St.James Control	+/- 1024	1060407.94	6.03
Home Range Control	+2048	N/a	Assigned 0.00 for comparison only)
Nina Control	-2048	1060407.94	Assigned 6.03 for comparison only

**Table 1;** Distance from edge in meters and extent of edge gradient and forest fragment area for the 14 sample sites and the 3 controls (Ewers et al., 2002). Fragment area is log10 transformed (Ewers et al., 2007) note the forest control site, Nina Control was assigned the log10 area value of 6.03 and Home Range Control was assigned the log10 area value of 6.03 for use in comparison and not used in analysis.



## 2.2 Environmental variables

Climatic, topological, soil property and landscape environmental variable data on the study region were obtained as raster layers of 25m<sup>2</sup> grid size for manipulation in ArcMap 9.2 GIS to be used as environmental layers for Maxent predictions (Table 2). Environmental variable values at each trap site were extracted from a GIS and all statistical analysis were conducted using R 2.7.2 software (R Development Core Team 2004).

<b>Variable</b>	<b>Abbreviation</b>	<b>Format</b>	<b>Unit</b>	<b>Source</b>
<b>Climatic</b>				
Annual rainfall deficit	deficit	Continuous	mm	Leathwick et al. 2003
June solar radiation	junes	Continuous	MJ/m <sup>2</sup> /day	Leathwick et al. 2003
Mean annual solar radiation	mas	Continuous	MJ/m <sup>2</sup> /day	Leathwick et al. 2003
Mean annual temperature	mat	Continuous	°C	Leathwick et al. 2003
Potential evapo-transpiration	r2pet	Continuous	ratio	Leathwick et al. 2003
Mean July Minimum temp	tmin	Continuous	°C	Leathwick et al. 2003
Mean October vapor pressure deficit	vpd	Continuous	kPa	Leathwick et al. 2003
<b>Soil properties</b>				
Acid soluble phosphorus	acidp	Categorical	Mg/100g	Leathwick et al. 2003
Age since last major rejuvenation	age	Categorical	years	Leathwick et al. 2003
Exchangeable calcium	calcium	Categorical	Mg/100g	Leathwick et al. 2003
Induration or hardness	induration	Categorical	descriptive	Leathwick et al. 2003
<b>Topographic</b>				
Altitude	altitude	Continuous	meters	Leathwick et al. 2003

<b>Variable</b>	<b>Abbreviation</b>	<b>Format</b>	<b>Unit</b>	<b>Source</b>
<b>Landscape Properties</b>				
Habitat classification	doccons	Categorical	descriptive	New Zealand Department of Conservation
Distance from forest edge	edgedist	Continuous	Log2-transformed distance in meters	Ewers et al., 2007
Size of fragmented area	fragarea	Continuous	Log10-transformed area in hectares	Ewers et al., 2007
Land Cover Data base	lcdb	Categorical	descriptive	New Zealand Ministry for the Environment
Distance from nearest river	riverdist	Continuous	Meters	New Zealand TopoMap
Distance from nearest road	roaddist	Continuous	Meters	New Zealand TopoMap

**Table 2;** Environmental variable layers obtained in raster format used and for Maxent predictions including the abbreviation, format, units and source (for correlation matrix, see Appendix A).

### *2.3 Modeling species distributions*

Maxent is a presence only modeling technique which estimates a species potential distribution by modeling the distribution of maximum entropy and has only recently been applied to ecology (Phillips et al., 2004). This technique has been shown to determine species niche requirements appropriately (Buermann et al., 2008), significantly outperforming other presence-only species distribution modelling methods (Elith et al., 2006) such as GARP (Phillips et al., 2004; Phillips et al., 2006). Maxent has been highlighted as one of the top performing techniques alongside more established methods relying on presence and absence data (Elith et al., 2006). Maxent also retains particular utility in providing more detailed predictions at a smaller scale than some other presence only methods such as GARP (Phillips et al. 2006). This is because even within overall high probability of occurrence areas, Maxent can further locate sites of higher suitability as well

as being able to utilize fine topographic data because the contribution of each variable to the model is added at each pixel (Phillips et al., 2006). Maxent also has an accessible user interface (Phillips et al. 2006) and easily interpreted output (Phillips et al., 2004; Phillips and Dudik 2008).

For analysis in Maxent 3.3.1, species from the Full dataset that were detected in less than 10 trap sites were excluded, retaining 245 of the 893 species for analysis. This subset of the Full data is henceforth referred to as the Subset data, the Subset data was converted to binary presence-absence data for use in Maxent (for Maxent format requirements see Phillips, 2009). Maxent models were fitted to all 245 species using the program default settings. The individual species model outputs were tested using a Jackknife. This consisted of comparing the individual species model performance in relation to 2 variants of: (1) a model using all variables excluding one variable in turn; and (2) a model using just one variable in turn and excluding all others (Phillips, 2009). This highlights which environmental variables contain the most useful information in isolation and which when excluded, reduce the models performance or “gain” (Phillips, 2009). Modeled distribution in comparison to a null hypothesis of random distribution is evaluated by Maxent by the Area Under the Curve (AUC), a value of 0.5 is considered as the same as random distribution with values over this value considered non-random, an absolute non random distribution would have a value of 1. Raster files representing the probability of occurrence for each individual species were transferred to ArcMap 9.2, where the probability value for each species occurring at each of the sampling sites were extracted. Modeled occurrence probabilities in their entirety are henceforth referred to as Maxent Whole. The Maxent Whole data was then converted to a format where species were considered to be absent at a site if their probability of occurrence was less than 0.5, this threshold data is referred to as Maxent Threshold and was used to investigate the utility of applying a threshold probability to the Maxent predictions.

#### *2.4 Data manipulation*

The goal of the analysis was to compare the community-level biodiversity patterns that were predicted for the Maxent (Maxent Whole and Maxent Threshold) occurrence probability values with the original data collected from the field (Full and Subset data). However, the spatial scale of the Hope River Forest Fragmentation project sampling design resulted in some grid squares in the GIS database having more than one trap site located within them. This meant that within-grid square replicates had exactly the same Maxent probability predictions, making them pseudoreplicates in any analysis. To avoid this problem, species

abundances from the Full and Subset data were converted from value per trap site to value per grid square. This was achieved by identifying multiple trap sites located in the same grid square and taking the average abundance of species from those traps (note- in Maxent Whole and Maxent Threshold matrices, multiple trap sites per grid square had the same values). Using a maximum of just one sample per grid square reduced the original 233 sampling sites to just 132 samples. Table 3 describes the data sets used for analysis.

<b>Data</b>	<b>Description</b>
<b>Full</b>	Surveyed data set of 893 species in an abundance format (flux density)
<b>Subset</b>	Subset of the Full data consisting of 245 species (species present in 10 or more trap sites) in an abundance format (flux density)
<b>Maxent Whole</b>	Data from the Maxent predictions of 245 species in an probability format
<b>Maxent Threshold</b>	Data from the Maxent predictions of 245 species in an probability format with a threshold of a 0.50 probability. Sites less than the threshold resulted in the species being considered as absent and assigned the value of 0

**Table 3;** the 4 data matrices consisting of the Full, Subset, Maxent Whole and Maxent Threshold data used for analysis

### *2.5 Conducting the Bray-Curtis index*

The Bray-Curtis distance (Beals, 1984; Bray and Curtis, 1957) is considered a robust measure for calculating community dissimilarity (Faith et al. 1987) and measures on a scale from 0 to 1, with a value of 1 indicating absolute dissimilarity and a value of 0 indicating complete similarity between communities. The Bray-Curtis distance was calculated between each of the 132 grid squares for the Full, Subset, Maxent Whole and Maxent Threshold datasets. The resulting 132 by 132 matrix was analyzed using a Mantel test (Mantel, 1967) with 1000 permutations based on the Pearson's product-moment correlation, as well as a paired Student's t-test. The Bray-Curtis matrices were compared between the Full, Subset data, Maxent Whole and Maxent Threshold predictions to see if they; (1) showed a similar pattern, and (2) had significantly different Bray-Curtis values.

## *2.6 Comparing the Bray-Curtis distance between sample sites and control sites*

As fragmentation has been shown to change community composition (Barbosa and Marquet, 2002; Davies et al., 2000; Didham et al., 1998a; Major et al., 2003), it would be expected that the observed data would reflect this and indicate a reasonably high Bray-Curtis dissimilarity. In order to ascertain if the Maxent predictions resulted in similar ecological patterns as the Full and Subset data, the Bray-Curtis indices between each grid square containing the “pristine” forest control (Nina Control) and grid squares containing the sample sites in the fragmented areas were calculated.

At each sampling site the average community dissimilarity relative to the 5 forest control sites were calculated in turn (Appendix A). These Bray-Curtis dissimilarity values were analyzed using Pearson's product-moment correlation and paired Student's t-test between the Full, Subset, Maxent Whole and Maxent Threshold predictions to see if they; (1) showed a similar pattern, and (2) had significantly different Bray-Curtis values.

## *2.7 Comparing species richness and diversity patterns*

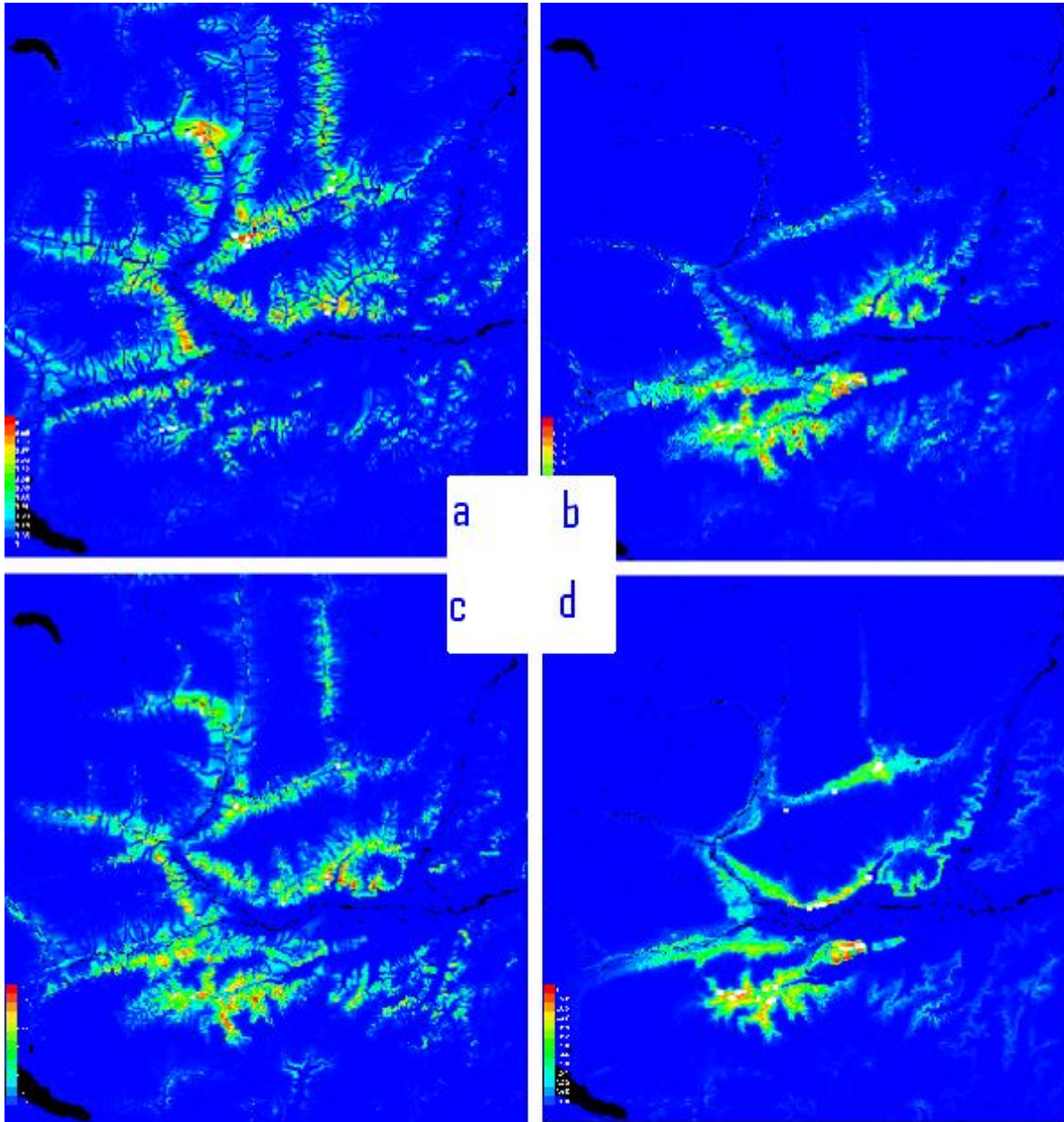
Species accumulation curves for the Subset, Maxent Whole and Maxent Threshold data sets of 245 species were generated to investigate differences in the landscape-scale estimate of species richness between observed and modeled data.

