MASTERS BY RESEARCH

# An integrated approach to high-resolution modelling of a species range expansion using presence-only data
# a case study on roe deer (Capreolus capreolus)

Barton, Owain

*Award date:*
2018

*Awarding institution:*
Bangor University

# An integrated approach to high-resolution modelling of a species range expansion using presence-only data: a case study on roe deer (*Capreolus capreolus*)

**Owain Barton**

*School of Biological Sciences, Bangor University, Bangor, Wales*

*Mae'r Ysgoloriaeth Sgiliau Economi Gwybodaeth (KESS 2) yn fenter sgiliau lefel uwch Cymru gyfan a arweinir gan Brifysgol Bangor ar ran y sector AU yng Nghymru. Fe'i cyllidir yn rhannol gan raglen cydgyfeirio Cronfa Gymdeithasol Ewropeaidd (ESF) ar gyfer Gorllewin Cymru a'r Cymoedd.*
*Knowledge Economy Skills Scholarships (KESS 2) is a pan-Wales higher level skills initiative led by Bangor University on behalf of the HE sector in Wales. It is part funded by the Welsh Government's European Social Fund (ESF) convergence programme for West Wales and the Valleys.*

**Table of contents**

**1. Introduction**

**1.1** Potential impacts of expanding deer populations

**1.2** Monitoring wild deer populations

**1.3** Correlative species distribution modelling

**1.4** Types of data used in species distribution modelling

**1.5** Modelling range expansions

**1.6** The European roe deer (*Capreolus capreolus*): a case study

**1.7** Modelling the distribution of roe deer

**1.8** Modelling of roe deer in Britain

**1.9** Developing a deer management strategy for Wales

**1.10** Study aims

**2. Materials and methods**

**2.1** Study area

**2.2** Study species

**2.3** Modelling approach

**2.4** Correlative species distribution modelling

2.4.1 *Distribution data*

2.4.2 *Ecogeographic predictor variables*

2.4.3 *MaxEnt model*

2.4.4 *Model evaluation*

2.4.5 *Predicting habitat suitability in the novel area*

**2.5** Mechanistic individual-based model

2.5.1 *RangeShifter*

2.5.2 *Population dynamics*

2.5.3 *Dispersal*

2.5.4 *Model initialisation*

**3. Results**

**3.1** Correlative model

3.1.1 *Habitat suitability*

3.1.2 *Relative occurrence ratio*

**3.2** Mechanistic individual-based model

**4. Discussion**

**4.1** Modelling presence-only data at a high spatial resolution

4.1.1 *Identifying the drivers of range expansion*

4.1.2 *Assumptions of the correlative model*

4.1.3 *Model evaluation*

**4.2** Integration of a mechanistic model

4.2.1 *Capturing ecological realism*

4.2.2 *Improving the RangeShifter model*

**4.3** Model application

4.3.1 *Confidence in modelling techniques*

4.3.2 *Developing a management plan for roe deer*

4.3.3 *Developing a pan-European ungulate management strategy*

**4.4** Conclusions

**5. References**

**Acknowledgements**

I would like to offer my sincerest thanks to everyone that has contributed to the project. I would especially like to thank my lead supervisor Dr Graeme Shannon for an outstanding level of support and guidance throughout. Co-supervisors Professor John Healey and Dr Line Cordes and collaborators Dr Andrew Davies and Dr Alastair Ward have also made a significant contribution for which I am very grateful. Thanks are extended to the Deer Initiative who have co-funded and supported the project. Steve Griffiths, in particular, has been invaluable in offering a practical perspective on the challenges of deer management. Finally, I would like to thank all of those involved in the KESS 2 program, without whom the study would not have been possible.

**Summary**

**1.** Predicting the potential distribution of species in novel areas is invaluable for conservation planning and the utility of distribution modelling for conservation decision-making is well recognised. However, concern over the uncertainties and ecological relevance of techniques have limited their use by practitioners. Model functionality is constrained by the quality of available species data which imposes a trade-off between the resolution and scale that species-environment relationships can be estimated at. Distribution modelling is particularly challenging in non-equilibrium contexts, such as range expansion, due to the paucity of data appropriate for evaluation. Roe deer (*Capreolus capreolus*) in the UK are used as a case study to explore the efficacy of an integrated approach to modelling species range expansion using low-quality, presence-only data.

**2.** A total of 3,843 sightings of roe deer and eight ecogeographic predictor variables were used to produce a habitat suitability map using a MaxEnt model. Predictions of habitat suitability in a novel region (Wales) at a 100 $m^2$ resolution were made using a model developed for a populated, neighbouring region (England and Scotland). The contribution of each variable to the model was assessed using a jackknife test, while the performance of the model in each region was evaluated using 10-fold and $n-1$ cross validation as well as a novel, qualitative method based on the relative occurrence ratio (ROR). The map was integrated with a recently developed, mechanistic model (RangeShifter) to estimate the pattern of range expansion over time.

**3.** The key, fine-scale drivers of roe deer distribution in the UK were identified by the MaxEnt model with a strong association with woodland habitat. The area under the receiver-operating-curve (AUC) values from the 10-fold and $n-1$ cross validation were $0.794 \pm 0.015$ (mean $\pm$ standard deviation) and $0.803 \pm 0.208$, respectively for England and Scotland and $0.664 \pm 0.073$ and $0.672 \pm 0.243$ for Wales. The AUC and ROR results indicated that suitability in Wales may have been under-predicted by the MaxEnt model, although both methods are likely to be sensitive to sample size. The RangeShifter model described the expected pattern of range expansion across Wales. Based on established estimates of expansion rates for free-ranging populations, it is predicted that 92% of suitable habitat will be occupied within 21 to 47 years.

