BIOLOGICAL CONSERVATION 142 (2009) 364–374

available at www.sciencedirect.com

ScienceDirect

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

Using community observations to predict the occurrence of malleefowl (*Leipoa ocellata*) in the Western Australian wheatbelt

Blair Parsons[a,c,*], *Jeff Short*[b], *J. Dale Roberts*[c]

[a]CSIRO, Sustainable Ecosystems, Private Bag 5, Wembley, WA 6913, Australia
[b]Wildlife Research and Management Pty Ltd., P.O. Box 1360, Kalamunda, WA 6926, Australia
[c]School of Animal Biology M092, University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia

A R T I C L E I N F O

Article history:
Received 11 June 2008
Received in revised form
14 October 2008
Accepted 20 October 2008
Available online 3 December 2008

Keywords:
Generalised additive model
Generalised linear model
Prediction
Presence-only
Species distribution
Bird habitat

A B S T R A C T

The Malleefowl is a ground-dwelling bird species that has declined in distribution and abundance in Australia since European settlement. These declines have been exacerbated in the Western Australian wheatbelt by the extensive clearing of native vegetation for agricultural development. A wealth of opportunistic, presence-only data exists for this species but absence data required for distribution modelling is lacking. This situation is typical of many species distribution datasets. We sought to establish the distribution of malleefowl within the Western Australian wheatbelt (210000 km^2) and their choice of habitat within this broad region. We supplemented a large presence-only dataset of malleefowl sightings with absence data derived from a bird atlas scheme and used these data to effectively predict the distribution of the species for the wheatbelt using a combined GAM/GLM approach. Both datasets were derived largely from community sightings. The distribution of malleefowl within the Western Australian wheatbelt was associated with landscapes that had lower rainfall, greater amounts of mallee and shrubland that occur as large remnants, and, lighter soil surface textures. This study illustrates how community knowledge, coupled with solid ecological understanding, can play a key role in developing the knowledge base to inform conservation and management of species in agricultural landscapes.

1. Introduction

Malleefowl *Leipoa ocellata* (Gould), like many bird species occurring in agricultural landscapes in Australia, have exhibited a decline in distribution and abundance since European settlement of Australia (Woinarski and Braithwaite, 1990; Benshemesh, 2000). Nationally, the species is listed as "vulnerable" under the *Environment Protection and Biodiversity Conservation Act 1999* and in Western Australia it is listed as "fauna that is rare or is likely to become extinct". Substantial investment has been directed towards studying the biology of malleefowl, but knowledge of the distribution and habitat preferences of the species is patchy. Localised, long-term studies fall within five biogeographical regions, whereas the range of the species is known to extend across at least 32 biogeographical regions (Benshemesh, 2000), each with contrasting characteristics (Benshemesh, 2000; Garnett and Crowley, 2000).

* *Corresponding author:* Address: CSIRO, Sustainable Ecosystems, Private Bag 5, Wembley, WA 6913, Australia. Tel.: +61 (0)8 9333 6451; fax: +61 (0)8 9333 6444.
E-mail addresses: blair.parsons@csiro.au (B. Parsons), jeff.short@wildliferesearchmanagement.com.au (J. Short), droberts@cyllene.uwa.edu.au (J.D. Roberts).

doi:10.1016/j.biocon.2008.10.032

Spatial modelling techniques are known to perform well for species that are relatively abundant, have relatively large breeding ranges, are territorial and are associated with terrestrial habitats (Garrison and Lupo, 2002), all features of the malleefowl. Spatial modelling of species distributions is typically conducted using presence–absence data but data availability often limits the extent to which such models can be created. Further, it is typically difficult or impossible to collect new occurrence data for a species across broad areas (Araújo and Guisan, 2006). To address a lack of species absence data researchers have modelled distributions using environmental envelope or profile techniques (e.g. Busby, 1986; Hirzel et al., 2002; Phillips et al., 2006). However, outcomes of these models are limited in their practical application (Austin, 2002), and can be inaccurate for species with broad ecological requirements (Brotons et al., 2004). The lack of absence data has sometimes been addressed by creating "pseudo-absences" randomly, or based on the outcome of an environmental profiling technique such as ecological niche factor analysis (Zaniewski et al., 2002; Engler et al., 2004; Pearce and Boyce, 2006). There is some concern that the use of such methods is potentially inaccurate, circular in nature, and merely acting to model the niche of presence data itself (Lütolf et al., 2006). More appropriate techniques must be employed to obtain absence data to complement the presence data if we are to extend our knowledge beyond that of describing environmental attributes of the occupied niche. A suitable approach is to create absences based on sites where collections or surveys have taken place but have not recorded the target species (Lütolf et al., 2006).

The malleefowl has become an iconic species due to its unique biology (Frith, 1956, 1962), and unmistakeable appearance. It is commonly sighted feeding along roads and venturing into paddocks, generating much presence data for the Western Australian wheatbelt. Because malleefowl are well known and relatively easy to detect, sites with extensive bird survey data can also generate reliable absence data thus allowing presence–absence modelling of the species' distribution.

The purpose of this study was to model the contemporary distribution of malleefowl within the Western Australian wheatbelt. A primary motivation was to demonstrate an approach for modelling a species' distribution by supplementing presence-only data with absences derived from an Atlas program. We also sought to provide useful information on malleefowl occurrence to inform conservation action for the species within the Western Australian wheatbelt and across other similar bioregions. More specifically, we aimed to predict the occurrence of the species across the region using broad-scale environmental variables, including climate, soil and vegetation extent using data at a 1 km^2 unit of resolution.

2. Materials and methods

2.1. Malleefowl – biology and current status

The malleefowl (*L. ocellata*) is a large (~2 kg), sedentary, ground-dwelling bird that uses a combination of fermentation and solar radiation to incubate its eggs in mounds (Frith, 1956). Detailed accounts of breeding, incubation, mound construction and temperature regulation are provided by Frith (1959).

The range of malleefowl spans much of southern Australia. Its range coves most of the southern half of Western Australia including much of the region primarily used for wheat production: the wheatbelt (Benshemesh, 2000; Barrett et al., 2003). The western margin of its range contracted during the mid twentieth century because of habitat loss and degradation (Parsons et al., 2008) but malleefowl are still widespread throughout the wheatbelt region. Historically, they were found in most vegetation communities within the Western Australian wheatbelt, but are most commonly observed in mallee (multi-stemmed *Eucalyptus* species), Acacia shrublands and scrub thickets (Storr, 1991). The species was known to occasionally occur in open woodlands (e.g. York gum *Eucalyptus loxophleba* and gimlet *E. salubris*) in Western Australia (Crossman, 1909; Leake, 1962), but occurrence in these vegetation types is uncommon today.