An aim of the analysis was to determine if the Maxent Whole and Maxent Threshold predictions retained similar spatial patterns of species richness and diversity in response to ecological gradients such as edge effects and fragment area to the observed survey data which have been identified in previous publications (Ewers et al., 2007; Ewers and Didham, 2008). To investigate this, the total number of different species per grid square and a Simpson's diversity index were calculated in response to distance from forest edge and fragment area. These values were then using the Pearson's product-moment correlation and paired Student's t-test.

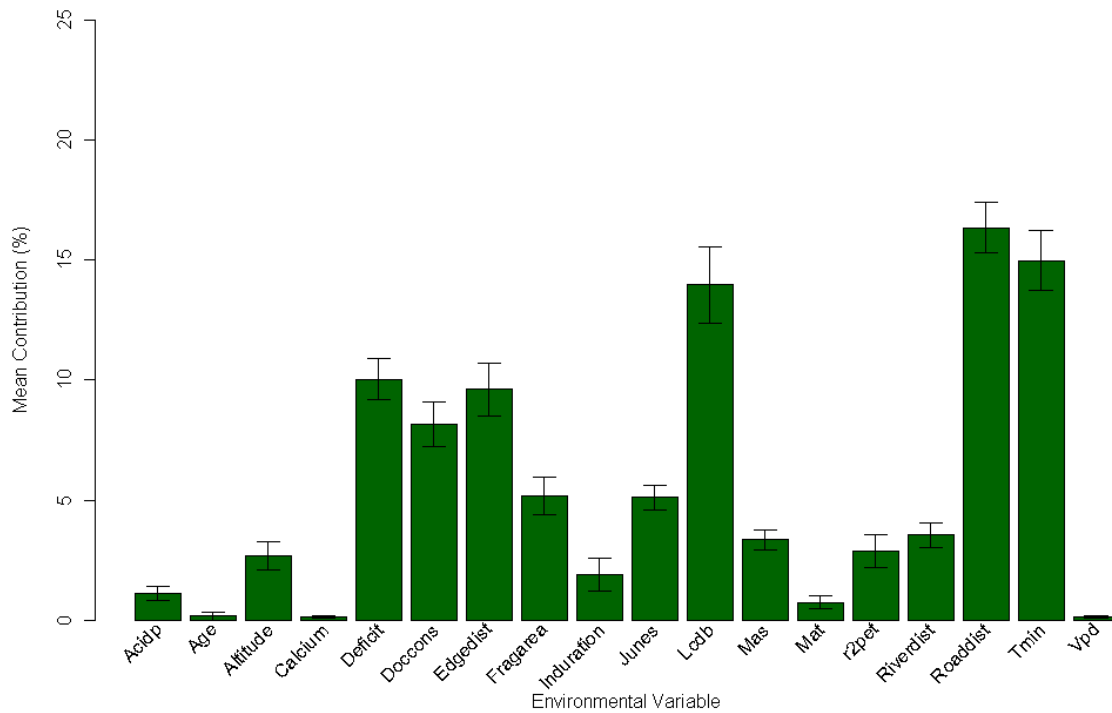
## 3.0 Results

### 3.1 *Species distribution models*

Maxent produced 245 species probability occurrence maps (examples Fig 2 and Appendix B) as well as variable contributions and jackknife output (Appendix B). The mean AUC score across all models was 0.982 +/- 0.014(3dp). The variable that contributed the most to all the models was distance to road with the highest average variable percentage contribution (16%) followed by minimum temperature and land cover type (15% and 14% respectively) (Fig 3). The variable that had the highest “gain” when used in isolation for the majority of the species models was minimum temperature (20% of models) followed by altitude then distance from edge (17% and 16% of models respectively) (Appendix C). The variable that decreased the “gain” when excluded in the majority of models was distance from road (60% of models) followed by land cover type and then annual rainfall deficit (12% and 11% of models respectively) (Appendix C).



**Fig 2;** Example of Maxent logistic output of species probability occurrence map (blue- lowest probability, red-highest occurrence probability) (see also Appendix B). Examples include (a) *Mecodema rugiceps* defined as a forest specialist, (b) *Mecodema fulgidum* defined as a matrix specialist, (c) *Costelytra brunneum* defined as a generalist and (d) *Costelytra zealandicum* defined as a matrix specialist. Habitat preferences defined by Ewers et al., 2007 Appendix *Ecological Archives* E088-007-A3.



**Fig 3;** Mean environmental variable percentage contribution (%), error bars show 95% confidence, Distance to road, minimum temperature and land cover type had the highest average percentage contribution to the 245 individual species models.

### 3.2 Bray-Curtis indices

The mean Bray-Curtis dissimilarity for the observed data was higher than the Maxent predictions (Table 4) and all Bray-Curtis matrices were significantly different from one another ( $p < 0.001$ ) (Table 5). Pearson's product-moment correlations were generally low with  $r$  values typically around 0.5 between the observed and the modelled data. Maxent Threshold data predicted higher mean Bray-Curtis dissimilarity than the Maxent Whole data but correlated marginally less than the Maxent Whole data with the observed data (Table 4).



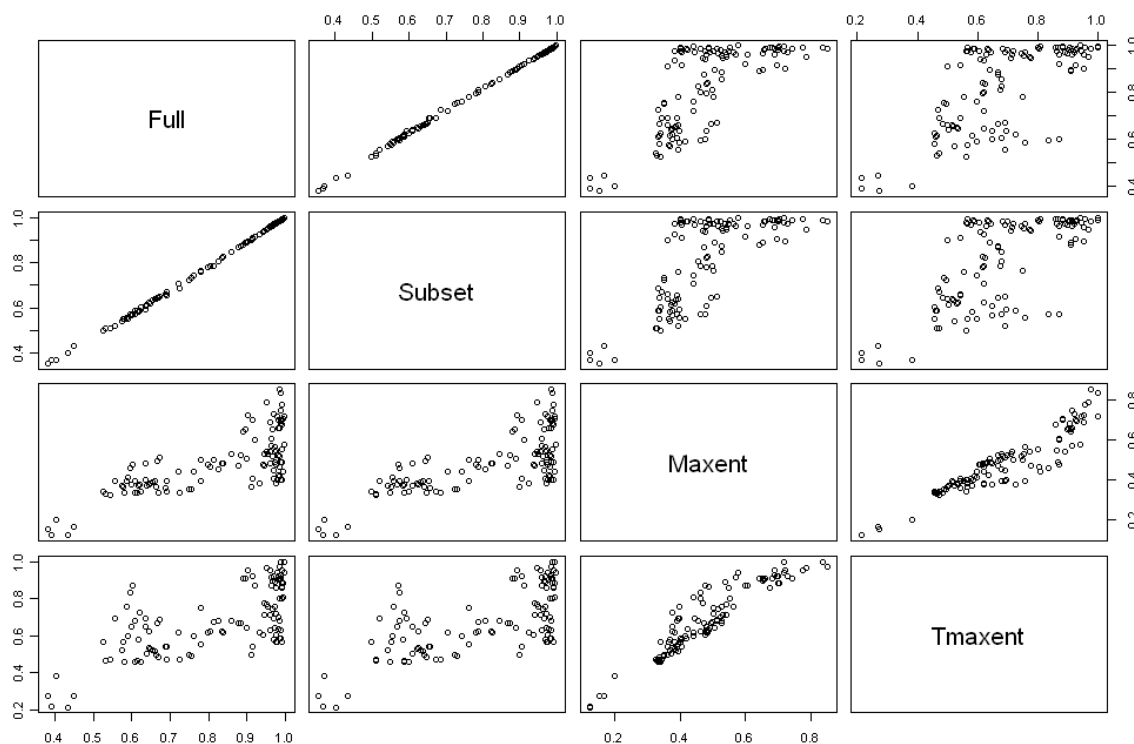
	<b>Bray-Curtis matrices Grid square dissimilarity</b>	<b>Forest control site-sample site dissimilarity</b>
<b>Full</b>	0.841 +/- 0.002	0.817 +/- 0.017
<b>Subset</b>	0.826 +/- 0.002	0.804 +/- 0.018
<b>Maxent Whole</b>	0.460 +/- 0.002	0.480 +/- 0.014
<b>Maxent Threshold</b>	0.671 +/- 0.002	0.681 +/- 0.017

**Table 4;** Mean and standard error Bray-Curtis dissimilarity values from the grid square to grid square comparison and the average forest control to sample site grid square comparison for the Full, Subset, Maxent Whole and Maxent Threshold. All to 3dp.

<b>Dataset</b>	<b>Subset</b>	<b>Maxent Whole</b>	<b>Maxent Threshold</b>
<b>Full</b>	<b>Correlation r value</b>	<b>0.999</b>	<b>0.504</b>
	<b>Mean of the differences</b>	<b>0.014</b>	<b>0.381</b>
<b>Subset</b>	<b>Correlation r value</b>	<b>N/A</b>	<b>0.502</b>
	<b>Mean of the differences</b>		<b>0.366</b>
<b>Maxent Whole</b>	<b>Correlation r value</b>	<b>r= 0.502</b>	<b>0.9231</b>
	<b>Mean of the differences</b>	<b>-0.366</b>	<b>-0.213</b>

**Table 5;** Mantel and paired Student's t-test results of Bray-Curtis dissimilarity matrices of grid square to grid square comparison between the Full, Subset, Maxent Whole and Maxent Threshold data sets. Bold values indicate show  $p < 0.05$ . Correlations from Pearson's product-moment correlation and mean of the differences from paired Student's t-test to 3dp. Results show all Bray-Curtis matrices were significantly different from each other ( $p < 0.05$ ) and the Maxent predictions were weakly correlated with the observed data ( $r = 0.5$ ).

The pattern of correlation between the Bray-Curtis indices was also reflected in the Pearson's product-moment correlation and the paired Student's t-test for Bray-Curtis dissimilarity between the forest control sites and the sample sites in the fragmented areas but with higher correlations ( $r=0.7$ ) between the Maxent predictions and the observed data (Table 4, 6 and Fig 4).



**Fig 4;** Pearson's Product-moment correlations between the Full, Subset, Maxent Whole (Maxent) and Maxent Threshold (Tmaxent) data for Bray-Curtis dissimilarity between the forest control and the sample sites. Results show Maxent predictions moderately correlating with the observed data in regards to dissimilarity between the forest control and the fragmented areas.