**4.** The model predictions of this study support the development of a proactive management strategy for roe deer in Wales. Integration of correlative and mechanistic models enabled the prediction of species distribution at a higher resolution across a larger geographic scale than is typically achievable using presence-only data. Details of ecologically-driven decisions in the modelling process are provided to promote greater confidence in techniques and encourage their application to a wide range of conservation objectives across taxa.

## 1. INTRODUCTION

Understanding species-environment relationships is integral to ecology and is essential for robust conservation management (Guisan & Zimmermann 2000; Guisan *et al.* 2013). Typically, modern conservation is founded on the preservation of specific, vulnerable or biologically valuable areas, which act as the basis for the development of regional strategies (Margules & Pressey 2000; Tittensor *et al.* 2014). Recently it has been recognised that the effectiveness of this approach is dependent on the maintenance of fundamental ecological processes such as movement, inter- and intra-specific interactions, natural disturbance regimes, climatic processes, variability in primary productivity and the formation of biophysical habitats (Bennett *et al.* 2009; Noss *et al.* 2012).

Ungulate species are often wide-ranging and act as keystone herbivores in a variety of ecosystems (West 1993; Martin *et al.* 2010; Holeski *et al.* 2016). The potential of ungulates to affect ecological processes on a large scale has prompted considerable research into the drivers of distribution patterns (e.g. Felix *et al.* 2007; Dawe, Bayne & Boutin 2014), movement behaviour (e.g. Leblond, Dussault & Ouellet 2010; Avgar *et al.* 2013) and population dynamics (e.g. Gaillard, Festa-Bianchet & Yoccoz 1998; Chitwood *et al.* 2015). Variation in the population status of ungulates is considerable. Some species, such as Barbary sheep (*Ammotragus lervia*), are rapidly expanding while others, such as hirola (*Beatragus hunter*), are at risk of extinction (Andanje & Ottichilo 1999; Cassinello, Acevedo & Hortal 2006). Advancing our understanding of how ungulates interact with their environment at an individual level is crucial for regional species conservation and the maintenance of functional ecosystems (Margules & Pressey 2000; Guisan *et al.* 2013).

### 1.1 Potential impacts of expanding deer populations

Deer (*Cervidae*) are a family of ungulates comprising 40 species distributed widely across the northern hemisphere, South America and Southeast Asia (Zhang & Zhang 2012). Ranges of many cervid species are expanding in response to land use modification, changes in hunting pressure, and the removal of natural predators (McShea, Underwood & Rappole 1997; Acevedo *et al.* 2005; Putman, Watson & Langbein 2011). Populations of deer often increase rapidly and studies have shown that browsing behaviour can have a number of

ecological and economic impacts (Putman & Moore 1998; Côté *et al.* 2004; Ward *et al.* 2008).

Deer can provide several ecosystem services such as nutrient cycling and biological control through trophic interactions (Costanza *et al.* 1998; Côté *et al.* 2004; Nishizawa *et al.* 2016). By foraging selectively, deer influence the competitive relationships between plants (Côté *et al.* 2004). The browsing of early successional plants, seedlings and saplings impedes and alters the composition of forest regeneration (Gill & Beardall 2001; Akashi, Unno & Terazawa 2011; White 2012). Foraging on dominant plants can enable the coexistence of competitively inferior species, thereby increasing species richness and ground cover (Côté *et al.* 2004; Nishizawa *et al.* 2016). Presence of deer on 'open sites' such as grasslands or meadows is generally perceived by managers as advantageous due to the suppression of encroaching scrub (Putman & Moore 1998). Browsing in nutrient-rich systems can improve the availability of nitrogen across the landscape through acceleration of nitrogen as well as carbon cycling (Côté *et al.* 2004). In addition to improving ecosystem functioning, deer can also be an economically valuable natural resource. Stalking for recreation can provide a profitable and sustainable source of high-quality game meat. In Scotland, recreational deer stalking is estimated to contribute £170 million to the economy (POST 2009).

The benefits of deer to ecosystems are frequently offset by the negative impact of their browsing on productivity (Côté *et al.* 2004). If dominant plant species are tolerant of browsing or resistant to herbivory then community structure is likely to converge towards a more simplified assemblage (Côté *et al.* 2004; Martin *et al.* 2010). Consequent reductions in tree growth and stand density reduce the economic value of forests and diminish the protection they provide from soil erosion (Mitchell & Kirby 1990; Cooke & Lakhani 1996; Putman & Moore 1998; Côté *et al.* 2004). Modifying the abundance, structure and diversity of plant species also has been shown to have several cascading effects on populations of birds, insects and other mammals (Feber *et al.* 2001; Flowerdew & Ellwood 2001; Allombert, Stockton & Martin 2005; Greenwald, Petit & Waite 2008; Bressette, Beck & Beauchamp 2012). A reduction in habitat availability, disruption to plant-pollinator associations and increased competition for resources are all impacts attributed to high deer densities, relative to the environment (Feber *et al.* 2001; Flowerdew & Ellwood 2001; Côté *et al.* 2004). Deer impacts cost the Forestry Commission in Scotland an estimated £4.5 million per

annum while agricultural damage in South East England is estimated at £3.2 million per annum (White & Smart 2004; POST 2009). Another significant socio-economic cost of deer overabundance arises from the increase in deer-vehicle collision rates (Côté *et al.* 2004).It is estimated that deer are responsible for over 30,000 road traffic accidents in Britain each year resulting in over 1,100 injuries and approximately 20 fatalities (The Deer Initiative 2007; POST 2009; Langbein 2011).