2.2. Study area

The Western Australian wheatbelt extends from north of Geraldton to east of Esperance in south-west Western Australia (20 618 000 ha study area, Fig. 1). It is bounded by the 300 mm and 600 mm annual rainfall isohyets to the east and west, respectively and land use consists largely of cropping (wheat, barley, and canola) and the grazing of sheep (Saunders et al., 1993). Over 93% of the native vegetation has been removed in less than 100 yr (Saunders et al., 1993) resulting in a highly fragmented landscape, consisting of small and isolated islands of native vegetation in a matrix of cropping and grazing lands. Woodlands on heavy soil were cleared first and most extensively, followed by areas dominated by mallee (Burvill, 1979). The "light lands", sandy or gravelly soils often supporting shrub-thickets, were the last to be cleared, typically after the 1950s. The Western Australian wheatbelt represents a high contrast landscape (Fischer and Lindenmayer, 2006) with habitat within the matrix remaining as discrete relatively intact remnants as opposed to scattered vegetation. The intensity of clearing has resulted in it becoming one of Australia's most stressed landscapes (National Land and Water Resources Audit, 2001). The soil and corresponding vegetation patterns of the region are complex and patchy (Burvill, 1979; Dirnbock et al., 2002) with patterns considered "strikingly unpredictable" at a local scale (Dirnbock et al., 2002).

2.3. Malleefowl dataset

This study made use of a comprehensive presence-only dataset of sightings of malleefowl obtained from community organisations and government agencies in Western Australia. These organisations included the Malleefowl Preservation Group and other regional community groups dedicated to malleefowl conservation, the Western Australian Department of Environment and Conservation, the Western Australian Museum and Birds Australia (New Atlas of Australian Birds, Barrett et al., 2003). Records were checked for reliability (e.g. adequate description of species and/or location) prior to

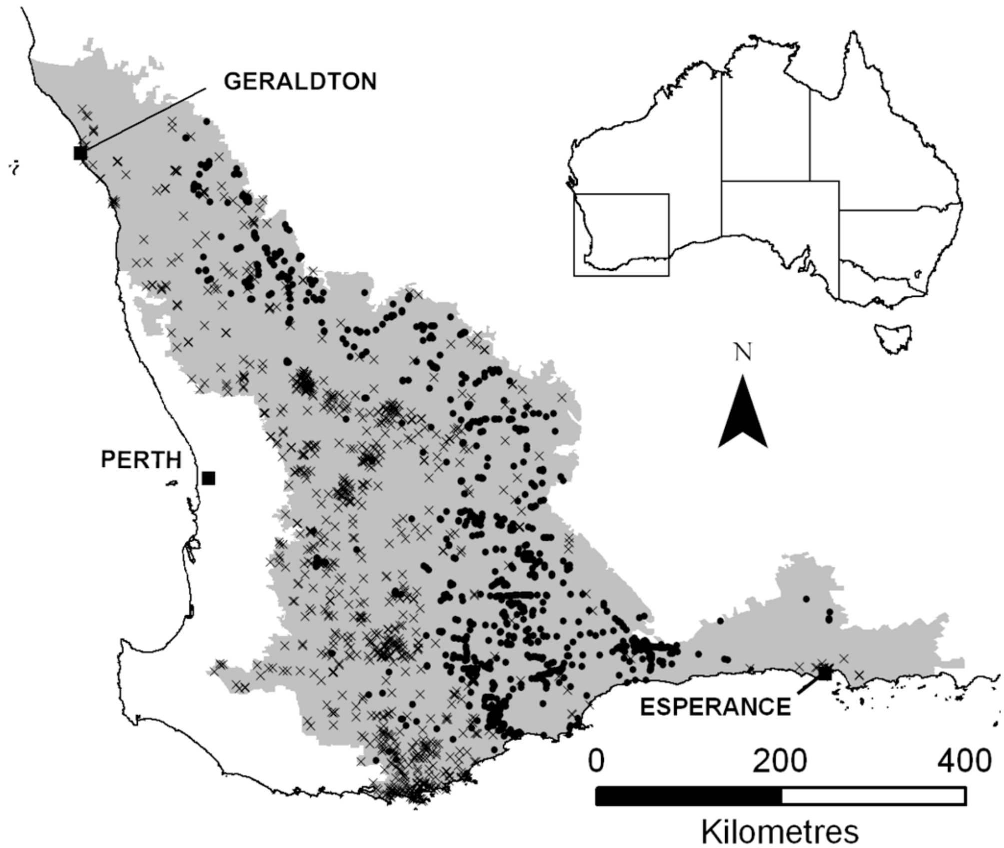

Fig. 1 – Map of south-west Western Australia showing the study area (shaded area) and distribution of malleefowl presences (circles) and absences (crosses) used in this study.

inclusion in analyses. Data used in analyses were accurate to within 1 km or less and included sightings of individual birds, mounds (both active and inactive) and road kills for the period 1990–2005. All locations of sightings within the bounds of the Western Australian wheatbelt were aggregated into 1 km^2 grid cells to remove duplicate records.

The presence-only dataset was supplemented with absence data obtained from the New Atlas of Australian Birds (Barrett et al., 2003). This atlas compiled community records for the period 1998–2002. We extracted all records within the bounds of the study area that were accurate to within 1 km, excluding nocturnal birds, waterbirds and birds of prey (i.e. species not likely to be reliably detected in surveys). Absences were generated from these records by selecting survey sites that met the following criteria: any cell greater than or equal to 2 km from any malleefowl presence location where there were Bird Atlas records for 25 or more bird species but no malleefowl records.

We selected an arbitrary record limit (25 bird species records) as there was no clear threshold for number of records above which malleefowl were more likely to be recorded. We randomly selected locations from this sample (i.e. to equal the number of presences) for use in modelling. This method of selecting locations where malleefowl were considered to be absent was similar to a method recommended by Graham et al. (2004); sampling locations for which surveys had been made, but the species was not recorded. The method was considered by Lütolf et al. (2006) to result in better model performance than selecting pseudo-absences randomly from background values (e.g. Milne et al., 2006). We also consider this to be an improvement over the seemingly circular technique of ENFA-weighted pseudo-absence (Zaniewski et al., 2002; Engler et al., 2004). We excluded 65 locations where environmental data were missing, resulting in a final dataset of 869 presences and 876 absences for use in modelling.

2.4. *Environmental variables*

It has been suggested that greater focus be given to the ecological basis for choosing candidate variables for species distribution models (Lehmann et al., 2002; Araújo and Guisan, 2006). We selected 19 variables that we believed were relevant to the ecology of malleefowl, for inclusion in an exploratory GAM modelling phase (Table 1). Relevant variables that were uniform across the study area (e.g. presence of the introduced red fox *Vulpes vulpes*) or highly variable over time (e.g. the presence of fire) were excluded from analysis.