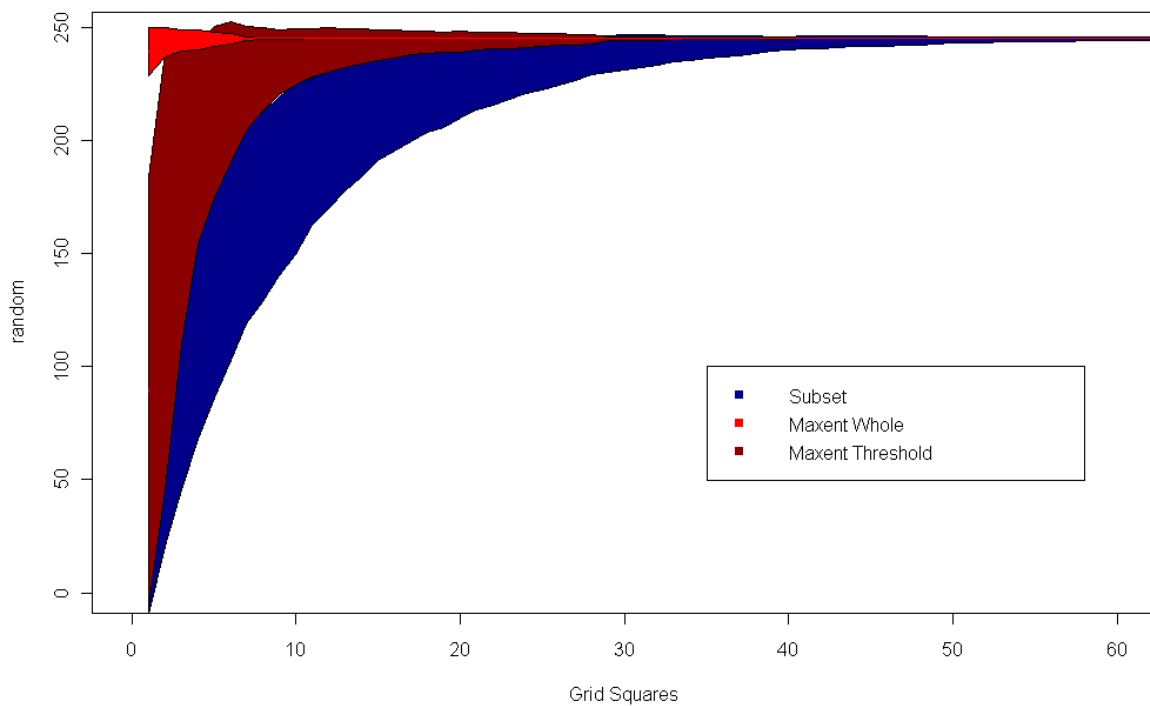
<b>Dataset</b>	<b>Subset</b>	<b>Maxent Whole</b>	<b>Maxent Threshold</b>
<b>Full</b>			
<b>Correlation r value</b>	<b>0.999</b>	<b>0.736</b>	<b>r= 0.687</b>
<b>Mean of the differences</b>	<b>0.014</b>	<b>0.337</b>	<b>0.136</b>
<b>Subset</b>	<b>N/A</b>		
<b>Correlation r value</b>		<b>0.735</b>	<b>0.684</b>
<b>Mean of the differences</b>		<b>0.324</b>	<b>mean 0.122</b>
<b>Maxent Whole</b>		<b>N/A</b>	
<b>Correlation r value</b>	<b>0.735</b>		<b>0.926</b>
<b>Mean of the differences</b>	<b>-0.324</b>		<b>-0.201</b>

**Table 6;** Pearson's product moment and paired Student's t-test results of Bray-Curtis dissimilarity between the forest control and sample sites in the fragmented area of the Full, Subset, Maxent Whole and Maxent Threshold data sets. Bold values indicate  $p < 0.05$ . Correlations from Pearson's product-moment correlation and mean of the differences from paired Student's t-test to 3dp. Results show all Bray-Curtis matrices were significantly different from each other ( $p < 0.05$ ) and the Maxent predictions were moderately correlated with the observed data ( $r = 0.7$ ).

### 3.3 Comparing richness and diversity patterns

#### 3.3.1 Species Accumulation

The species accumulation curves of the Subset, Maxent Whole and Maxent Threshold data showed the Maxent Whole exhibited accelerated species accumulation with an asymptote within just a few grid squares whereas the Maxent Threshold predictions asymptote later than the Maxent Whole predictions but earlier than the Subset data (Fig 5).

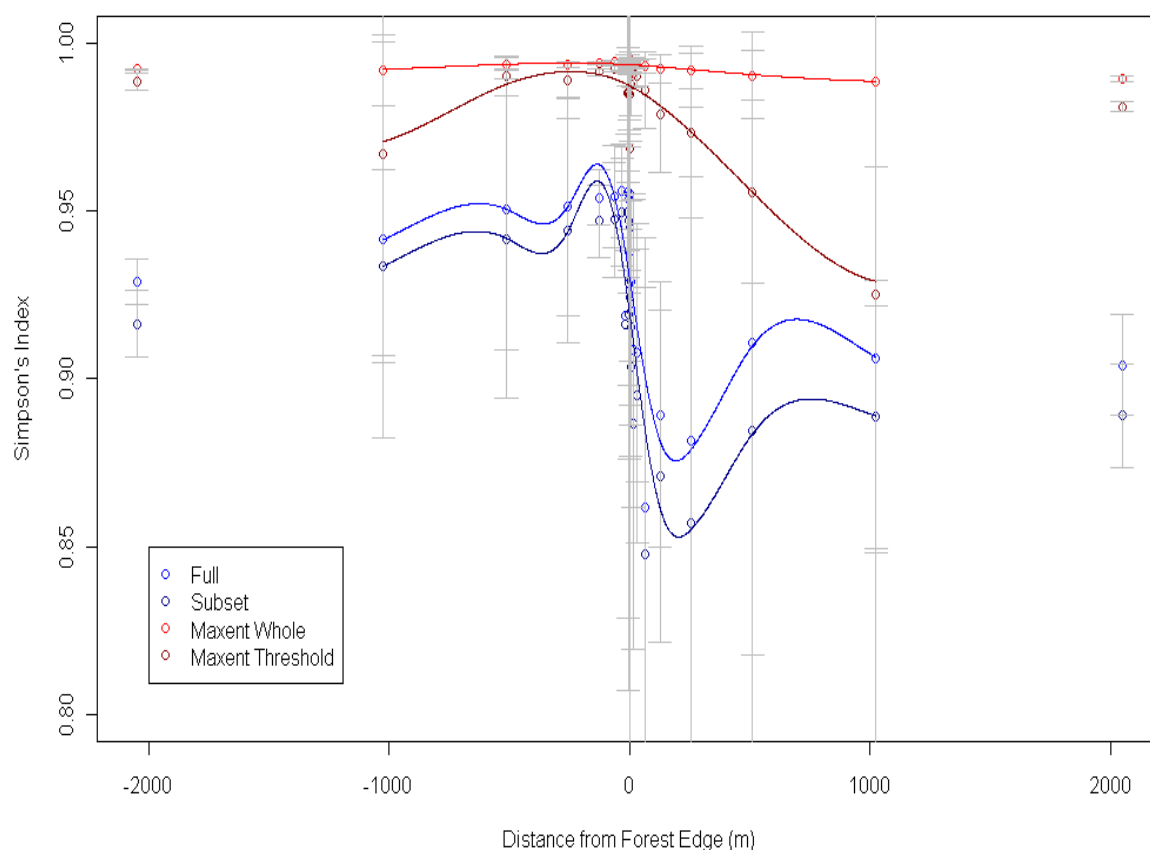


**Fig 5;** species accumulation curved for Subset, Maxent Whole and Maxent Threshold data. Based on 100 permutations and sites added in random order. Polygons shows 95% confidence limits.

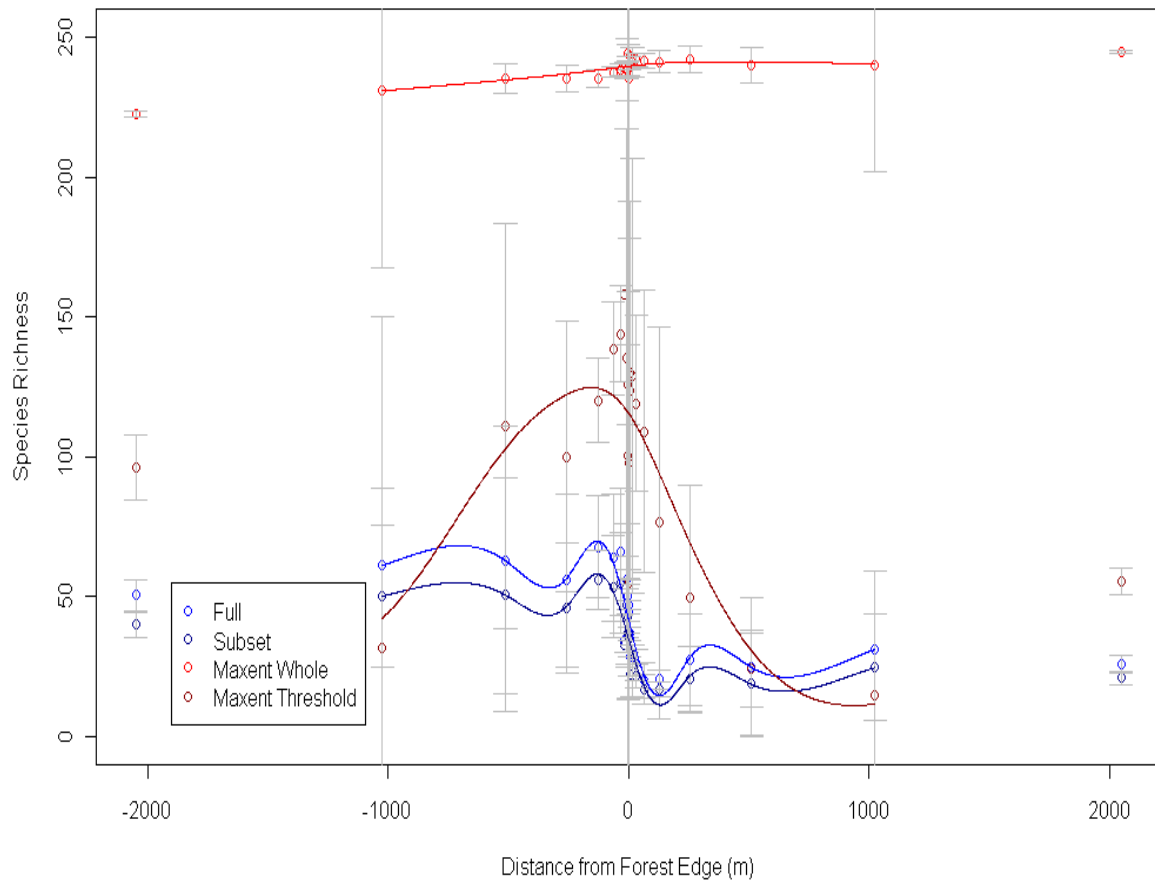
### 3.3.2 Distance from edge

Simpson's index in response to distance from edge for all data sets showed a significant ( $p < 0.01$ ) weak negative correlation (Table 7). This was slightly stronger for the Maxent Threshold data ( $r = -0.4$ ) (Fig 6).

Absolute richness as measured by number of species in each grid square showed a significant ( $p < 0.05$ ) weak negative correlation in all but the Maxent Whole data which predicted the opposite pattern in response to distance from forest edge ( $r = 0.4$ ) (Fig 7, Table 7).