It is generally recognised that low to moderate densities of deer, relative to the environment, are economically and ecologically favourable (Weisberg *et al.* 2002; Côté *et al.* 2004). A review by Putman *et al.* (2011) introduced a decision-support framework to assist in the development of effective management strategies. Emphasis was placed on deer densities and range use assessed on a landscape scale, equivalent to the effective home-range area of the population (Putman, Watson & Langbein 2011). Historically, forestry and farm managers have operated on a local scale (Putman, Watson & Langbein 2011). However, the disassociation between such property-scale management of habitat and the appropriate landscape scale of deer management is detrimental to the operations of each, as eradication of deer from one area is likely to displace browsing pressure to a neighbouring site (Palmer *et al.* 2003; Austin *et al.* 2010; Putman, Watson & Langbein 2011; Wäber & Dolman 2015).

### 1.2 Monitoring wild deer populations

Estimating deer abundance and space use on a large scale is typically achieved using: (1) direct methods such as distance sampling (e.g. Wäber & Dolman 2015), (2) indirect methods such as pellet counts (e.g. Acevedo *et al.* 2010) and track counts (Stephens *et al.* 2006), or (3) indices of abundance such as hunting bag records (e.g. Burbaite & Csányi 2009) and the Kilometric Index (e.g. Vincent *et al.* 1995). Such methods facilitate the assessment of observed impacts in relation to local densities and can inform important management decisions such as the estimation of cull targets or the location of protective fencing (Putman, Watson & Langbein 2011). Although useful for managing impact, population monitoring alone is unlikely to be sufficient for impact prevention. Detection is often imperfect and there is considerable uncertainty in estimated density thresholds at which

impacts start to occur (Putman *et al.* 2011) . Observations of presence in a novel area could be evidence of a larger population that has already started to affect the local environment (Putman *et al.* 2011). Physical monitoring techniques are also limited geographically, due to the inherent costs and logistics of fieldwork and specialist equipment such as camera traps or thermal imaging cameras (Turner *et al.* 1995; Putman, Watson & Langbein 2011). Predictive tools, such as species distribution models, present an effective and economical approach to estimating large scale patterns of range expansion and identifying priority habitats that are most vulnerable to increased browsing pressure (Elith & Leathwick 2009).

### 1.3 Correlative species distribution modelling

Species distribution models (SDMs) are empirical models relating observations of species presence or abundance with ecogeographic variables (Guisan & Thuiller 2005; Elith & Leathwick 2009). In recent decades, SDMs have become the most commonly used tool for the spatially explicit investigation of species-environment relationships (Guisan & Thuiller 2005; Elith & Leathwick 2009; Guisan *et al.* 2013). Such research has provided important insights into species responses to climate change and human activity (Acevedo *et al.* 2011; Kuemmerle *et al.* 2014). Predictive models have been used to assess disease spread (Williams, Dechen Quinn & Porter 2014), identify habitat vulnerable to browsing (DeVore *et al.* 2016), estimate the performance of mitigation strategies (Radeloff, Pidgeon & Hostert 1999), and predict the expansion patterns of invasive species (Acevedo *et al.* 2007). The utility of distribution modelling is considerable but several review papers have cautioned that the robustness of models is contingent on the comprehensive understanding of their limitations during development (Guisan & Thuiller 2005; Heikkinen *et al.* 2006; Elith & Graham 2009; Elith & Leathwick 2009). In reality, species distributions may be influenced by a vast number of biotic and abiotic factors and even well-developed models are unlikely to encapsulate all of the complexity of a natural system, which inherently engenders uncertainties in model output (Elith & Leathwick 2009; Guillera-Arroita *et al.* 2015). Management based on SDMs should specify a level of acceptable uncertainty suitable to the objectives of the model (Guisan *et al.* 2013). Each stage of model development should be informed by: (1) the type and quality of occurrence data available, (2) the ecology of the

focal species and (3) the intended application of the output (Guisan and Thuiller 2005, Elith and Leathwick 2009).

### 1.4 Types of data used in species distribution modelling

SDMs are founded on the premise that a species' observed distribution is indicative its environmental preferences or requirements (Guisan & Thuiller 2005; Elith & Leathwick 2009). Occurrence data used in models can be described as either (i) presence-only, (ii) presence-absence or (iii) detection (Guillera-Arroita *et al.* 2015). The type of available data governs the selection, resolution and predictive ability of models, which consequently defines their usefulness as guidance tools for management decisions (Johnson & Gillingham 2005; Guisan *et al.* 2013; Guillera-Arroita *et al.* 2015).

Presence-only data are typically the most commonly available but they are generally attained through non-systematic processes such as incidental recordings by the public or from museum records (Elith & Leathwick 2009; Guillera-Arroita *et al.* 2015). Simply observing correlations between observed presences and ecogeographic variables is insufficient to describe species-environment relationships as a lack of observations in an area may be due to historical factors or a paucity of sampling and not because the environment is unsuitable (Guillera-Arroita *et al.* 2015). Consequently, popular methods to model presence-only data, such as genetic algorithm for rule set production (GARP), ecological niche factor analysis (ENFA) and maximum entropy modelling (MaxEnt), rely on the creation of pseudoabsence points to provide a background sample of ecogeographic variables (Elith & Leathwick 2009). The distribution of pseudoabsences in ecogeographic space is then compared to that of presences to identify any biases which may be interpreted as selection for, or avoidance of, specific environmental conditions (Elith & Leathwick 2009; Phillips 2009).