Soil quality variables (variables 1–8, Table 1) were derived from proportional mapping of soil-landscape subsystems (Western Australian Department of Agriculture and Food, 2004). Proportional mapping has unmapped components which are described as a percentage of the map unit polygon. Using the soil quality GRAVEL as an example, a map unit polygon may have a 10% probability of having very few gravel fragments in the profile, 15% probability of few, 20% of common,

Table 1 – Environmental predictor variables used in exploratory GAM analysis.

			Units
Soil variables			
1	Flood risk	An index of the temporary covering of land by flood waters	n/a
2	Gravels	An index of the abundance of gravel present in the soil profile	n/a
3[a]	Surface stones/gravels	An index of the abundance of gravel and coarse fragments (>20 mm) present in the soil profile	n/a
4	Permeability	An index of the capacity of a material to transmit a fluid such as water	n/a
5[a]	Absorption	An index of absorption ability of the soil.	n/a
6	Rooting depth	An index of depth to which plant roots can penetrate	n/a
7	Surface salinity	An index intended to reflect current salinity status	n/a
8	Surface texture	An index reflecting the amount of clay in the top 10 cm of the soil profile.	n/a
Climate variables			
9	Mean annual rainfall	Mean annual rainfall as derived by BIOCLIM	mm
10	Mean annual temperature	Mean annual temperature as derived by BIOCLIM	°C
Vegetation variables			
11	Distance 500 ha	Distance to the nearest 500 ha patch of vegetation	m
12	Distance 200 ha	Distance to the nearest 200 ha patch of vegetation	m
13	Distance 100 ha	Distance to the nearest 100 ha patch of vegetation	m
14	Distance 75 ha	Distance to the nearest 75 ha patch of vegetation	m
15	Distance 50 ha	Distance to the nearest 50 ha patch of vegetation	m
16	Mallee/shrub 5 km	Proportion of the surrounding circle (radius 5 km) containing mallee or shrubland	%
17	Mallee/shrub 2 km	Proportion of the surrounding circle (radius 2 km) containing mallee or shrubland	%
18	Mallee/shrub 1 km	Proportion of the surrounding circle (radius 1 km) containing mallee or shrubland	%
19	Distance to road	Distance to the nearest road	m

a Denotes correlated variables removed prior to modelling.

25% of many, 30% of abundant, and a 0% probability of having none. To assign proportional soil quality information to a cell falling within a polygon, we assigned a value to these categories (e.g. very few = 1, few = 2, common = 3, many = 4, and abundant = 5), and assigned their contribution to the cell proportionally. The index for GRAVEL for the hypothetical cell is calculated below:

$$\mathrm{GRAVEL} = 0.1 \times 1 + 0.15 \times 2 + 0.2 \times 3 + 0.25 \times 4 + 0.3 \times 5 = 3.5$$

Mean annual temperature and mean annual rainfall (variables 9 and 10) were estimated using the BIOCLIM module within ANUCLIM 5.1 (Houlder et al., 2000).

All vegetation variables (variables 11–18) were based on mapped vegetation extent data for 2004 (Commonwealth of Australia, 2004), excluding areas modelled as at risk of salinity (Evans and Caccetta, 2000), as much of the vegetation in these areas was unlikely to be suitable for malleefowl (e.g. saline flats). Individual vegetation associations were simplified into four broad categories: (1) mallee; (2) shrubland/thicket; (3) woodland; and (4) scrub/heath. Inspection of the mallee and shrub/thicket extent variables showed the two habitats to be complementary: mallee existed primarily in the southern half of the study area, shrub/thicket primarily in the north. Malleefowl are known to inhabit both mallee and shrub-thicket vegetation (Storr, 1991; Benshemesh, 2000) so it was decided to combine these two categories into one variable (variables 16–18 in Table 1). We excluded woodland and scrub-heath extent variables because (a) woodland extent showed a negative relationship with malleefowl occurrence and would not result in predictions that could be easily applied to management and (b) scrub-heath extent did not span the study area adequately. Variables developed for modelling fell into two classes: (1) vegetation isolation (i.e. distance to the nearest remnant of 50, 75, 100, 200, or 500 ha) and (2) mallee/shrubland extent (i.e. amount of vegetation in the surrounding circle of radius 1, 2, or 5 km^2).

There is likely to be a significant bias towards malleefowl along roads given the opportunistic method of data collection. As such, distance to road (Variable 19) was estimated from a 1:1000000 scale map (Geoscience Australia, 2005), and included in the modelling to eliminate this source of bias.

2.5. *Exploratory GAM analysis*

We modelled the occurrence of malleefowl in the Western Australian wheatbelt using a combination of generalised additive models (GAM, Hastie and Tibshirani, 1990) and generalised linear models (GLM, McCullagh and Nelder, 1983). GAM was used in an exploratory sense: (1) to select variables that explained the greatest amount of total variation and (2) to examine the shape of the response curves for variables prior to input into a GLM. The benefit of using GAM was that it did not require an *a priori* knowledge of the shape of the response curves (Hastie and Tibshirani, 1990). Rather, the shape of the response is guided by the data itself (Ferrier et al., 2002; Denoël and Lehmann, 2006).

Prior to modelling, correlated variables ($r > 0.70$) were identified and the variable least relevant to malleefowl (i.e. least direct; *sensu* Austin and Meyers, 1996) removed. *Absorption* and *Permeability* were correlated but as the former was a derivative of the latter it was removed (Table 1). *Gravels* and *Surface Stones and Gravels* were also correlated but the latter variable included rocks and large boulders as part of the measure, an element not considered relevant to the habitat preferences of malleefowl (Benshemesh, 2000), and so was discarded.

GAM models were produced using GRASP (Lehmann et al., 2002) in the statistical package R (http://www.r-project.org/). Models were fitted using a binomial distribution with a backwards stepwise selection method employing Bayesian information criterion (BIC). BIC was chosen as it is known to impose heavier penalties on including additional terms in a model than Akaike's information criterion (Burnham and Anderson, 2002). It was our intent to create a model with fewer terms to remove unnecessary complexity.

We created 15 GAM models by entering all soil, road and climate variables with all combinations of one habitat proximity variable and one habitat extent variable. The habitat variables were added singly as they were not strictly independent. The combination that showed the lowest amount of residual deviation and lowest BIC value was selected as the final model. This selection process represented a method of "fine-tuning" the habitat extent and proximity variables to the scale most appropriate for malleefowl.

A weakness of GAM is that of over-fitting the data (Guisan et al., 2002). We addressed this problem by reducing the degrees of freedom (i.e. increasing the smoothness of response curves) used in the modelling process where possible. All variables selected in the final model were entered into GRASP using two, three and four degrees of freedom. If explanatory power was maintained despite a reduction in degrees of freedom, the model with lower degrees of freedom was used.

Model validation was conducted by measuring the area under the curve (AUC) of a receiver operating characteristic (ROC) plot (Fielding and Bell, 1997), thus quantifying the ability of the GAM model to describe the training data, as recommended by Lehmann et al. (2002).