**Fig 6;** Simpson's index in response to distance from edge (m). Controls at -2048, 2048 (m) for comparison. Trend line shows General Additive Model (GAM) to represent detailed pattern of the relationship between species diversity and distance from edge. Maxent models predict higher diversity than recorded at the study site.



**Fig 7;** Absolute species richness in response to distance from edge (m). Controls at -2048, 2048 (m) for comparison. Trend line shows General Additive Model (GAM) to represent detailed pattern of the relationship between species richness and distance from edge. The Maxent Whole model overall predicts higher species richness than recorded at the study site and shows a dissimilar pattern in response to edge

For the Full, Subset, Maxent Whole and Maxent Threshold data, all diversity and richness values and means were significantly different from one another ( $p < 0.001$ ) (Table 7 and 8). Maxent predicted consistently higher mean diversity and richness than what was calculated for the observed data and correlated weakly with the observed data (Table 8).

	<b>Simpson's index</b>	<b>Richness</b>
<b>Full</b>		
<b>Mean +/- SD</b>	0.926 +/- 0.005	43.312 +/- 2.400
<b>Correlation r value</b>	<b>-0.278</b>	<b>-0.390</b>
<b>Subset</b>		
<b>Mean +/- SD</b>	0.914 +/- 0.005	35.638 +/- 1.943
<b>Correlation r value</b>	<b>-0.315</b>	<b>-0.408</b>
<b>Maxent Whole</b>		
<b>Mean +/- SD</b>	0.993 +/- <0.001	239.292 +/- 0.400
<b>Correlation r value</b>	<b>-0.308</b>	<b>0.433</b>
<b>Maxent Threshold</b>		
<b>Mean +/- SD</b>	0.985 +/- 0.002	106.856 +/- 5.432
<b>Correlation r value</b>	<b>-0.452</b>	<b>-0.219</b>

**Table 7;** Mean and standard deviation (3dp) Simpson's index value and absolute species richness calculated in grid squares for the Full, Subset, Maxent Whole and Maxent Threshold datasets. Pearson's product-moment correlation r value (3dp) for Simpson's index in response to distance from edge and absolute species richness in response to distance from edge. Bold values indicate  $p < 0.05$ . Note- Maxent Whole predictions for richness in response to distance from edge show opposite trends.

	Subset	Maxent Whole	Maxent Threshold
<b>Full</b>			
<b>Simpson's</b>			
Correlation r value	<b>0.987</b>	0.127	0.138
Mean of the differences	<b>0.012</b>	<b>-0.067</b>	<b>-0.058</b>
<b>Richness</b>			
Correlation r value	<b>0.991</b>	<b>-0.327</b>	<b>0.306</b>
Mean of the differences	<b>7.674</b>	<b>-195.980</b>	<b>-63.544</b>
<b>Subset</b>			
<b>Simpson's</b>			
Correlation r value	N/A	0.125	0.167
Mean of the differences		<b>-0.079</b>	<b>-0.070</b>
<b>Richness</b>			
Correlation r value	N/A	<b>-0.356</b>	<b>0.330</b>
Mean of the differences		<b>-203.654</b>	<b>-71.219</b>
<b>Maxent Whole</b>			
<b>Simpson's</b>			
Correlation r value	0.125	N/A	<b>0.858</b>
Mean of the differences	<b>0.079</b>		<b>0.009</b>
<b>Richness</b>			
Correlation r value	<b>-0.356</b>	N/A	-0.045
Mean of the differences	<b>203.654</b>		<b>132.436</b>

**Table 8;** Pearson's product moment and paired Student's t-test results of Simpson's index and richness values between the Full, Subset, Maxent Whole and Maxent Threshold data sets. Bold values indicate  $p < 0.05$ . r values from Pearson's product-moment correlation and mean of the differences from paired Student's t-test to 3dp. Maxent predictions weakly correlate with the observed data.



### 3.3.3 Forest fragment area

All data sets did not show a significant correlation between forest fragment area and species diversity or richness apart from Maxent Whole data which showed a significant ( $p < 0.001$ ) negative correlation of species richness (log) ( $r = -0.7$ ), where increasing fragment area subsequently decreased species richness, a pattern opposite to the other data sets (Table 9).

	<b>Simpson's index</b>	<b>Richness</b>
<b>Full</b>		
<b>Mean +/- SD</b>	0.949 +/- 0.003	1.731 +/- 0.024
<b>Correlation r value</b>	0.156	0.227
<b>Subset</b>		
<b>Mean +/- SD</b>	0.941 +/- 0.004	1.652 +/- 0.023
<b>Correlation r value</b>	0.194	0.255
<b>Maxent Whole</b>		
<b>Mean +/- SD</b>	0.994 +/- <0.001	2.376 +/- <0.001
<b>Correlation r value</b>	0.190	<b>-0.687</b>
<b>Maxent Threshold</b>		
<b>Mean +/- SD</b>	0.989 +/- 0.001	2.033 +/- 0.033
<b>Correlation r value</b>	0.229	0.190

**Table 9;** Mean and standard deviation (3dp) Simpson's index value and species richness (log) calculated for grid squares in forest fragments for the Full, Subset, Maxent Whole and Maxent datasets. Pearson's product-moment correlation r value (3dp) between Simpson's index in response to forest fragment area and species richness in response to forest fragment area. Bold values indicate  $p < 0.05$ .

For the Full, Subset, Maxent Whole and Maxent Threshold data, all diversity and richness values for forest sites were significantly different from one another ( $p < 0.001$ ). Maxent predicted consistently higher mean diversity and richness than the observed data and correlated weakly with the observed data (Table 10).

	<b>Subset</b>	<b>Maxent Whole</b>	<b>Maxent Threshold</b>
<b>Full</b>			
<b>Simpson's</b>			
<b>Correlation r value</b>	<b>0.990</b>	<b>0.405</b>	<b>0.371</b>
<b>Mean of the differences</b>	<b>0.008</b>	<b>-0.045</b>	<b>-0.040</b>
<b>Richness (log)</b>			
<b>Correlation r value</b>	<b>0.986</b>	-0.102	<b>0.391</b>
<b>Mean of the differences</b>	<b>0.08</b>	<b>-0.645</b>	<b>-0.302</b>
<b>Subset</b>			
<b>Simpson's</b>	<b>N/A</b>		
<b>Correlation r value</b>		<b>0.410</b>	<b>0.394</b>
<b>Mean of the differences</b>		<b>-0.053</b>	<b>-0.047</b>
<b>Richness log</b>	<b>N/A</b>		
<b>Correlation r value</b>		-0.137	<b>0.428</b>
<b>Mean of the differences</b>		<b>-0.72</b>	<b>-0.381</b>
<b>Maxent Whole</b>			
<b>Simpson's</b>		<b>N/A</b>	
<b>Correlation r value</b>	<b>0.410</b>		<b>0.910</b>
<b>Mean of the differences</b>	<b>0.053</b>		<b>0.005</b>
<b>Richness log</b>		<b>N/A</b>	
<b>Correlation r value</b>	-0.137		0.125
<b>Mean of the differences</b>	<b>0.724</b>		<b>0.343</b>

**Table 10;** Pearson's product moment and paired Student's t-test results (3dp) of Simpson's index and richness between the Full, Subset, Maxent Whole and Maxent Threshold data sets. Bold values indicate  $p < 0.05$ . r values from Pearson's product-moment correlation and mean of the differences from paired Student's t-test to 3dp. Maxent predictions weakly correlate with the observed data.

## 4.0 Discussion

### 4.1 *Species distribution models*

Distribution predictions for the 245 species produced by Maxent had a high mean AUC (0.982 +/-0.014). These results are in accordance with the literature, where Maxent has been shown to consistently produce models with high accuracy and been highlighted as a top performer even with small sample sizes (Elith et al., 2006; Elith and Graham, 2009, Hernandez et al., 2006; Wisz et al., 2008).

Since the publication of Maxent as an ecological tool for predicting species distributions (Phillips et al., 2004), the high suitability of this technique has been demonstrated for a range of applications and taxa. Maxent has predominantly been applied to the field of biological conservation by predicting species distributions in order to establish conservation priorities and allocate actions for a range of species (Boubii and Lima, 2009; DeMatteo and Loiselle, 2008; Hinojosa-Diaz et al., 2009; Pawar et al., 2007; Suarez-Seoane et al., 2008; Yost et al., 2008). Recently in its most high-profile use, Maxent has been utilized to allocate species richness priority locations for conservation purposes in Madagascar (Kremen et al., 2008). As well as conservation priority setting, Maxent has particular applicability in detecting new locations of rare species (Boubii and Lima, 2009; Williams et al., 2009) and has been applied to climate change projections in order to assign future management decisions and policies (Fuller et al., 2008). Maxent has also been used to model the potential distribution of invasive alien species (Crossman et al., 2008; Kadoya et al., 2009; Mingyang et al., 2008; Thuiller et al., 2005) and pathogens and parasites (Puschendorf et al., 2009; Rose et al., 2009; Wang et al., 2007). In conclusion, the support for Maxents accuracy and utility across a range of taxa and applications is highlighted in the literature and in the present study. In extension to this, the present study expands upon former work as it has shown Maxents utility to model species distributions with accuracy at finer spatial scales than previously tested.

#### *4.2 Patterns of diversity and community dissimilarity*

Maxent Whole and Maxent Threshold results showed similar trends in Simpson's diversity in response to distance from edge ( $r = -0.34$ ,  $r = -0.20$  respectively) to the real survey dataset. This reflects the strong edge effects on the Coleoptera species richness that has been illustrated previously with this dataset (Ewers et al., 2007; Ewers and Didham, 2008). However, in contrast, Maxent Whole and Maxent Threshold produced significantly different richness and diversity results than the surveyed data and alarmingly, Maxent Whole predicted opposite trends for richness in response to distance from edge and forest fragment area ( $r = 0.81$ ,  $r = -0.69$  respectively). In addition to this, model predictions did not adequately correlate with real community dissimilarity between grid squares according to the Bray-Curtis matrices for both Maxent Whole and Maxent Threshold data ( $r = 0.50$ ), with lower and significantly different dissimilarity values than the observed survey data. These results indicate that overall Maxent predictions did not accurately reflect the patterns of alpha and beta diversity that were exhibited at the study area. Despite this, the inclusion of thresholds however, did increase Maxent's ability to reflect real surveyed patterns of richness and diversity. Additionally, although Maxent Threshold predictions did not display a strong correlation with the observed data, the pattern of species richness to distance from edge was similar to that reported previously for a subset (just those species which were affected by distance from edge) of the dataset (Ewers and Didham, 2008). A pattern also reported in other Coleoptera diversity studies in response to forest fragmentation (Magura, 2002).

With regards to compositional dissimilarity between the pristine forest control sites and the fragmented sample sites, the distribution model predictions were moderately congruent with the real survey data for both Maxent Whole and Maxent Threshold data (correlation  $r = 0.73$ ,  $r = 0.68$  respectively). These changes in community composition attest the results previously reported for the dataset (Ewers et al., 2007; Ewers and Didham, 2008). Additionally, these results reflect how habitat fragmentation changes insect community assemblages inducing marked differences between forest interior, edge and matrix species compositions as well as community dissimilarity along a gradient of fragment sizes (Barbosa and Marquet, 2002; Davies et al., 2000; Didham et al., 1998a; Major et al., 2003). Changes in community assemblages are the result of a combination and interaction of contracting fragment area subsequently decreasing insect biodiversity (Didham et al., 1998a; Larsen et al., 2005), increased edge habitat adjusting trophic structures (Didham et al., 1998b; Valladares et al., 2006) and the decline of fragmentation sensitive species (Davies et al., 2000; Didham et al., 1998b; Larsen et al., 2005). Maxent predictions showed a similar pattern to the surveyed data concerning community dissimilarity between sites indicating that the models were able

to predict patterns of dissimilarity in an environmentally heterogeneous landscape. Despite this, mean dissimilarity values were overall lower than and significantly different from the real survey data.