True absence data collected from surveys provides more reliable information about the environmental characteristics of unsuitable habitat (Elith & Leathwick 2009; Guillera-Arroita *et al.* 2015). As such, presence-absence or detection data can be modelled using regression methods including generalized linear models (GLMs), generalized linear mixed models (GLMMs) and generalized additive models (GAMs) (Guisan & Zimmermann 2000; Venables

& Dichmont 2004). Detection data collected through repeated surveys also facilitates the explicit modelling of detectability; a potential source of inaccuracy in models using presence-absence data (Mackenzie *et al.* 2002; Lahoz-Monfort, Guillera-Arroita & Wintle 2014; Guillera-Arroita *et al.* 2015).

It is imperative to understand how data types and sampling processes influence model utility (Guillera-Arroita *et al.* 2015). Ideally, occurrence data should be collected using a randomised sampling strategy to ensure that all combinations of environmental predictors are sampled with equal probability and pseudo-replication through autocorrelation is avoided (Guisan and Zimmermann 2000). Non-systematic sampling reduces model output from probability estimates of species occurrence to relative likelihood of species occurrence or to a simplified ranking of sites by occurrence probability (Guillera-Arroita *et al.* 2015). Although informationally more valuable, repeat surveys can be logistically challenging to conduct and any data obtained are likely to be more spatially and temporally constrained than a large-scale, presence-only dataset (Elith & Leathwick 2009; Guillera-Arroita *et al.* 2015).

### 1.5 Modelling range expansions

Challenges arise when predicting future species' distributions in a novel area as two key assumptions of SDMs are often invalidated (Elith & Leathwick 2009; Elith, Kearney & Phillips 2010). SDMs assume that species are at equilibrium with their environment and that all relevant environmental gradients in the study area have been equally and adequately sampled (Guisan & Thuiller 2005; Elith & Leathwick 2009; Menke *et al.* 2009; Elith, Kearney & Phillips 2010). At the time of modelling, the potential region of expansion has not been populated. Consequently, predictions of habitat suitability must be extrapolated from sampled locations outside the study area, which may be considered representative of stable populations (Miller *et al.* 2004; Dormann 2007; Elith & Leathwick 2009; Elith, Kearney & Phillips 2010). Caution is strongly advised when extrapolating predictions to new geographic regions, as space use in the novel area may be limited by environmental factors and biotic interactions that are different from those in the sampled region (Miller *et al.* 2004; Dormann 2007; Elith, Kearney & Phillips 2010).

It has been suggested that integration of correlative SDMs with mechanistic, individual-based models (IBMs) may improve the robustness of predicted range shifts in non-equilibrium contexts (With 2002; Kearney & Porter 2009; Gallien *et al.* 2010). IBMs describe population dynamics through hierarchal, rule-based simulations of each individual in a population (DeAngelis & Mooij 2005). The behaviour of simulated animals is characterised as a series of probabilities. Constraints are imposed by rules based on existing knowledge of life history traits, movement behaviour and habitat selection (DeAngelis & Mooij 2005; López-Alfaro *et al.* 2012). The spatially explicit modelling of demography and dispersal processes in IBMs facilitates investigation of the drivers of range expansion on a local scale (Gallien *et al.* 2010; Dormann *et al.* 2012; Bocedi *et al.* 2014). Population dynamics on a landscape scale are then observable as a function of the interactions between individuals and their environment (DeAngelis & Mooij 2005; López-Alfaro *et al.* 2012). However, parameterising an IBM can be challenging and requires expert knowledge of life history traits or detailed demographic data (Kearney & Porter 2009; Elith, Kearney & Phillips 2010; Gallien *et al.* 2010).

### 1.6 The European roe deer (*Capreolus capreolus*): a case study

Roe deer are the most widespread cervid species in Europe with a range that extends east to Russia and central Asia (Lorenzini *et al.* 2014; Valente *et al.* 2016). Populations are increasing in Europe and numbers are estimated to exceed 15 million individuals (Lovari *et al.* 2016). Presence of roe deer has been associated with agricultural damage (White & Smart 2004) , reduction of biodiversity (Flowerdew & Ellwood 2001) and crop yield in forests (Ward *et al.* 2004) and an increase in deer-vehicle collisions (Girardet, Conruyt-Rogeon & Foltête 2015; Hothorn *et al.* 2015). Over the last three decades, recognition of the impacts associated with roe deer presence has stimulated investigation into the ecology of the species (e.g. Gaillard *et al.* 1993; Nilsen *et al.* 2009), its population dynamics (e.g. Wahlström & Liberg 1995; Focardi *et al.* 2002) and methods of monitoring its populations (e.g. Putman, Watson & Langbein 2011; Valente *et al.* 2016b).