2.6. *Predictive GLM analysis*

We predicted the occurrence of malleefowl in the Western Australian wheatbelt using GLM, after exploring the data and obtaining a final set of variables using GAMs. We used GLM because of its simplicity, ease of interpretation, and the ready availability of a model formula and the capacity to calculate confidence intervals (Wintle et al., 2005). These confidence intervals can be plotted spatially to provide a better idea of how uncertain a model is on the ground, which is important when model outcomes are to be applied for management (Guisan et al., 2006).

We inspected predictor response curves in the final GAM to identify which functional forms may be most appropriate for the GLM (Wintle et al., 2005). Where several candidate functional forms existed for a variable (e.g. linear, log and reciprocal), all were tested using an all-subsets approach and the form that explained the greatest amount of deviation retained. Model validation was conducted using ROC plots and employed bootstrapping as detailed by Wintle et al. (2005). Response plots for each of the variables were examined to determine whether they reflected an ecologically plausible relationship. The model residuals were also tested for spatial autocorrelation by calculating Global Moran's I (Fortin and Dale, 2005) on the points in a GIS. A Moran's I value of 1 indicates strong positive autocorrelation and −1 strong negative autocorrelation (Fortin and Dale, 2005). Finally, we created spatial predictions of malleefowl occurrence by plotting the final GLM equation (with upper and lower 95% confidence intervals) using a GIS.

3. Results

3.1. *Exploratory GAM analysis*

Of the 11 soil, climate and road variables initially developed for inclusion in the candidate model, all but two soil variables remained after removing correlated variables. These and all combinations of one habitat proximity and one habitat extent variable were entered into GRASP, resulting in the selection of four variables: (1) *mallee/shrub within surrounding circle of radius* 5 km, (2) *mean annual rainfall*, (3) *distance to nearest* 500 ha *remnant* and (4) *surface texture* (Table 2). All other variables were dropped from the model, including *distance to road*. For the habitat proximity variables, *distance to the nearest* 500 ha *remnant* explained the most variation, with the other variables (i.e. *distance to the nearest* 200, 100, 75, and 50 ha *remnant*) showing higher levels of residual deviation (Fig. 2). Similarly, *mallee/shrubland within the surrounding circle of radius 5 km* explained more variation than the 2 and 1 km variables.

Our GAM model explained 51% of the variation associated with malleefowl presence–absence within the Western Australian wheatbelt using four degrees of freedom for all variables. After reducing the number of degrees of freedom for each variable to two, the final GAM explained 49% of the variation. Partial response curves for the GAM are shown in Fig. 3. Cross-validation of the model resulted in an area under the ROC curve of 0.92, indicating that the model was able to effectively discriminate between areas of malleefowl presence and absence.

3.2. *Predictive GLM analysis*

The model with the lowest residual deviation (48.9%, Table 3) took the following form:

$$Malleefowlpresence/absence \sim \log(a) + b + c + d + d^2.$$

where a = *mallee/shrub within surrounding circle of radius* 5 *km*, b = *mean annual rainfall,* c = *distance to nearest* 500 *ha remnant*, and d = *surface texture.*

Cross-validation using bootstrapping resulted in an area under the ROC curve (AUC) of 0.93, indicating good discriminatory ability between malleefowl presence and absence.

Table 2 – Summary of the GAM model for malleefowl occurrence within the Western Australian wheatbelt, selected by a backward stepwise procedure.

Selected variable	d.f.	P-value	Functional form
Mallee/shrub 5 km	2	<0.001	+log
Annual rainfall	2	<0.001	−Linear
Distance 500 ha	2	<0.001	−Linear
Surface texture	2	<0.001	−Quadratic
Null deviation	2412		
Degree of freedom	1742		
Residual deviation	1202		
Residual d.f.: all variables	1734		

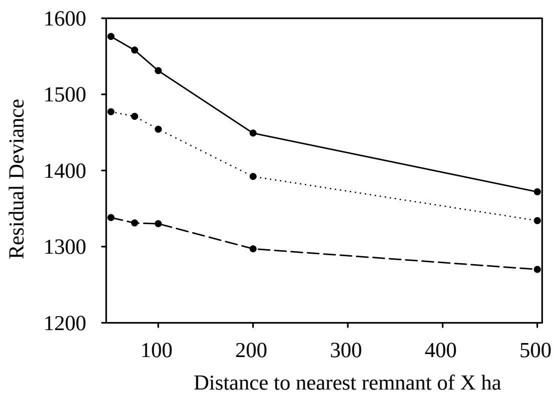

Fig. 2 – Effect of using different combinations of vegetation variables on the residual deviation for GAM of malleefowl occurrence within the Western Australian wheatbelt. Solid, dotted, and dashed lines = mallee/shrubland in surrounding circle of radius 1, 2, and 5 km, respectively.

Drop and alone contributions of predictors to the model are presented in Table 4. Fig. 4 displays plots of the predicted probability of occurrence of malleefowl against each variable (with all other variables held at their mean value). These plots show a positive relationship between malleefowl occurrence and the amount of mallee/shrubland in the surrounding 5 km with a negative relationship observed for both mean annual rainfall and the distance to the nearest 500 ha remnant. The relationship with soil surface texture was less obvious, with malleefowl occurrence predicted to decrease as the soil index (i.e. amount of clay in the soil) increased. The 95% confidence intervals illustrate that there is uncertainty associated with *surface texture* within the model. The residuals of the model were not spatially autocorrelated (Moran's $I = 3.25 \times 10^{-2}$).

We created a spatial plot of the model and 95% confidence intervals for the entire study area (Fig. 5). Our model identified 8689500 ha of the study area as having a 50% or greater chance of containing a malleefowl presence. At the upper and lower 95% confidence limits, this area could be as large as 15605200 ha or small as 577700 ha, respectively. If we restrict the above calculations to areas of remnant vegetation only and exclude farmland, we find that the model predicts 2016000 ha of remnant vegetation as having a 50% or greater chance of containing malleefowl, with 2752600 ha and 370600 ha representing upper and lower 95% confidence intervals, respectively. Of the 2016000 ha identified above, approximately 55% is part of the public estate (e.g. reserves, unallocated crown land and unmanaged reserves) with approximately 45% on private lands. Remnant vegetation formally reserved as part of the conservation estate (i.e. managed by the Western Australian Department of Environment and Conservation) makes up approximately 36% of the area predicted.

4. Discussion

The malleefowl is a threatened species that has persisted within the Western Australian wheatbelt, despite intense loss and fragmentation of its habitat over the past century.