#### *4.3 Possible explanations for congruence and potential mitigation methods*

A suggestion for the incongruence between the model predictions and the real data set could be due to Maxent's tendency to over-predict the sites that may be suitable for the species modeled, a factor evident in the species accumulation curves. In concordance to these results, Maxent predictions have been reported to contain errors of commission (Boubii and Lima, 2009; Williams et al., 2009) and to be the least discriminatory when assigning high probability occurrence locations in comparison to other modeling techniques such as GLMs (Williams et al., 2009). This tendency to produce "Type I errors of false positives" may be a particular problem when predicted species distributions are to be applied to conservation actions if species do not occur in areas that are selected for protection according to the model (Loiselle et al., 2003).

One of the underlying mechanisms for the Maxent models' tendency to over-predict the extent of species distributions may be due to biotic interactions such as competition being absent from the predictor variables. Competition is an important factor in determining the distribution of species which is usually absent from species distribution models (Guisan et al., 2006; Guisan and Thuiller, 2005). The inclusion of competition in species distribution models has been highlighted as a general limitation of current models and a basis for future development (Guisan and Thuiller, 2005; Guisan and Zimmermann, 2000; Pulliam, 2000). Competition has been shown to influence the distribution of species (Costa et al., 2008; Ritchie et al., 2009) and the inclusion of competition as a model parameter has been shown to improve model performance (Anderson et al., 2002; Leathwick and Austin, 2001; Pondchong et al., 2009; Ritchie et al., 2009). Techniques for the inclusion of competition include either adding the distribution of a known competitor as a predictor variable (Anderson et al., 2002; Leathwick and Austin, 2001), partitioning environmental and biotic variables to obtain a predictive map by averaging the models (Pondchong et al., 2009) or using each species model as a predictor in an iterative process until each model has obtained the highest accuracy (Guisan and Zimmermann, 2000).

Over a landscape there has been evidence of interspecific competition at a small scale but not at a large scale (Zhang et al., 2009), suggesting that this process may be more applicable to small scale distributional modelling. However, the effects of competition at a small scale may be more potent for sessile organisms (Hart and Marshall, 2009) in addition to this, at particularly small scales ( $m^2$ ) the negative effects of competition may be overestimated (Jakobsson et al., 2009). This means that the inclusion of competition must be critically evaluated and applied appropriately to the scale to which the models are to be projected on and in consideration of the dispersal limitations of the species involved.

Due to these insights, the inclusion of each species layer as a predictor variable or the occurrence of a known predator or competitor to the modeled species is recommended to enhance the predictive ability of species distribution models and in particular of Maxent's model accuracy. A major constraint of this proposal is that this technique may only be feasible with species that are well documented in terms of taxonomy and ecology (Ewers, 2009). This limitation may result in the benefits of community modeling being subsequently impaired as species that lack biotic interaction information may be excluded or inaccurately modeled. This is a highly probable consequence when confronted with large community datasets especially as the inclusion of all biotic interactions such as apparent competition, mutualism, and pollination for example, could be an immeasurable task. In conclusion, the emergent characteristics of a community may never be absolutely reconstructed when using the aggregated properties of individual species models to construct community composition. This may be an ongoing limitation of this technique and may be an obstacle which is more adequately approached by other community modeling techniques such as "assemble first and predict later" or "assemble and predict together" (Ferrier and Guisan, 2006)

#### *4.4 The utility of occurrence probability thresholds*

The utility of adding restrictions to Maxent suitability predictions by using a threshold value of occurrence probability was investigated through the Maxent Threshold data. Results showed that even though species accumulation more closely represented that of the Subset data, Maxent Threshold predictions still accelerated species accumulation in comparison to the observed data. Thresholds also did not improve Bray-Curtis distance dissimilarity concordance between the Maxent predictions and the observed data. The use of Thresholds did however, moderately increase the models ability to reflect ecological patterns of species richness and diversity patterns in response to habitat edge and area effects than the Maxent Whole predictions. In conclusion, thresholds for determining a species presence or absence

from Maxent predictions from a community can be used to partly mitigate the effects of over-prediction on species accumulation but do not necessarily enhance concordance with community dissimilarity. This technique retains particular utility if the agenda of the model is to ascertain alpha diversity response patterns to environmental gradients such as distance from fragment edge as illustrated in the present study. In conclusion the use and evaluation of threshold at different probability values is advocated.

#### *4.5 Appropriate selection of environmental variables*

The choice of predictor variables is a vital consideration in order to obtain accurate and biologically meaningful distribution predictions and depends on the intended purpose of the models. If the aim is to model species distributions at a fine spatial scale as in the current investigation, the inclusion of more small scale environmental variables may need to be included. For example, Carabid community assemblages have been shown to be affected by shrub cover and soil organic matter, canopy structure, vegetation diversity, leaf litter, shrub and herb cover and micro-habitat type (Barton et al., 2009; Jukes et al., 2001; Magura, 2002; Taboada et al., 2006). Canopy height has also been identified as the most significant variable in determining species composition at the study site (Ewers et al., 2007). These fine scale environmental factors may exhibit more heterogeneous patterns over the landscape, possibly aiding more fine scaled predictions and therefore increasing the accuracy of the models. This means that predictors such as these examples may need to be incorporated into small scale species distribution models in order to improve prediction accuracy.

The environmental variables used in the present study had mixed contributions to each of the individual species models (Fig 3). The main unexpected result was that distance to the road had the highest average percentage variable contribution despite lack of correlation between other environmental variables or evidence of road-side bias even within species models where distance to the road reduced the “gain” when isolated (Appendix C). Visual inspection of the logistic output from species with distance from road as the highest variable contribution suggests that the area around the road is not suitable for most species as it predicts low occurrence probability in this region even when distance from road was excluded from the model (Appendix C). The roads are situated in the valleys of the study area so a general correlation with altitude should be expected (Ewers, 2009). This could have accounted for the unexpected result especially as altitude was one of the highest rated variables when used in isolated along with minimum temperature, both of which are highly

correlated. However, distance from the road was only weakly correlated with altitude ( $r=-0.2$ ) and minimum temperature ( $r=0.3$ ) even in individual species models where distance from the road reduced the “gain” when isolated from the models (Appendix C). Alternatively, or in combination with this, the unanticipated result could be an artifact of the sampling design (no sites were located closer than 480m to a road) or an undetected environmental or habitat factor and therefore, more quantitative investigation is recommended.

## **5.0 Limitations and further applications**

### *5.1 The utility of Maxent and presence-only data*

The potential utility is high for Maxent and presence-only data to model species distributions at fine spatial scales as in the present study, and this technique could be applied to reserve or protected area allocation where meaningful and relevant actions could be implemented which require more defined scale habitat suitability identification. Maxent's assessable user interface and the logistic output could enable those without prior modeling experience to utilize the software for as examples; reserve managers, or those in planning departments. With the accelerating use of Maxent in modeling species distributions and growing evidence for its utility and accuracy (Elith et al., 2006; Elith and Graham, 2009; Wisz et al., 2008), the indication is evident that ecological modeling may be on the brink of a presence-only paradigm shift. This is most prevalent for vagile species where obtaining true absence data is not always achievable.

### *5.2 Modeling communities*

Modeling the collective properties of biodiversity such as community composition and similarity as opposed to individual species or alpha diversity is highly advocated (Ewers et al., 2009; Ferrier et al., 2002; Guisan et al., 2006; Su et al., 2004). However, the allocation of conservation priorities to biologically diverse areas may be obstructed if modeled community structure is incongruent with real community assemblages such as in the present study. This means that conclusions drawn from both the spatial patterns of species and communities need to be considered with caution and critical evaluation in order to assign appropriate, meaningful and effective policy decisions, priority setting and conservation actions (Austin, 2007; Ferrier et al., 2002; Guisan et al., 2006). Maxent overall had a tendency to over-



predict alpha diversity and under-estimate beta diversity, relative to observed field data and this could be a particular concern if inferences made from these conclusions subsequently misdirect conservation efforts.

### *5.3 Concluding remarks*

Despite the limitations that have been acknowledged and discussed, Maxent and species distribution modeling retain high utility. This is particularly pronounced where species distributions can be projected to future climate scenarios in order to mitigate potential negative effects, however this is only relevant if biotic interactions are accounted for (Araujo and Luoto, 2007; Preston et al., 2008). The modeling of individual species prior to aggregation into community assemblages is especially applicable to predicting the impacts of climate change on biodiversity as this technique allows the idiosyncratic responses of individual species to be accounted for (Ferrier and Guisan, 2005). The consideration of biotic interactions and the inclusion of appropriate predictor variables are highly important in order to produce robust, applicable and biologically meaningful models. With more explicit and accurate predictions, species distributional modeling techniques such as Maxent are powerful tools that can be utilized to anticipate, locate and mitigate the potential negative effects of climate change and ecosystem fragmentation on vulnerable species, populations and communities and the indispensable ecosystem functions and services they provide.

## **6.0 Acknowledgments**

With thanks to Dr R. M Ewers for data, guidance, advice and discussion.

Thanks to Dr T. Collins for advice on statistics and to my friends and family for ongoing support.

## 7.0 References

- Anderson, R. P., Peterson, A. T., Gomez-Laverde, M., 2002. Using niche-based GIS modeling to test geographic predictions of competitive exclusion and competitive release in South American pocket mice. *OIKOS*. 98, 3-16.
- Andresen, E., 2003. Effects of forest fragmentation on dung beetle communities and functional consequences for plant regeneration. *Ecography*. 26, 87-97.
- Araujo, M. B and Luoto, M., 2007. The importance of biotic interactions for modelling species distributions under climate change. *Global Ecology and Biogeography*. 16, 743-753.
- Austin, M., 2007. Species distribution modes and ecological theory: A critical assessment and some possible new approaches. *Ecological Modelling*. 200, 1-19.
- Balmford, A., Bennun, L., ten Brink, B., Cooper, D., Cote, I. M., Crane, P., Dobson, A., Dudley, N., Dutton, I., Green, R., Gregory, R. D., Harrison, J., Kennedy, E., Kremen, C., Leader-Williams, N., Lovejoy, T. E., Mace, G., May, R., Mayaux, P., Morling, P., Phillips, J., Redford, K., Ricketts, T. H., Rodriguez, J. P., Sanjayan, M., Schei, P. J., van Jaarsveld, A. S., Walther, B. R., 2005. The Convention on Biological Diversity's 2010 Target. *Science*. 307, 212-213.
- Balvanera, P., Pfisterer, A. B., Buchmann, N., He, J., Nakashizuka, T., Raffaelli, D., Bernhard, S., 2006. Quantifying the evidence for biodiversity effects on ecosystem functioning and services. *Ecology Letters*. 9, 1146-1156.
- Barbosa, O and Marquet, P.A., 2002. Effects of forest fragmentation on the beetle assemblage at the relict forest of Fray Jorge, Chile. *Oecologia*. 132, 296-306.
- Barton, P. S., Manning, A. D., Gibb, H., Lindenmayer, D. B., Cunningham, S. A. 2009. Conserving ground-dwelling beetles in an endangered woodland community: Multi-scale habitat effects on assemblage diversity. *Biological Conservation*. 142, 1701-1709.
- Beale, C. M., Lennon, J. J., Gimona, A., 2008. Opening the climate envelope reveals no macroscale association with climate in European birds. *PNAS*. 105(39), 14908-14912.
- Beals, E. W., 1984. Bray-Curtis ordination: an effective strategy for analysis of multivariate ecological data. *Advances in Ecological research*. 14, 1-55.