It is estimated that roe deer have been a component of the British fauna for the past 10,000 years (Baker & Rus Hoelzel 2013). Changes to forest laws in the fourteenth century severely

reduced populations through increased hunting pressure and deforestation (Hewison 1995). By the sixteenth century roe deer were extirpated throughout southern Britain with only a few remnant populations remaining in the Scottish highlands (Hewison 1995; Baker & Rus Hoelzel 2013). Reintroductions to East Anglia and the south of England during the nineteenth century, together with afforestation have since contributed to their recovery (Hewison 1995; Ward 2005). Further recent changes in habitat structure appear to have accelerated population growth and, for the last four decades, roe deer have expanded significantly both in abundance and geographic distribution (Hewison 1995; Putman & Moore 1998; Ward 2005). Impacts on agriculture and forestry have been observed and roe deer contribute approximately 32% and 69% of all deer-vehicle collisions in England and Scotland, respectively (White & Smart 2004; The Deer Initiative 2007; Ward *et al.* 2008). The national population size is estimated to exceed 800,000 individuals with an annual growth rate of approximately 2.3% per year (Ward 2005; POST 2009). Roe deer are now the most widely distributed deer species in Britain and remain either absent or present at low densities only in the Midlands and parts of Kent in England and in Wales (Ward 2005; Palmer 2014). Current distributions are consistent with a history of radial expansion from sites of reintroduction (Hewison 1995; Ward 2005; Baker & Rus Hoelzel 2013). As ranges continue to expand, these areas are, therefore, most likely to be the locations of future recolonization (Ward 2005; Palmer 2014).

### 1.7 Modelling the distribution of roe deer

Some behavioural and biological characteristics of roe deer make them a favourable study species for correlative population modelling. Roe deer are income breeders, they store limited body reserves and acquire energy for reproduction during the reproductive period (Andersen *et al.* 2000). Therefore, distributions are likely to be strongly influenced by the availability of resources such as cover and palatable forage (Andersen *et al.* 2000; Toigo *et al.* 2006; Nilsen *et al.* 2009; Tablado *et al.* 2016). For highly mobile species, resource selection can be difficult to quantify if the available occurrence data do not describe the behavioural mode of the animal at the time of observation (Johnson 1980; Pearce & Boyce 2006; Elith & Leathwick 2009). It is not possible to determine if the presence of deer at observed loactions are due to their selection of specific resources present there or if the

animal is simply moving through the landscape (Johnson 1980; Pearce & Boyce 2006; Elith & Leathwick 2009). However, territoriality in male roe deer and high home range fidelity observed in adults of both sexes, indicate that fine-scale temporal changes in habitat use are unlikely to occur over a large geographic area (Wahlström & Liberg 1995; Kjellander *et al.* 2004; Debeffe *et al.* 2013). Seasonal migrations between winter and summer ranges have been observed but these are typically only for juveniles and yearlings (Wahlström & Kjellander 1995; Wahlström & Liberg 1995). In most cases, permanent settlement occurs once individuals reach three years of age (Wahlström & Kjellander 1995; Wahlström & Liberg 1995). Home ranges of roe deer are relatively small, varying in size from 10 to 140 ha (Vincent *et al.* 1995; Wahlström & Liberg 1995; Kjellander *et al.* 2004; Saïd *et al.* 2005; Morellet *et al.* 2013). Provided that the occurrence data are recorded to a resolution that is the same as, or finer than, this home range area, characterisation of suitable habitat based on correlative models should be robust (Ferrier *et al.* 2002; Elith & Leathwick 2009).

Extensive research on roe deer has also provided a wealth of information that can be used to parameterise an individual-based model. Data are available for key aspects of their ecology,behaviour and population dynamics including age structure and survival rates (Gaillard *et al.* 1993; Loison *et al.* 1999; Davis, Stephens & Kjellander 2016; Sonnichsen *et al.* 2017), fecundity (Hewison & Gaillard 1996, 2001; Wäber, Spencer & Dolman 2013a), dispersal characteristics (Wahlström & Liberg 1995; Pettorelli *et al.* 2003; Gaillard *et al.* 2008; Debeffe *et al.* 2013) and their density dependence (Putman *et al.* 1996; Andersen & Linnell 2000; Cobben *et al.* 2009).

A systematic review of studies, from 1996 to 2017, that applied species distribution modelling to roe deer identified 23 publications and one unpublished thesis (see Appendix 1). A strong bias was observed towards the use of regression modelling (n = 16) over resource and step selection functions (n = 4), machine learning methods (n = 3) and IBMs (1). A relationship was observed between the type of occurrence data used in each model, the resolution of the occurrence data and the extent of the study area (Table 1). Studies that modelled occurrence data at higher resolutions typically covered smaller areas and relied on data collected through surveys or tracking using GPS or VHF radio collars. Occurrence data at a resolution ≤0.01 $km^2$ were modelled on a maximum area of 2,058 $km^2$

(Wu, Li & Hu 2016). These results suggest that species distribution modelling of roe deer is limited by a trade-off between study area and the quality of occurrence data available.

Table 1. Data sources used in studies of roe deer (*Capreolus capreolus*) species distribution modelling (n=24, 1996-2017, Appendix 1) included in a systematic review. DET=detection, PA=presence-absence, PO=presence-only, MOV=movement data obtained from GPS or VHF radio collars (note: details of one study; Torres *et. al.* 2012, were not available).

| Resolution of species occurrence data ($km^2$) | Number of studies | Mean study area ($km^2$) | Types of occurrence data (number of studies) |
|---|---|---|---|
| ≤0.01 | 11 | 373 | DET (1); PA (4); MOV (6) |
| 0.01 – 1 | 6 | 17,300 | PA (2); MOV (1); PO (3) |
| 1 – 100 | 5 | 153,200 | PA (1); PO (4) |
| >100 | 1 | 311,900 | PO (1) |