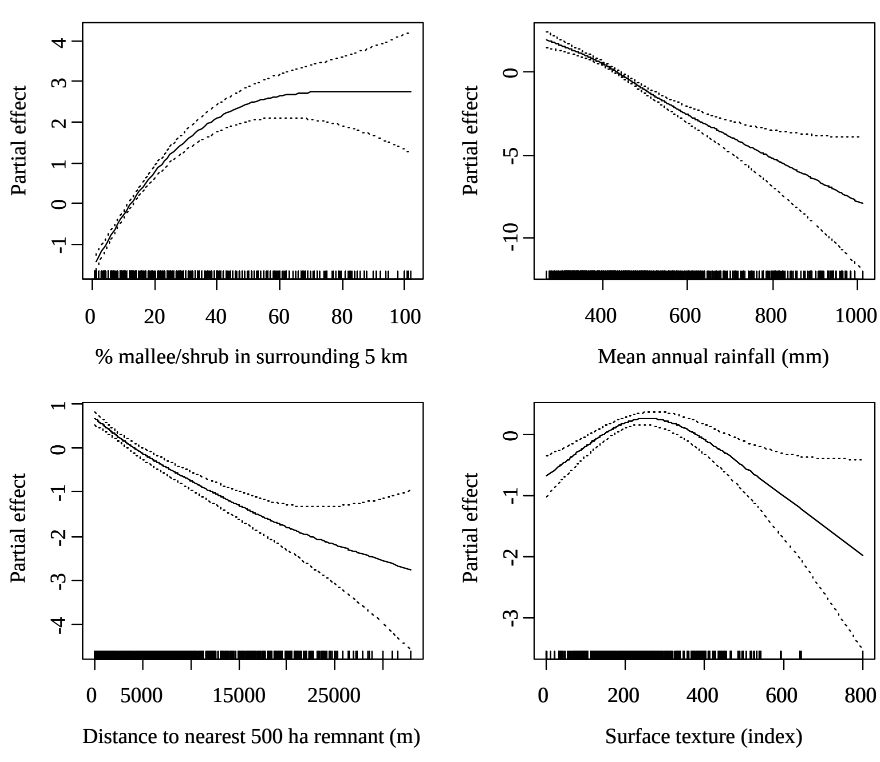

Fig. 3 – The response of malleefowl presence–absence to predictor variables in the GAM. y-axis = partial effect of each variable on presence of malleefowl (given the effect of other variables, S = smoothed, degrees of freedom given). Solid black lines = response; dashed lines = 95% confidence intervals for the mean response.

Table 3 – GLM model of malleefowl occurrence (869 presences, 876 absences, deviation explained = 51.1%, residual deviation 48.9%).

Term	Estimate	Standard error	P
Constant	2.53×100	5.03×10^{-1}	<0.001
Mean annual rainfall	-1.22×10^{-2}	1.07×10^{-3}	<0.001
log (distance 500 ha)	-1.43×10^{-4}	1.49×10^{-5}	<0.001
log (mallee/shrub 5 km)	2.39×10^{0}	1.45×10^{-1}	<0.001
Surface texture	1.03×10^{-2}	1.60×10^{-3}	<0.001
Surface texture^2	-1.98×10^{-5}	3.07×10^{-6}	<0.001
Null deviation	2419		
Degree of freedom	1744		
Residual deviation	1183		
Residual d.f.: all variables	1739		

Table 4 – Contributions of selected variables in the GLM model of malleefowl occurrence in the Western Australian wheatbelt.

Variables	Percent change in deviation	
	Drop contribution	Alone contribution
log (mallee/shrub 5 km)	15.2	40.1
Annual rainfall	7.6	16.1
log (distance 500 ha)	4.6	10.5
Surface texture^2	1.8	1.7

Significant effort is currently being invested in the conservation of the species within the region but ecological knowledge for planning and design of management has been lacking. For management action to be effective, practitioners require information on where malleefowl are most likely to occur and what governs their distribution.

4.1. *Use of unstructured data in spatial analyses*

The primary advantage of the approach employed in this study is in the utilisation of existing data without the need to conduct further survey effort. Unstructured datasets can be problematic when originating from remote or unevenly populated areas, as this usually results in heavily biased observer effort across the study area (Araújo and Guisan, 2006; Barry and Elith, 2006). Our data originated from a well-settled area with consistent land use and a comprehensive road network, resulting in the environmental space of the study area being sampled representatively. The fact that our "distance to road" variable was dropped from the model is evidence that bias in observer effort was not an overwhelming feature of the dataset.

The malleefowl is a species for which the collection of sightings data is appropriate as it is taxonomically stable and easily identifiable so error associated with misidentification is negligible. Our approach is not novel and is suitable for other conspicuous taxa with detailed presence data (e.g. butterflies – Lütolf et al., 2006; plants – Engler et al., 2004; mammals – Ward and Close, 2004) and shows potential for making use of the abundance of Atlas data that currently exists (Dunn and Weston, 2008). However, the reliability of absence data represents a major limitation for using the approach with less conspicuous taxa.

Our model described less than half the variation in the distribution of malleefowl sightings, suggesting that other factors not included in the model are also responsible for influencing distribution. It is possible that the use of sightings data in our model has limited its explanatory power somewhat as it is a data source that is inherently geographically variable. Factors that we were unable represent spatially may have also influenced malleefowl distribution substantially (e.g. predation by the introduced red fox).

By plotting the confidence intervals of the model spatially, we illustrated the substantial on-ground variability in model outcomes. An explicit recognition of how model uncertainty translates into on-ground spatial variability must be appreciated when applying model outcomes for management. Despite these shortcomings, we believe that this approach is useful in quantifying the relative influence of different variables on a species when traditional presence/absence data is lacking, which is important when attempting to manage a species for conservation.

4.2. *Malleefowl distribution*

Our analyses showed that malleefowl are widespread throughout the Western Australian wheatbelt and are most closely associated with mallee and shrubland vegetation assemblages. The species has managed to persist despite extensive clearing of its habitat perhaps due to the selective nature of clearing. Comparatively less mallee, shrubland and thicket associations were cleared in the wheatbelt compared with woodlands (Burvill, 1979; National Land and Water Resources Audit, 2001). Much of the vegetation remaining, particularly in the eastern wheatbelt, contains habitat identified as suitable for malleefowl.

Malleefowl are considered a sedentary species (Frith, 1962) and at risk from habitat loss and fragmentation (Benshemesh, 2000) but the scale at which these processes become significant has not been previously studied. The outcomes of this study suggest that mallee/shrubland cover in the surrounding 5 km should remain above 30% and patches of 500 ha or greater are required to increase the probability of the species occurring. These findings are not unexpected as studies have established that malleefowl may have home ranges up to 400 ha in size (Booth, 1985) and so are quite likely to respond at this scale. Also, individuals are known to move distances of 15 km or greater over a matter of weeks (Frith, 1959; Sims, 2000).