Bojorquez-Tapia, L. A., Azuara, I., Ezcurra, E., 1995. Identifying conservation priorities in Mexico through Geographic Information Systems and modeling. *Ecological Applications*. 5(1), 215-231.

Boubii, J. P and de Lima, M. G., 2009. Modeling the Geographic Distribution and Fundamental Niches of *Cacajao* spp. and *Chiropotes israelita* in Northwestern Amazonia via a Maximum Entropy Algorithm. *International Journal Primatology*. 30, 217-228.

Bray, J. R and Curtis, J. T., 1957. An ordination of the upland forest communities of southern Wisconsin. *Ecological Monographs*. 27, 325-329.

Buermann, W., Saatchi, S., Smith, T. B., Zutta, B. R., Chaves, J. A., Mila, B., Graham, C. H., 2008. Predicting species distributions across the Amazonian and Andean regions using remote sensing data. *Journal of Biogeography*. 35, 1160-1176.

Costa, G. C., Wolfe, C., Shepard, D. B., Caldwell, J. P., Vitt, L. J., 2008. Detecting the influence of climatic variables on species distributions: a test using GIS niche-based models along steep longitudinal environment gradient. *Journal of Biogeography*. 35, 637-646.

Crossman, N. D., Bran, B. A., Cooke, D. A., 2008. An invasive plant and climate change threat index for weed risk management: Integrated habitat distribution pattern and dispersal processes. *Ecological Indicators*. In Press.

Davies, K.F., Margules, C.R., Lawrence, J. F., 2000. Which traits of species predict population declines in experimental forest fragments. *Ecology*. 81(5), 1450-1461.

DeMatteo, K. E and Loiselle, B. A., 2008. New data on the status and distribution of the bush dog (*Speothos venaticus*): Evaluating its quality of protection and directing research efforts. *Biological Conservation* 141, 2494-2505.

Didham, R.K., Hammond, P.M., Lawton, J.H., Eggleton, P., Stork, N.E., 1998. Beetle species responses to tropical forest fragmentation. *Ecological Monographs*. 68(3), 295-323.

Didham, R. K., Lawton, J. H., Hammond, P. M., Eggleton, P., 1998. Trophic structure stability and extinction dynamics and beetles (Coleoptera) in tropical forest fragments. *Philosophical Transactions of the Royal Society of London Series B*. 353, 437-451.

Elith, J., Graham, C. H., Anderson, R. P., Dudik, M., Ferrier, S., Hijmans, R. J., Huettmann, F., Leathwick, J. R., Lehmann, A., Li, J., Lohmann, L. G., Loiselle, B. A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J. M., Peterson, T., Phillips, S. J., Richardson, K., Scachettie-Pereira, R., Schapire, R. E., Soberson, J., Williams, S., Wisz, M. S.,

- Zimmermann, N. E., 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography*. 29(2), 129-151.
- Elith, J and Graham, C. H., 2009. Do they? How do they? WHY do they differ? On finding reasons for differing performances of species distribution models. *Ecography*.32, 66-77.
- Ewers, R. M., Didham, R. K., Ranson, L. H., 2002. The Hope River forest fragmentation project. *The Weta*. 24(1), 25-34.
- Ewers, R. M., Thorpe, S., and Didham, R. K., 2007. Synergistic interactions between edge and area effects in a heavily fragmented landscape. *Ecology*.88(1), 96-106.
- Ewers, R. M., 2008. Spatio-temporal variation in mortality rates of *Mecodema* spp. (Coleoptera: Carabidae) across a forest-grassland edge in New Zealand. *Insect Conservation and Diversity*. 1, 40-47.
- Ewers, R. M. and Didham, R.K., 2008. Pervasive impact of large-scale edge effects on a beetle community. *PNAS*.105(14), 5426-5429.
- Ewers, R.M., Kapos, V., Coomes, D.A., Laforteza, R., Didham, R.K., 2009. In Press. Mapping community change in modified landscapes. *Biological Conservation*.
- Ewers, R. M. 2009. Personal Communication.
- Faith, D. P., Minchin, P. R., Belbin, L., 1987 Compositional dissimilarity as a robust measure of ecological distance. *Vegetatio*. 69, 57-68.
- Ferraz, G., Russell, G.J., Stouffer, P.C., Bierregaard, R.O., Pimm, S.L. Lovejoy, T.E., 2003. Rates of species loss from Amazonian forest fragments.
- Ferrier, S., 2002. Mapping Spatial Patterns in Biodiversity for Regional Conservation Planning: Where to from Here? *Systematic Biology*. 51(2), 331-363.
- Ferrier, S., Drielsma, M., Manion, G., Watson, G., 2002. Extended statistical approaches to modelling spatial pattern in northeast New South Wales. II Community-level modelling. *Biodiversity and Conservation*. 11, 2309-2338.
- Ferrier, S., and Guisan, A., 2006. Spatial modeling of biodiversity at the community level. *Journal of Applied Ecology*. 43, 393-404.

- Fuller, T., Morton, D., Sarkar, S., 2008. Incorporating uncertainty about species' potential distribution under climate change into the selection of conservation area with a case study from the Arctic Coastal Plain of Alaska. *Biological Conservation*. 141, 1547-1559.
- Funk, V. A. and Richardson, K. S., 2002. Systematic Data in Biodiversity Studies: Use It or Lose It. *Systematic Biology*. 51 (2), 303-316.
- Graham, C. H., Ferrier, S., Huettmann, F., Moritz, C. Pereson, A. T., 2004. New developments in museum-based informatics and applications in biodiversity analysis. *TRENDS in Ecology and Evolution*. 19(9), 497-503.
- Gonzalez, A. and Chaneton, E.J., 2002. Heterotroph Species Extinction, Abundance and Biomass Dynamics in an Experimentally Fragmented Microecosystem. *Journal of Animal Ecology*. 71(4), 594-602.
- Guisan, A., Broennimann, O. Enger, R., Vust, M., Yoccoz, N. G., Lehmann, A., Zimmermann, N. E., 2005. Using Niche-Based Models to Improve the Sampling of Rare Species. *Conservation Biology*. 20(2), 501-511
- Guisan, A., Lehmann, A., Ferrier, S., Austin, M., Overton, J. M. C., Aspinall, R., Hastie, T., 2006. Making better biogeographical predictions of species' distributions. *Journal of Applied Ecology*. 43, 386-392.
- Guisan, A and Thuiller, W., 2005. Predicting species distribution: offering more than simple habitat models. *Ecology Letters*. 8, 993-1009.
- Guisan, A., and Zimmermann, N. E., 2000. Predictive habitat distribution models in ecology. *Ecological modeling*. 135, 147-186.
- Hart, S. P., Marshall, D. J. 2009. Spatial arrangement affects population dynamics independent of community competition. *Ecology*. 90(6), 1485-1491.
- Henle, K., Davies, K.F., Kleyer, M., Margules, C., Settele, J., 2004. Predictors of species sensitivity to fragmentation. *Biodiversity and Conservation*. 13, 207-251.
- Hernandez, P. A., Graham, C. H., Master, L. L., Albert, D. L., 2006. The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography*. 29, 773-785.
- Hinojosa-Diaz, I. A., Feria-Arroyo, T. P., Engel, M. S., 2009. Potential distribution of orchid bees outside their native range: The case of *Eulaema polychroma* (Mocsary) and *Euglossa*

*viridissima* Friese in the USA (Hymenoptera: Apidae). Diversity and Distributions. 15, 421-428.

Hirzel, A., Le Lay, G., Helfer, V., Randin, C., Guisan, A., 2006. Evaluating the ability of habitat suitability models to predict species presences. Ecological modeling. 199, 142-15.

Hooper, D. U., Chapin, F. S., Ewel, J. J., Hector, A., Inchausti, P., Lavorel, S., Lawton, J. H., Lodge, D. M., Loreau, M., Schmid, B., Setälä, H., Symstad, A. J., Vandermeer, J., Wardle, D. A., Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. Ecological Monographs. 75(1), 3-35.

Jakobsson, A., Padron, B., Traveset, A., 2009. Competition for pollinators between invasive and native plants: Effects of spatial scale of investigation. Ecoscience. 16(1), 138-141.

Jukes, M. R., Peace, A. J., Ferris, R., 2001. Carabid beetle communities associated with coniferous plantations in Britain: the influence of site, ground vegetation and stand structure. Forest Ecology and Management. 148, 271-86.

Kadoya, T., Ishii, H. S., Kikuchi, R., Suda, S., Washitani, I., 2009. Using monitoring data gathered by volunteers to predict the potential distribution of the invasive alien bumblebee *Bombus terrestris*. Biological Conservation. 142, 1011-1017.

C. Kremen, C., Cameron, A., Moilanen, A., Phillips, S. J., Thomas, C. D., Beentje, H., Dransfield, J., Fisher, B. L., Glaw, F., Good, T. C., Harper, G. J., Hijmans, R. J., Lees, D. C., Louis, E., Nussbaum, R. A., Raxworthy, C. J., Razafimpahanana, A., Schatz, G. E., Vences, M., Vieites, D. R., Wright, P. C., Zjhra, M. L., Aligning Conservation Priorities Across Taxa in Madagascar with High- Resolution Planning Tools. Science. 320, 222-226.

Larsen, T.H., Williams, N.M., Kremen, C., 2005. Extinction order and altered community structure rapidly disrupt ecosystem functioning. Ecology letters. 8, 538-547.

Leathwick, J. R. and Austin, M. P., 2001. Competitive interactions between tree species in New Zealand's old growth indigenous forests. Ecology. 82(9), 2560-2573.

Leathwick, J.R., McC.Overton, J., McLeod, M., 2003. An Environmental Domain Classification of New Zealand and Its Use as a Tool for Biodiversity Management. Conservation Biology. 17(6), 1612-1623.

Loiselle, B. A., Howell, C. A., Graham, C. H., Goerck, J. M., Brooks, T., Smith, K. G., Williams, P. H., 2003. Avoiding Pitfalls of Using Species Distribution Models in Conservation Planning. Conservation Biology. 17(6), 1591-1600.

- Loreau, M., Naeem, S., Inchausti, P., Bengtsson, J., Grime, J. P., Hector, A., Hooper, D. U., Huston, M. A., Raffaelli, D., Schmid, B., Tilman, D., Wardle, D. A., 2001. Biodiversity and Ecosystem Functioning: Current Knowledge and Future Challenges. *Science*. 294, 804-808.
- Magura, T., 2002. Carabids and forest edge: spatial pattern and edge effect. *Forest Ecology and Management*. 157, 23-37.
- Major, R.E., Christie, F.J., Gowing, G., Cassis, G., Reid C.A.M., 2003. The effects of habitat configuration on arboreal insects in fragmented woodlands of south-eastern Australia. *Biological Conservation*. 113, 35-48.
- Mantel, N., 1967. The detection of disease clustering and a generalized regression approach. *Cancer Research*. 27, 209-220.
- Mingyang, L., Yunwei, J., Kumar, S., Stohlgren, T. J., 2008 .Modeling potential habitats for alien species *Dreissena polyorpha* in Continental USA. *Acta Ecologica Sinica*. 28(9), 4253-4258.
- Murcia, C., 1995. Edge effects in fragmented forests: implications for conservation. *TREE*. 10(2), 58-62.
- Pawar, S., Koo, M. S., Kelle, C., Ahmed, M. F., Chauduri, S., Sarkar, S., 2007. Conservation assessment and prioritization of areas in Northeast India: priorities for amphibians and reptiles. *Biological Conservation*. 136, 346-361.
- Pearce, J. L. and Boyce, M. S., 2006. Modelling distribution and abundance with presence-only data. *Journal of Applied Ecology*. 43, 45-412.
- Pearce, J. and Ferrier, S., 2000. Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological modeling*. 133, 225-245.
- Phillips, J., 2009, A Brief Tutorial on Maxent. (Online) [www.cs.princeton.edu/~schapire/maxent/tutorial/tutorial.doc](http://www.cs.princeton.edu/~schapire/maxent/tutorial/tutorial.doc). Date accesses 15/03/09.
- Phillips, J., Dudik, M., Schapire, R. E., 2004. A Maximum Entropy Approach to Species Distribution Modelling. *Proceedings of the 21<sup>st</sup> International Conference on Machine Learning*, Banff, Canada.
- Phillips, S. J., Anderson, R. P., Schapire, R. E., 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling*. 190, 231-259.