**1.8 Modelling of roe deer in Britain**

Populations of roe deer in the UK have been modelled on a national scale in three studies (Acevedo *et al.* 2010b; Rodríguez-Rey, Jiménez-Valverde & Acevedo 2013; Palmer 2014). In each study, habitat suitability was estimated across a 10 x 10 km UTM grid using data obtained from surveys conducted by the British Deer Society, National Biodiversity Network (NBN) Gateway (since renamed NBN Atlas, https://nbnatlas.org/) and published sources (Ward 2005). In each case it was assumed that lack of presence within a grid cell equated to absence. Although the majority of occurrence data used were collected through non-systematic processes, Acevedo *et al.* (2010) argue cogently that such an assumption is appropriate given the considerable sampling effort and the large scale at which the data are modelled (Acevedo *et al.* 2010b). The study by Palmer (2014) developed an additional mechanistic model and presented the first example of an integration between correlative and mechanistic models for the prediction of deer distributions (Palmer 2014).

An individual-based model was developed using demographic data taken from published literature and predictions of habitat suitability produced by the correlative model (Palmer 2014). Populations were simulated using historic distribution data and model performance

was assessed based on the ability of the model to predict current and future distributions (Palmer 2014). Although the study demonstrated the utility of an integrated modelling approach, predictive performance was low for novel areas such as Wales and the English Midlands (Palmer 2014). Under-prediction by the IBM was most likely due to inaccuracies in the prediction of habitat suitability by the correlative model (Palmer 2014). As for the other two national studies, the UK was modelled as a whole with the implicit assumption that the species was in equilibrium with its environment. Lack of presences in Wales and the English Midlands consequently produced models that predicted low habitat suitability for these regions (Palmer 2014). It is also likely that the consideration of ecogeographic variables on a 10 x 10 km scale overlooked important, fine-scale details of the landscape which may drive the patterns of observed distributions (Ferrier *et al.* 2002; Elith & Leathwick 2009).

### 1.9 Developing a deer management strategy for Wales

In 2011, the Welsh Government responded to the emerging threat of growing deer numbers and introduced the Wild Deer Management strategy (Welsh Government 2017). The recently updated five-year action plan, in support of the strategy, was developed by the Wales Deer Forum, a collection of 19 member organisations including public bodies, non-governmental organisations and private landowners (Welsh Government 2017). The action plan describes the wide range of potential impacts of deer on agriculture, and the objectives set out in several existing government policies and strategies such as The Environment (Wales) Act 2016, Woodlands for Wales (2009) and the Nature Recovery Plan for Wales 2015 (Welsh Government 2017). Evidence-based policy is listed as a guiding principle for future management of deer (Welsh Government 2017). Population control of deer is achieved through lethal action and it is recognised in the strategy that successful implication of such measures requires robust techniques for population monitoring and forecasting (Wäber, Spencer & Dolman 2013a; Welsh Government 2017).

### 1.10 Study aims

The intention of the Wales Deer Forum is to take a proactive approach to deer management in Wales by monitoring and controlling deer populations before they reach unfavourably high densities (Welsh Government 2017). The purpose of the present study is to support the objectives of the Wales Deer Forum by using species distribution modelling techniques to: (1) predict the distribution of suitable roe deer habitats across Wales and (2) estimate the pattern of future range expansion, at a scale that is appropriate to inform population management.

As the ecological implications of global climate change and anthropogenic disturbance have been recognised, species distribution modelling has become an increasingly popular tool for ecosystem conservation (Elith & Leathwick 2009; Guisan *et al.* 2013). The reliability and geographic range of models is often limited by the availability of high quality occurrence data (Pearce & Boyce 2006). Models developed using low resolution data have provided useful insights into the broad patterns of distributions at a national scale (Guillera-Arroita *et al.* 2015). Studies using higher resolution data typically cover much smaller areas but offer more accurate assessments of the species-environment relationship (Boyce *et al.* 2002; Guillera-Arroita *et al.* 2015). The resolution of models is usually constrained by logistical difficulties in acquiring occurrence data through systematic processes over a large area (Elith & Leathwick 2009; Guillera-Arroita *et al.* 2015). This study aims to build on the coarse, 10 x 10 km predictions of existing roe deer models and demonstrate a robust, integrated approach to modelling a species range expansion across a large area using presence-only data. Improving methods for modelling low quality data should help to improve our understanding of resource use and habitat suitability for a wide range of species across multiple geographic scales.

## 2. MATERIALS AND METHODS

### 2.1 Study area

The study area covered mainland Great Britain (218,819 km$^2$), divided into two regions: England and Scotland (198,569 km$^2$) and Wales (20,250 km$^2$) (Fig. 1). The most abundant habitat in mainland Britain is improved and semi-natural grassland (39.2%) followed by arable land (25.7%), broadleaved and coniferous woodland (13.2%), upland, heath and bog (12.8%) and urban areas (7.4%; Centre for Ecology and Hydrology (CEH) Land Cover Map, 2015). The habitats are fragmented by approximately 77,000 km of roads and mountainous regions, which are present in Scotland, the North of England and North Wales (Ordnance Survey (OS) Meridian $^{TM}$ 2, OS Terrain 50 Digital Terrain Map). The climate is temperate with gradients of increasing temperatures towards the south and increasing precipitation towards the northwest.