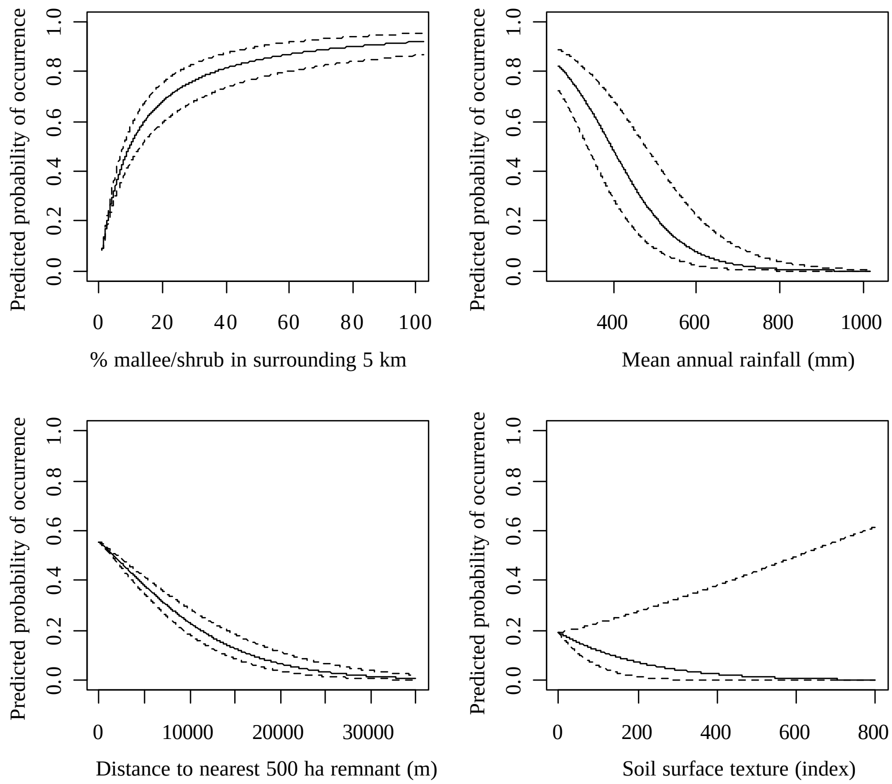

Fig. 4 – The response of malleefowl presence-absence to predictor variables in the GLM. Solid black lines = response (all other variables held at mean value); and dashed lines = 95% confidence intervals for the response.

Given the uncertainty surrounding predictions of minimum patch size (Lindenmayer and Luck, 2005), the pursuit of an area threshold above which the species can persist may not be ecologically informative (*sensu* Brooker, 2002; Radford and Bennett, 2004). Like many bird species (see Villard et al., 1999), our analysis illustrated that the relationship between malleefowl occurrence and the loss and fragmentation of vegetation does not lend itself to a specific threshold. Further, the association between presence of malleefowl and distance to nearest remnant of 500 ha (as opposed to remnants of 200, 100, 75, or 50 ha) suggests that the species simply requires large areas of habitat in the surrounding landscape in order to persist. Therefore, rather than aiming to reach a minimum threshold value for habitat cover (e.g. minimum patch size), management should seek to increase habitat cover in the context of the broader landscape (Westphal et al., 2003; Bennett et al., 2006).

Of various soil predictors, surface texture was the only one retained in the GLM model. We found that the probability of malleefowl occurrence was greater in lighter soils (i.e. less clay content), which is consistent with work conducted in eastern Australia (Frith, 1959). This is likely to be related to mound drainage, which may be important in maintaining adequate temperatures for egg incubation via fermentation. Heavier soils also tend to be very hard when dry and it is unlikely that a malleefowl would be capable of working such a soil in order to create a mound. Furthermore, heavier soils tend to support open woodlands (Burvill, 1979), a habitat which may render the species vulnerable to aerial predation. However, the surface texture variable contributed little to the model and displayed substantial uncertainty (Fig. 4); hence it is possible that its selection in the model is a result of statistical chance rather than a representation of ecological processes. Furthermore, it may be that surface texture has limited explanatory power because it displays considerable heterogeneity within each study unit (e.g. if there are dunes and swales within each area).

Our model identified a relatively strong negative correlation between the occurrence of malleefowl and mean annual rainfall. Whilst not unexpected, explaining the processes underpinning such a relationship is difficult. Vegetation associations are broadly correlated with rainfall in the Western Australian wheatbelt (Burvill, 1979) with greater amounts of non-favoured habitat (e.g. jarrah, wheatbelt woodlands) occurring in higher rainfall areas. Therefore the rainfall variable may be acting as a surrogate for vegetation type. Also, higher rainfall areas of the wheatbelt were cleared for agriculture earlier than drier areas (Jarvis, 1986), so mean annual rainfall may also be acting as a surrogate for land use history. Higher amounts of rainfall may also have a direct influence on the nesting ability of the species, with mound temperature perhaps becoming compromised after receiving greater amounts of rainfall, particularly in spring and summer when mounds are active. The correlative nature of this and other modelling studies is a shortcoming, as it is unable to draw a causal link between explanatory and response variables, particularly when dealing with broad climatic variables (Westphal et al., 2003; Austin, 2007). We suggest that hypotheses generated by our model be further investigated by, for example, monitoring of individual populations.