- Phillips, S. J. and Dudik, M., 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*. 31, 161-175.
- Podchong, S., Schmidt-Vogt, D., Honda, K., 2009. An improved approach for identifying suitable habitat of Sambar Deer (*Cervus unicolor* Kerr) using ecological niche analysis and environmental categorization: Case study at Phu-Khieo Wildlife Sanctuary, Thailand. *Ecological Modelling*. 220, 2103-2114.
- Preston, K. L., Rotenberry, J. T., Redak, R. A., Allen, M. F., 2008. Habitat shifts of endangered species under altered climate conditions: importance of biotic interactions. *Global Change Biology*. 14, 2501-2515.
- Pulliam, H. R., 2000. On the relationship between niche and distribution. *Ecology Letters*. 3, 349-361.
- Puschendorf, R., Carnaval, A. C., VanDerWal, J., Zumbado-Ulate, H., Chaves, G., Bolanos, F., Alford, R. A., 2009. Distribution models for the amphibian chytrid *Batrachochytrium endrobatidis* in Cost Rica: proposing climate refuges as conservation tools. *Diversity and Distributions*. 15, 401-408.
- Randin, C. F., Dirnbock, T., Dullinger, S., Zimmermann, N. E., Zappa, M., Guisan, A., 2006. Are niche-based species distribution models transferable in space? *Journal of Biogeography*. 33, 1689-1703.
- Ritchie, E. G., Martin, J. K., Johnson, C. N., Fox, B. J., 2009. Separating the influences of environment and species interactions on patterns of distribution and abundance: competition between large herbivores. *Journal of Animal Ecology*. 78, 724-731.
- Rose, H., Learmount, J., Taylor, M., Wall, R., 2009. Mapping risk foci for endemic sheep scab. *Veterinary Parasitology*. In Press.
- Rosenlew, H. and Roslin, T., 2008. Habitat fragmentation and the functional efficiency of temperate dung beetles. *Oikos*. 117, 1659-1666.
- Rushton, S. P., Ormerod, S. J., Kerby, G., 2004. New paradigms for modelling species distributions? *Journal of applied zoology*. 41, 193-200.
- Singh, N. J., Yoccoz, N. G., Bhatnager, Y. V., Fox, J. L., 2009. Using habitat suitability models to sample rare species in high-altitude ecosystems: a case study with Tibetan argali. *Biodiversity Conservation*. Published online 10 March 2009.



- Soberon, J., 2005. Interpretation of models of fundamental ecological niches and species distributional areas. *Biodiversity Informatics*. 2, 1-10.
- Su, J. C., Debinski, D. M., Jakubauskas, M. E., Kindscher, K., 2004. Beyond Species Richness: Community Similarity as a Measure of Cross-Taxon Congruence for Coarse-Filter Conservation. *Conservation Biology*. 18 (1), 167-173.
- Suarez-Seoane, S., Garcia de la Morena, E. L., Prieto, M. B., Osborne, P. E., de Juana, E., 2008. Maximum entropy niche-based modelling of season changes in little bustard (*Tetrax tetrax*) distribution. *Ecological Modelling*. 219, 17-29.
- Taboada, A., Kotze, D. J., Tarrega, R., Salgado, J. M., 2006. Traditional forest management: Do carabid beetles respond to human-created vegetation structures in an oak mosaic landscape? *Forest Ecology and Management*. 237, 436-449.
- Thuiller, W., Richardson, D. M., Pysks, P., Midgley, G. F., Hughes, G. O., Rouget, M., 2005. Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. *Global Change Biology*. 11, 2234-2250.
- Tscharntke, T., Steffan-Dewenter, I., Kruess, A., Thies, C., 2002. Characteristics of insect populations on habitat fragments: A mini review. *Ecological research*. 17, 229-239.
- Valladares, G., Salvo, A., Cagnolo, L., 2006. Habitat Fragmentation Effects on Tropic Processes of Insect-Plant Food Webs. *Conservation Biology*. 20(1), 212-217.
- Wang, Y., Xie, B., Wan, F., Xiao, Q., Dai, L., 2007. The Potential Geographic Distribution of *Radopholus similis* in China. *Agricultural Sciences in China*. 6(12), 1444-1449.
- Watling, J.I. and Donnelly, M.A., 2006. Fragments as Islands: a synthesis of Faunal Responses to Habitat Patchiness. *Conservation Biology*. 20(4), 1016-1025.
- Williams, J N., Seo, C., Thorne, J., Nelson, J. K., Erwin, S., O'Brien, J. M., Schwartz, M. W., 2009. Using species distribution models to predict new occurrences for rare plants. *Diversity and Distributions*. 15, 565-576.
- Wintle, B. A. and Elith, J., Potts, J. M., 2005. Fauna habitat modeling and mapping: A review and case study in the Lower Hunter Central Coast region of NSW. *Austral Ecology*. 30, 719-738.

Wisz, M. S., Hijmans, R. J., Li, J., Peterson, A. T., Graham, C. H., Guisan, A., 2008. Effects of sample size on the performance of species distribution models. *Diversity and Distributions*. 14, 763-773.

Yost, A. C., Petersen, S. L., Gregg, M., Miller., 2008 . Predictive modeling and mapping sage grouse (*Centrocercus urophasianus*) nesting habitat using Maximum Entropy and a long-term dataset from Southern Oregon. *Ecological Informatics*. 3, 375-386.

Zhang, J., Hao, Z. Q., Song, B., Li, B. H., Wang, X. G., Ye, J., 2009. Fine-scale species co-occurrence patterns in an old-growth temperate forest. *Forest Ecology and Management*. 257(10), 2115-2120.

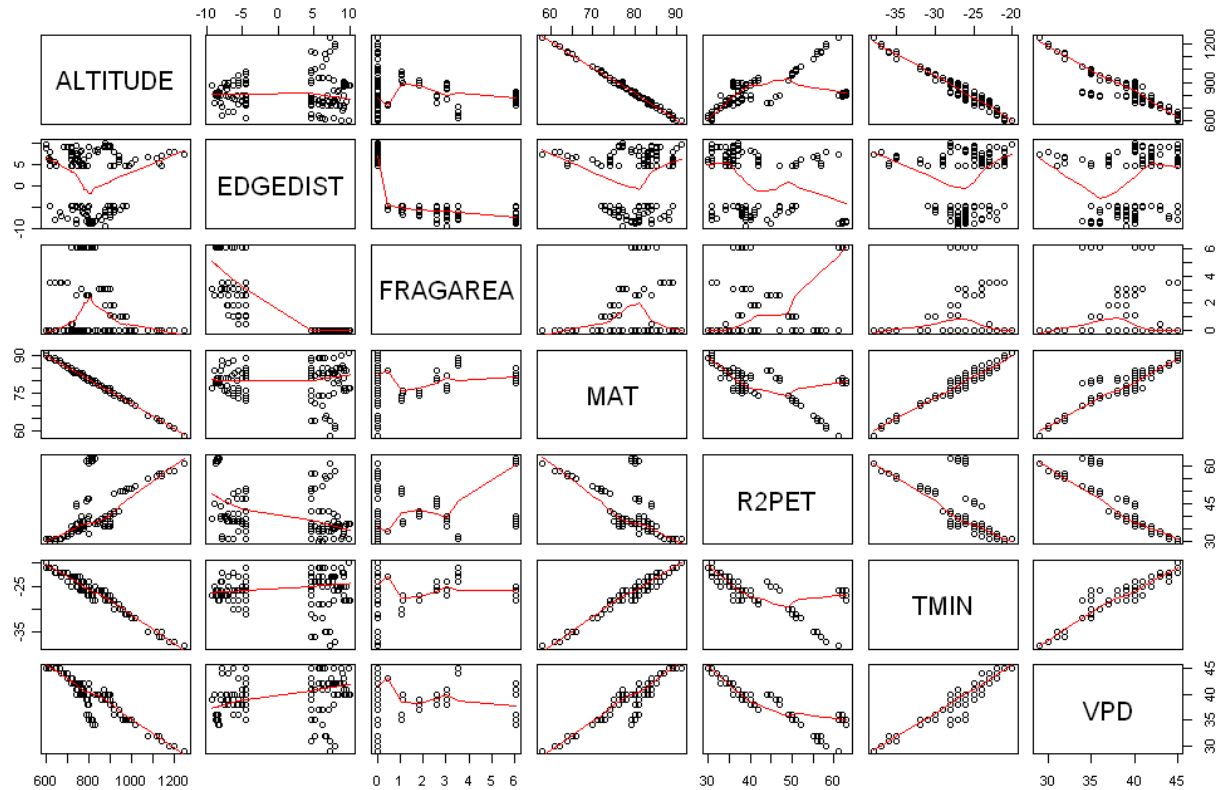
## **8.0 Software**

Maxent 3.3.1. (Online) [www.cs.princeton.edu/~schapire/maxent/](http://www.cs.princeton.edu/~schapire/maxent/). Date downloaded 14/03/09.

R 2.7.1. (Online) <http://cran.r-project.org/bin/windows/base/old/2.7.1/>. Date downloaded 10/11/08.

## 9.0 Appendices

### 9.1 Appendix A



**Fig 8;** Pearson's product-moment correlations between environmental variables that show a moderate to strong correlation ( $r = \pm 0.6-1$ ). These variables include altitude, distanced to edge (edgedist), fragment area (fragarea), maximum annual temperature (mat), potential evapotranspiration (r2pet), minimum temperature (tmin) and mean October vapor pressure deficit (vpd).

	ALTITUDE	DEFICIT	EDGEDIST	FRAGAREA	JUNES	MAS	MAT	R2PET	ROADDIST	RIVERDIST	TMIN
DEFICIT	-0.5864										
EDGEDIST	0.0346	0.2288									
FRAGAREA	-0.1194	-0.2260	<b>-0.8331</b>								
JUNES	0.4639	-0.4804	-0.3596	0.3874							
MAS	-0.5351	0.4574	0.3580	-0.2428	-0.3580						
MAT	<b>-0.9977</b>	0.5739	-0.0307	0.1227	-0.4638	0.5423					
R2PET	0.6236	-0.5014	-0.4063	0.4455	0.6763	-0.7935	-0.6272				
ROADDIST	-0.2337	-0.1535	0.0684	-0.0758	-0.2958	0.1209	0.2359	-0.4530			
RIVERDIST	-0.0489	-0.0795	-0.4694	0.6127	0.3123	-0.4157	0.0560	0.5246	-0.2639		
TMIN	<b>-0.9745</b>	0.5823	0.0214	0.0251	-0.5866	0.5247	<b>0.9725</b>	-0.6966	0.3114	-0.0370	
VPD	<b>-0.8941</b>	0.6189	0.2154	-0.1936	-0.6707	0.7285	<b>0.8963</b>	<b>-0.8970</b>	0.3571	-0.2571	<b>0.9267</b>

**Table 11;** Correlation matrix for all continuous environmental variables using Pearson's product-moment correlation. Bold indicates a strong correlation ( $r \geq 0.8, < -0.8$ ).