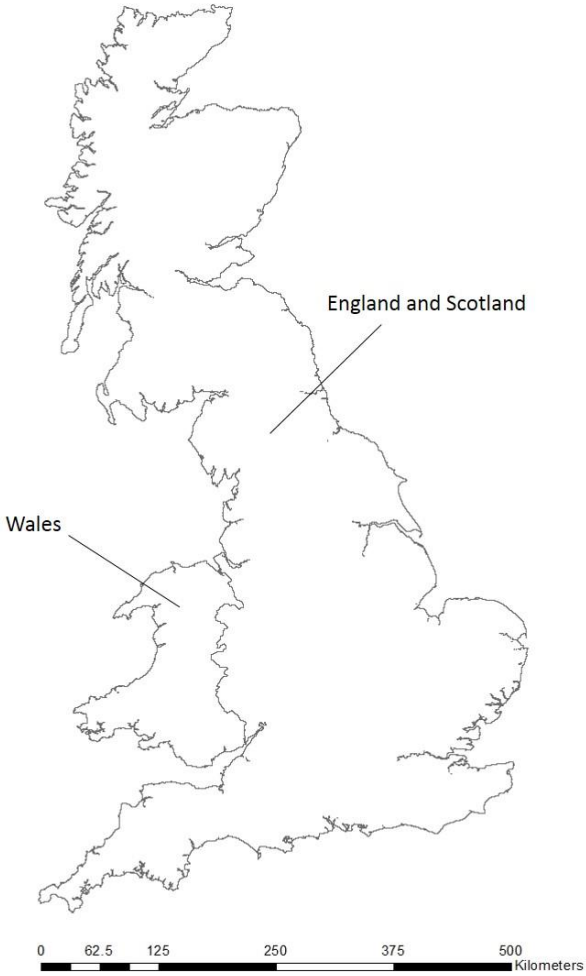


Figure 1. Map of the study area showing the boundaries of the two regions defined for the analyses.

## 2.2 Study species

Roe deer are medium sized forest-living cervids. During the breeding season (March-April to late August-early September) males defend territories and typically mate with up to three different females (Vanpé *et al.* 2008). Males become sexually mature after one year but territory defence does not usually occur before three years of age (Gaillard *et al.* 1993; Vincent *et al.* 1995). Females give birth, usually to twins, in May-June (Hewison & Gaillard 2001). They are not territorial and live solitarily or with dependent fawns in overlapping home ranges (Hewison & Gaillard 1996). Small groups of less than five females may form during winter, occasionally accompanied by one or two males (Kjellander *et al.* 2004). In open areas, such as agricultural plains, groups of over 50 individuals have been observed (Gerard *et al.* 1995). In the absence of woodland cover, individuals in larger groups most likely benefit from higher group vigilance (Ruckstuhl & Festa-Bianchet 2001; Jepsen & Topping 2004). Behavioural plasticity in response to landscape structure, high rates of adult survival and fecundity, and the capacity to tolerate a wide range of climatic conditions enable roe deer to live in a variety of fragmented landscapes (Gaillard *et al.* 1993; Jepsen *et al.* 2005; Torres *et al.* 2012a; Wäber, Spencer & Dolman 2013b).

## 2.3 Modelling approach

This study describes the prediction of roe deer range expansion in Wales through an integration of two modelling components: (1) a correlative species distribution model and (2) a mechanistic, individual-based model. A MaxEnt model was used to produce a habitat suitability map which provided a landscape for the prediction of expansion patterns using an individual-based model, RangeShifter (Bocedi *et al.* 2014).

## 2.4 (1) Correlative species distribution model

### 2.4.1 *Distribution data*

Data on roe deer sightings from 1960-2015 were taken from the NBN Gateway during December 2016. As it was critical to predict suitability at a high resolution, only occurrence records with a locational precision of 100 m (n = 5,652) were considered for analysis.

Records in Wales were limited (n=20), so data collected by regional wildlife trusts were obtained from the Biodiversity Information Service (http://www.bis.org.uk/, n = 122), the Wildlife Trust of South West Wales (https://www.welshwildlife.org/, n = 4) and the North Wales Environmental Information Service (http://www.cofnod.org.uk/Home, n = 31). All records used for analyses were characteristic of presence-only data as they were collected from a wide range of sources and sampling effort was not determinable. Data were filtered to remove duplicates (n = 1,172), records at locations with incomplete environmental data (n = 89) and any records within 100 m of a road (n = 599), as it is likely that these originated from deer-vehicle collisions (see Appendix 2a).

It was assumed from the literature (Ward 2005; Palmer 2014) and the distribution of the data that presences in Wales were indicative of the early stages of recolonization by a population expanding from England. Roe deer in Wales were, therefore, considered to be in disequilibrium with their environment. Records from Wales (n = 106) were used for evaluation of model performance in the Wales region only. Predictions of habitat suitability in Wales were derived from projections of a model trained and tested using environmental and species data (n = 3,843) from the England and Scotland region.