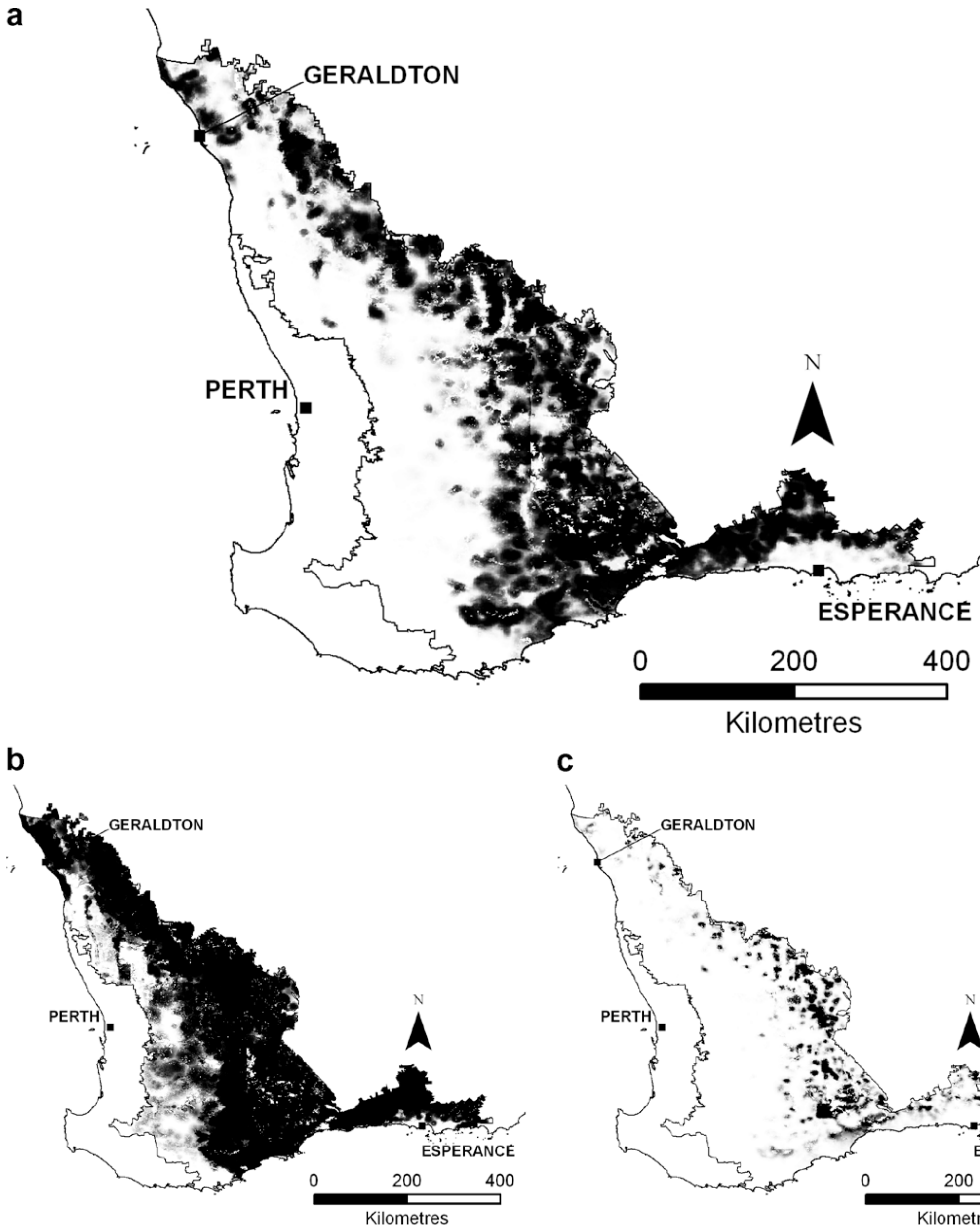

Fig. 5 – Spatial prediction plot of the final GLM and 95% confidence intervals (a) prediction at mean values; (b) upper 95% confidence interval; and (c) lower 95% confidence interval, darker indicates a higher probability of occurrence.

4.3. *Conclusion*

This study has demonstrated a method for using presence-only species data (e.g. records of sightings) to investigate the ecology of a species by combining it with absences derived from an Atlas program. Our statistical modelling approach was effective at broad scales and may have generality to any conspicuous and accessible species where there is strong community interest and an established Atlas scheme.

Our spatial modelling identified more than two million hectares of remnant vegetation within the wheatbelt as having a >50% probability of containing malleefowl. Nearly half of the vegetation identified as malleefowl habitat within the Western Australian wheatbelt exists on private land. Consequently, there is a need to incorporate the wider farming community in malleefowl related actions for conservation. This study, which used community-sourced data, provides key stakeholders (i.e. community groups, State agencies) with relevant knowledge for use in on-ground management of the species. The successful conservation and management of this species in this agricultural landscape will require concerted action across both the public and private estate.

Acknowledgements

We thank the Malleefowl Preservation Group and the North Central Malleefowl Preservation Group for providing access to their data and for their support during this study. We also thank Friends of North Eastern Malleefowl, Western

Australian Museum, and Western Australian Department of Environment and Conservation for provision of location data. Drs. Roger Lawes and Nick Nicholls (CSIRO Sustainable Ecosystems) provided advice on modelling techniques. Dr. Jeremy Wallace (CSIRO Mathematical and Information Sciences) and Dennis van Gool (Department of Agriculture and Food, Western Australia) provided vegetation and soil datasets, respectively. We thank them for their support. Drs. Mike Austin and Geoff Barrett provided comments on an earlier draft of this manuscript. This project was undertaken while BP was the recipient of an Australian Postgraduate Award. Financial support was provided by the Avon Catchment Council (via the Natural Heritage Trust), CSIRO Sustainable Ecosystems and the University of Western Australia School of Animal Biology.

REFERENCES

Araújo, M.B., Guisan, A., 2006. Five (or so) challenges for species distribution modelling. Journal of Biogeography 33, 1677–1688.

Austin, M.P., 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. Ecological Modelling 157, 101–118.

Austin, M., 2007. Species distribution models and ecological theory: a critical assessment and some possible new approaches. Ecological Modelling 200, 1–19.

Austin, M.P., Meyers, J.A., 1996. Current approaches to modelling the environmental niche of eucalypts: implication for management of forest biodiversity. Forest Ecology and Management 85, 95–106.

Barrett, G.W., Silcocks, A., Barry, S., Cunningham, R., Poulter, R., 2003. The New Atlas of Australian Birds. Birds Australia, Hawthorn East.

Barry, S., Elith, J., 2006. Error and uncertainty in habitat models. Journal of Applied Ecology 43, 413–423.

Bennett, A.F., Radford, J.Q., Haslem, A., 2006. Properties of land mosaics: implications for nature conservation in agricultural environments. Biological Conservation 133, 250–264.

Benshemesh, J., 2000. National Recovery Plan for Malleefowl. Department for Environment and Heritage, Adelaide.

Booth, D.T., 1985. Ecological Physiology of Malleefowl (*Leipoa ocellata*). Ph.D. Thesis, University of Adelaide, Adelaide.

Brooker, L., 2002. The application of focal species knowledge to landscape design in agricultural lands using the ecological neighbourhood as a template. Landscape and Urban Planning 60, 185–210.

Brotons, L., Thuiller, W., Araujo, M.B., Hirzel, A.H., 2004. Presence–absence versus presence-only modelling methods for predicting bird habitat suitability. Ecography 27, 437–448.

Burnham, K.P., Anderson, D.R., 2002. Model Selection and Multimodel Inference. A Practical Information-theoretic Approach. Springer, New York.

Burvill, G.H., 1979. Agriculture in Western Australia 1829–1979. UWA Press, Nedlands.

Busby, J.R., 1986. A biogeoclimatic analysis of *Nothofagus cunninghamii* (Hook.) Oerst. in southeastern Australia. Australian Journal of Ecology 11, 1–7.

Commonwealth of Australia. 2004. Australian Natural Resources Atlas – Native Vegetation Types and Extent. <http://www.audit.ea.gov.au/ANRA/vegetation/vegetation_frame.cfm?region_type=AUS&region_code=AUS&info=veg_type> (accessed 2006).

Crossman, A.F., 1909. Birds seen at Cumminin Station, Western Australia. Emu 9, 84–90.

Denoël, M., Lehmann, A., 2006. Multi-scale effect of landscape processes and habitat quality on newt abundance: implications for conservation. Biological Conservation 130, 495–504.

Dirnbock, T., Hobbs, R.J., Lambeck, R.J., Caccetta, P.A., 2002. Vegetation distribution in relation to topographically driven processes in southwestern Australia. Applied Vegetation Science 5, 147–158.