	ACIDP	AGE	CALCIUM	DOCCONS	INDURATION
AGE	-0.2356				
CALCIUM	0.2356	<b>-1</b>			
DOCCONS	0.0394	0.0808	-0.0808		
INDURATION	<b>-0.9473</b>	-0.0882	0.0882	-0.0670	
LCDB	0.3067	-0.0712	0.0712	-0.4495	-0.2909

**Table 12;** Correlation matrix for all categorical environmental variables using Pearson's product-moment correlation. Bold indicates a strong correlation ( $r \geq 0.8, < -0.8$ ).

<b>GRID ID.</b>	<b>SITES</b>	<b>GRID ID</b>	<b>SITES</b>	<b>GRID ID</b>	<b>SITES</b>
1	1:3	45	77:79	89	150
2	4,5	46	83,84,85,87,88,89	90	151,152,153,162
3	6	47	86	91	154,155
4	7	48	90:94	92	156
5	8,9	49	95	93	157
6	10,11	50	96	94	158
7	12	51	97	95	159
8	13	52	98	96	160
9	14	53	99	97	161
10	15:17	54	100	98	163:165
11	18,19	55	101:104	99	166
12	20	56	105	100	167
13	21	57	106	101	168
14	22:24	58	107	102	169
15	25:26	59	108	103	170
16	27	60	109	104	171
17	28	61	110	105	172,179,180,181
18	29,40,41	62	111	106	173:176
19	30:32	63	112:115	107	177
20	33,34	64	116	108	178
21	35	65	117	109	182
22	36	66	118	110	183
23	37	67	119	111	184
24	38	68	120	112	185:189
25	39	69	121	113	190,201,202,203
26	42,43	70	122:125	114	191:193
27	44	71	126	115	194:195
28	45	72	127	116	196
29	46	73	128	117	197
30	47	74	129	118	198
31	48	75	130	119	199
32	49	76	131	120	200
33	50:58	77	132,142,143,144	121	204:205
34	59,68,69,70,71	78	133	122	206
35	60:63	79	134:136	123	207
36	64	80	137	124	208
37	65	81	138	125	209
38	66	82	139	126	210
39	67	83	140	127	211:215,217:218
40	72	84	141	128	216
41	73	85	145:146	129	219:220
42	74	86	147	130	221
43	75	87	148	131	222:226
44	76,80,81,82	88	149	132	227:233

**Table 13;** grid square ID and the trap sites that are located in that individual grid square which were subsequently averaged for analysis.

### *Bray-Curtis dissimilarity between forest control sites and sample sites*

Grid square community values for trap sites at control sites (both forest and matrix controls) situated at -1024, -512, 0, +512 and +1024 were retained with the grid squares that contained the remaining control site values deleted, this left 113 grid squares for analysis. For the control sites at 0 distances, the averaged data for the grid square that contained sites 90:94 was replaced with the community values for site 90 (0 distance at the matrix control) and the grid square that contained the averaged values for sites 151:153 and 162 were replaced with the community values from site 151(0 distance at the forest control).

### *9.2 Appendix B*

CD containing 245 individual species distribution maps

Examples include from Fig 2;

(a) *Mecodema rugiceps* defined as a forest specialist (sp0011)

(b) *Mecodema fulgidum* defined as a matrix specialist (sp0018),

The effects and mechanisms of distance from edge on both a and b have been further investigated in Ewers, (2008).

(c) *Costelytra brunneum* defined as a generalist (sp1082)

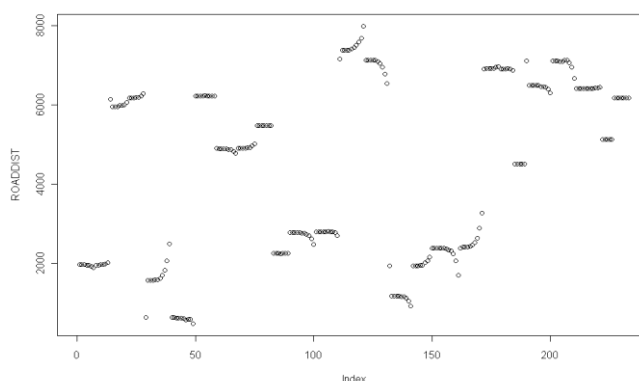
(d) *Costelytra zealandicum* defined as a matrix specialist (sp0946)

Habitat preferences defined by Ewers et al., 2007 Appendix *Ecological Archives* E088-007-A3.

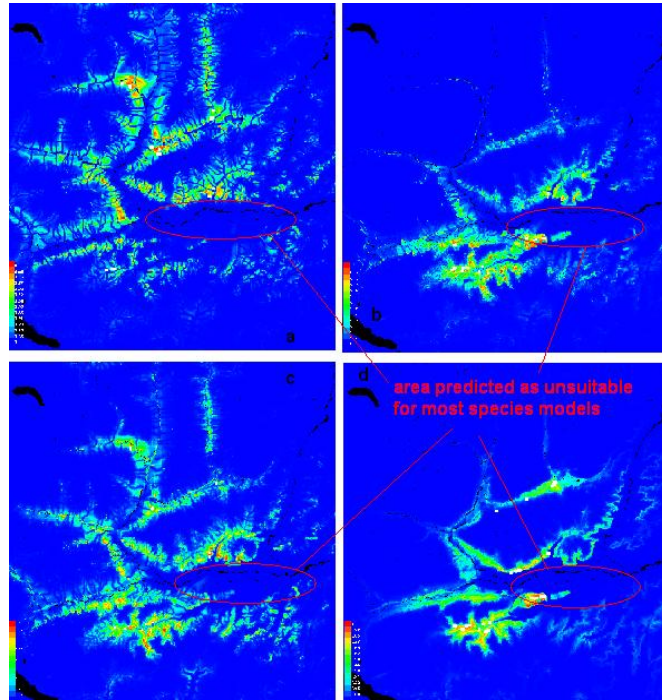
9.3 Appendix C

Variable	Contribution%	HighGain%	DecreaseGain%
Acidp	1.11	0	0
Age	0.17	0	0
Altitude	2.67	16.73	0
Calcium	0.13	0	0
Deficit	10.04	2.86	11.43
Doccons	8.14	4.49	0.82
Edgedist	9.61	15.51	9.39
Fragarea	5.19	2.04	1.22
Induration	1.89	1.22	0
Junes	5.1	1.63	0.41
Lcdb	13.97	11.43	11.84
Mas	3.35	1.22	1.63
Mat	0.73	1.63	0
r2pet	2.89	9.39	2.04
Riverdist	3.54	0	1.22
Roaddist	16.36	12.24	60
Tmin	14.99	19.59	0

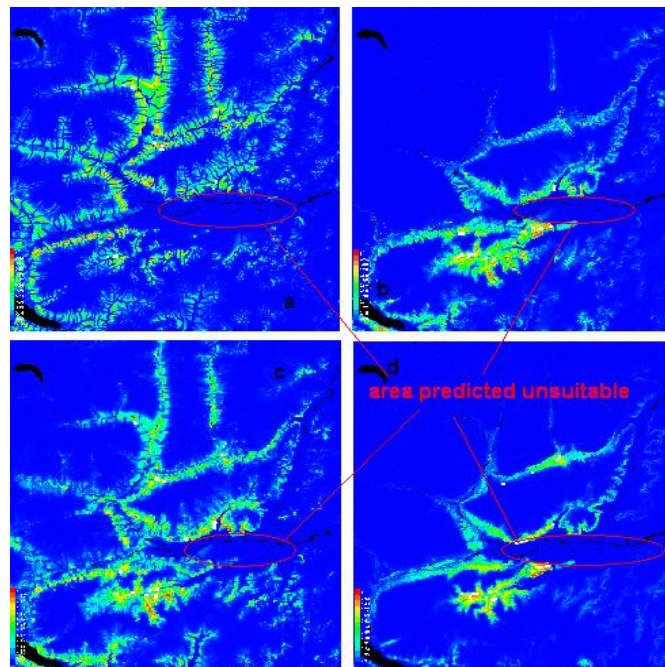
**Table 14;** mean variable percentage contribution, variable which had the highest “gain” when used in isolation and variable that decreased the “gain” the most when excluded from the model for each of the 245 species.



**Fig 9;** plot of distance from road indicates no explicit roadside or distance bias in sampling.



**Fig 10;** Example of the area around the road predicted to be unsuitable for most of the species modelled. Blue indicates low probability of occurrence.



**Fig 11;** Example of the area around the road predicted to be unsuitable for most species modelled even when distance from road was excluded from the model.



<b>Correlations</b>			
Acidp	0.288	Junes	0.289
Age	0.080	Lcdb	-0.139
Altitude	0.085	Mas	-0.401
Calcium	-0.119	Mat	0.022
Deficit	-0.032	r2pet	-0.150
Doccons	-0.046	Riverdist	0.058
Edgedist	-0.247	Roaddist	1.00
Fragarea	-0.273	Tmin	-0.201
Induration	-0.201	Vpd	0.105

**Table 15;** Distance from road correlated with the other environmental variables. Environmental variable values taken from the 147 models where distance from road when isolated from models decreased the “gain” the most. Pearson’s product-moment correlation  $r$  values to 3dp. Shows no strong correlation even with sites where distance from road when excluded decreased the models “gain” the most.

## 10.0 Suggested additional reading and the future of species distribution modelling

Araujo, M. B. and Guisan, A., 2006. Five (or so) challenges for species distribution modelling. *Journal of Biogeography*.33, 1677-1688.

Araujo, M. B. and Pearson, R. G., 2005. Equilibrium of species' distributions with climate. *Ecography*. 28, 693-695.

Barbosa, A. M., Real, R., Vargas, J. M., 2009. Transferability of environmental models in geographic space: The case of the Iberian desman (*Galemys pyrenaicus*) in Portugal and Spain. *Ecological Modelling*. 220, 747-754.

Ewers, R.M. and Didham, R.K., 2006. Confounding factors in the detection of species responses to habitat fragmentation. *Ecological Reviews*. 81, 117-142.

Faith, D.P., Ferrier, S., Williams, K.J., 2008. Getting biodiversity intactness indices right: ensuring that “biodiversity” reflects “diversity”. *Global Change Biology*. 14, 207-217.

Ferrier, S., Manion, G., Elith, J., Richardson, K., 2007. Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Diversity and Distributions*. 13, 252-264.

Hanski, I. and Ovaskainen, O., 2000. The metapopulation capacity of a fragmentation landscape. *Nature*. 404, 755-758.

Hirzel, A. H., Hausser, J., Chessel, D., Perrin, N., 2002. Ecological-Niche Factor Analysis to Compute Habitat-Suitability Maps without Absence Data? *Ecological Society of America*. 83(7), 2027-2036.

Ovaskainen, O. and Hanski, I., 2003. How much does an individual habitat fragment contribute to metapopulation dynamics and persistence?. *Theoretical Population Biology*. 64, 481-495.

Peterson, A. T., Papes, M., Eaton, M., 2007. Transferability and model evaluation in ecological niche modeling: a comparison of GARP and Maxent. *Ecography*. 30, 550-560.

Phillips, S. J., 2008. Transferability, sample selection bias and background data in presence-only modelling: in response to Peterson et al. (2007). *Ecography*. 31, 272-278.

Thuiller, W., 2004. Pattern and uncertainties of species' range shifts under climate change. *Global Change Biology*. 10, 2020-2027.

VanDerWal, J. Shoo, L. P., Graham, C. Williams, S. E., 2009. Selecting pseudo-absence data for presence only distribution modeling: How far should you stray from what you know?. *Ecological Modelling*. 220, 589-594.