Dunn, A.M., Weston, M.A., 2008. A review of terrestrial bird atlases of the world and their application. Emu 108, 42–67.

Engler, R., Guisan, A., Rechsteiner, L., 2004. An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. Journal of Applied Ecology 41, 263–274.

Evans, F.H., Caccetta, P.A., 2000. Broad-scale spatial prediction of areas at risk from dryland salinity. Cartography 29, 33–40.

Ferrier, S., Watson, G., Pearce, J., Drielsma, M., 2002. Extended statistical approaches to modelling spatial pattern in biodiversity in northeast New South Wales. I. Species-level modelling. Biodiversity and Conservation 11, 2275–2307.

Fielding, A.H., Bell, J.F., 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental Conservation 24, 38–49.

Fischer, J., Lindenmayer, D.B., 2006. Beyond fragmentation: the continuum model for fauna research and conservation in human-modified landscapes. Oikos 112, 473–480.

Fortin, M., Dale, M., 2005. Spatial Analysis: A Guide for Ecologists. Cambridge University Press, Cambridge.

Frith, H.J., 1956. Breeding habits in the family Megapodiidae. Ibis 98, 620–640.

Frith, H.J., 1959. Breeding of the mallee-fowl, *Leipoa ocellata* Gould (Megapodiidae). CSIRO Wildlife Research 4, 31–60.

Frith, H.J., 1962. Conservation of the mallee fowl, *Leipoa ocellata* Gould (Megapodiidae). CSIRO Wildlife Research 7, 33–49.

Garnett, S., Crowley, G., 2000. The Action Plan for Australian Birds 2000. Environment Australia, Canberra.

Garrison, B.A., Lupo, T., 2002. Accuracy of bird range maps based on habitat maps and habitat relationship models. In: Scott, J.M. (Ed.), Predicting Species Occurrences: Issues of Accuracy and Scale. Island Press, Washington, pp. 367–375.

Geoscience Australia, 2005. Maps, Images, Data and Publications. <https://www.ga.gov.au/bridge/products.jsp> (accessed November 2005).

Graham, C.H., Ferrier, S., Huettman, F., Moritz, C., Peterson, A.T., 2004. New developments in museum-based informatics and applications in biodiversity analysis. Trends in Ecology and Evolution 19, 497–503.

Guisan, A., Edwards Jr., T.C., Hastie, T., 2002. Generalized linear and generalized additive models in studies of species distributions: setting the scene. Ecological Modelling 157, 89–100.

Guisan, A., 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.

Hastie, T.J., Tibshirani, R.J., 1990. Generalized Additive Models. Chapman and Hall, London.

Hirzel, A.H., Hausser, J., Chessel, D., Perrin, N., 2002. Ecological niche factor analysis: how to compute habitat-suitability maps without absence data? Ecology 83, 2027–2036.

Houlder, D.J., Hutchinson, M.F., Nix, H.A., McMahon, J.P., 2000. ANUCLIM User Guide, Version 5.1. Centre for Resource and Environmental Studies, Australian National University, Canberra.

Jarvis, N., 1986. Western Australia: An Atlas of Human Endeavour. Department of Lands and Surveys, Perth.

Leake, B.W., 1962. Eastern Wheatbelt Wildlife. B.W. Leake, Perth.
Lehmann, A., Overton, J.M., Austin, M.P., 2002. Regression models for spatial prediction: their role for biodiversity and conservation. Biodiversity and conservation 11, 2085–2092.
Lindenmayer, D.B., Luck, G., 2005. Synthesis: thresholds in conservation and management. Biological Conservation 124, 351–354.
Lütolf, M., Kienast, F., Guisan, A., 2006. The ghost of past species occurrence: improving species distribution models for presence-only data. Journal of Applied Ecology 43, 802–815.
McCullagh, P., Nelder, J.A., 1983. Generalised Linear Models. Chapman and Hall, London.
Milne, D.J., Fisher, A., Pavey, C.R., 2006. Models of the habitat associations and distributions of insectivorous bats of the Top End of the Northern Territory, Australia. Biological Conservation 130, 370–385.
National Land and Water Resources Audit, 2001. Australian Native Vegetation Assessment 2001. National Land and Water Resources Audit, Turner.
Parsons, B.C., Short, J.C., Roberts, J.D., 2008. Contraction in range of malleefowl (*Leipoa ocellata*) in Western Australia: a comparative assessment using presence-only and presence–absence datasets. Emu 108, 221–231.
Pearce, J.L., Boyce, M.S., 2006. Modelling distribution and abundance with presence-only data. Journal of Applied Ecology 43, 405–412.
Phillips, S.J., Anderson, R.P., Schapire, R.E., 2006. Maximum entropy modeling of species geographic distributions. Ecological Modelling 190, 231–259.
Radford, J.Q., Bennett, A.F., 2004. Thresholds in landscape parameters: occurrence of the white-browed treecreeper *Climacteris affinis* in Victoria, Australia. Biological Conservation 117, 375–391.
Saunders, D.A., Hobbs, R.J., Arnold, G.W., 1993. The Kellerberrin Project on fragmented landscapes: a review of current information. Biological Conservation 64, 185–192.
Sims, C., 2000. The reintroduction of malleefowl to Shark Bay, Western Australia. Reintroduction News 19, 38–41.
Storr, G.M., 1991. Birds of the South-west Division of Western Australia. Western Australian Museum, Perth.
Villard, M.A., Trzcinski, M.K., Merriam, G., 1999. Fragmentation effects on forest birds: relative influence of woodland cover and configuration on landscape occupancy. Conservation Biology 13, 774–783.
Ward, S., Close, R.L., 2004. Southern Sydney's urban koalas: community research and education at Campbelltown. In: Lunney, D., Burgin, S. (Eds.), Urban Wildlife: More Than Meets the Eye. Royal Zoological Society of New South Wales, Mosman, pp. 44–54.
Western Australian Department of Agriculture and Food, 2004. Soil-landscape Mapping. <http://www.agric.wa.gov.au/content/lwe/rpm/landcap/lra_soil_mapping.htm> (accessed 2006).
Westphal, M.I., Field, S.A., Tyre, A.J., Paton, D., Possingham, H.P., 2003. Effects of landscape pattern on bird species distribution in the Mt. Lofty Ranges, South Australia. Landscape Ecology 18, 413–426.
Wintle, B.A., Elith, J., Potts, J.M., 2005. Fauna habitat modelling and mapping: a review and case study in the Lower Hunter Central Coast Region of NSW. Austral Ecology 30, 719–738.
Woinarski, J.C.Z., Braithwaite, R.W., 1990. Conservation foci for Australian birds and mammals. Search 21, 65–67.
Zaniewski, A.E., Lehmann, A., Overton, J.M., 2002. Predicting species spatial distributions using presence-only data: a case study of native New Zealand ferns. Ecological Modelling 157, 261–280.