### 2.4.2 *Ecogeographic predictor variables*

To improve the robustness and realism of models, studies have argued for the selection of functionally relevant predictors based on existing knowledge and theory (MacNally 2000, Elith and Leathwick 2009). Ecogeographic variables were grouped into four categories: land cover, roads, terrain and climate. Land cover variables were derived from the Centre for Ecology and Hydrology's Land Cover Map 2015 (https://www.ceh.ac.uk/services/land-cover-map-2015). Information is presented for 23 habitat types which were grouped into 10 aggregate classes (Table 2). Data for roads and terrain were taken from OS Meridian™ 2 and Terrain 50 maps respectively (http://digimap.edina.ac.uk/datadownload/osdownload) and climate data were downloaded from an online database, Worldclim (http://www.worldclim.org/). All variables were resampled to 100 m pixel rasters to be consistent with the roe deer distribution data (see Appendix 2a). As home ranges of roe deer are likely to be larger than 100 x 100 m, buffer zones of 500 m radius were defined for each cell to include environmental information of the area which may feasibly have been

occupied at the time of recording. Collinearity between variables was assessed prior to model building through Variance inflation factor (VIF) analysis and the calculation of pairwise Pearson's correlation coefficients in R (https://www.r-project.org/). Variable selection was also informed based on the contribution and visual assessment of response curves derived from trial models (see Appendix 2b). If two variables were highly correlated ($r_p \geq 0.6$ or VIF > 2) then the lowest contributing variable was removed. A final set of eight variables was selected for model development from a candidate list of 23 (Table 2, see Appendix 2b). All values for predictor variables in the Wales region were within the limits of the England and Scotland region.

Table 2. Ecogeographic predictor variables used in the correlative species distribution model of roe deer distribution in England and Scotland

| Predictor variable | Description | Units |
|---|---|---|
| Land cover | Habitat type | 10 aggregated classes; Broadleaved woodland; Coniferous woodland; Arable and horticulture; Improved grassland; Semi-natural grassland; Mountain, heath & bog; Saltwater; Freshwater; Coastal; Urban and suburban |
| Woodland cover | Proportion of broadleaved and coniferous woodland within a 500-m radius buffer | 0-1 |
| Forage cover | Proportion of non-woodland forage (Improved grassland, Semi-natural grassland and Arable land) within a 500-m radius buffer | 0-1 |
| Woodland distance | Euclidean distance to broadleaved or coniferous woodland | m |
| Road cover | Area of motorways, A roads and B roads within a 500-m radius buffer | $m^2$ |
| Altitude | Height above sea level | m |
| Temperature annual range | Maximum temperature of warmest month – Minimum temperature of coldest month | °C x 10 |
| Precipitation seasonality | Coefficient of variation† | mm |

†Calculated as the ratio of the standard deviation to the mean of annual precipitation ranges for each year from 1960 to 1990

### 2.4.3 *MaxEnt model*

Presence records of roe deer in England and Scotland were modelled using MaxEnt 3.4.0 (Phillips *et al.* 2017). MaxEnt is a machine learning technique that estimates a probability distribution by comparing the distribution of presences in ecogeographic space to a user-

defined number of randomly generated background points (Phillips, Anderson & Schapire 2006). It has been shown that MaxEnt performs as well as or better than other methods when modelling presence-only data (Elith *et al.* 2006). MaxEnt is a versatile tool, applicable across all scales and can be analysed mathematically in the same way as regression models (Elith *et al.* 2006; Phillips, Anderson & Schapire 2006).

Various types of output can be produced using MaxEnt, all of which, give a value ranging from 0 to 1 for each cell. The 'raw' output is an estimation of actual or relative rate of occurrence, if total population size is known or unknown, respectively (Merow, Smith & Silander 2013). A logistic (log) or complementary log-log (cloglog) transformation of this output is also provided which can be interpreted as a probability of presence (Merow, Smith & Silander 2013; Phillips *et al.* 2017). The reliability of inferences from MaxEnt depends on important assumptions of sampling effort and the probability of species detection (Royle *et al.* 2012; Merow, Smith & Silander 2013). In this study, the cloglog transformation was used because the output more closely matched the observed distribution of the species than predictions made using the raw output. However, as the sampling process was non-systematic, this output was interpreted only as an index of habitat suitability, not a probability of presence.

A MaxEnt model was constructed using 10,000 randomly generated background points and presence data divided into a training dataset (90% of presences, n = 3,459) and test dataset (10% of presences, n = 384). Ten replicates were run, all model parameters were set to default values and the relative contribution of each ecogeographic variable to the suitability distribution was estimated. The relationship between each variable and the predicted index of habitat suitability was also assessed using a jackknife test (Phillips, Anderson & Schapire 2006). Variable importance was calculated by comparing the increase in regularised training gain when the variable was used in isolation to the decrease in gain when the variable was excluded from the full model. Environmental relationships estimated by the model for the England and Scotland region were used to predict habitat suitability in the novel region of Wales.

### 2.4.4 *Model evaluation*

Model evaluation was achieved using 10-fold and $n - 1$ cross validation techniques. Performance was estimated based on the ability to correctly rank presences in the test dataset higher than background points as given by the mean area under the receiver-operating-characteristic curve (AUC). For the England and Scotland region an AUC value was observed for 10 replicate models. Performance in Wales was estimated based on projections from the model developed for the England and Scotland region. Predictive ability was evaluated using presence records (n = 106) partitioned into 10 test datasets and 1,000 randomly generated background points. Secondary analysis was also performed for a single model in each region using an $n - 1$ cross validation method; a variant of *k*-fold cross validation which partitions one data point as a test dataset for n (England and Scotland: n = 3,843; Wales: n = 106) iterations (Cawley & Talbot 2003). Model performance was based on the mean of the AUC scores produced by each iteration.

### 2.4.5 *Modification of the habitat suitability map*

Following satisfactory evaluation, the habitat suitability map was used as the landscape for the individual-based model. Carrying capacity in the individual-based model is linearly related to habitat suitability. However, it was unreasonable to interpret the output from the MaxEnt model as a measure of actual rate of occurrence given that the total population size of roe deer in either study region was indeterminable (Merow, Smith & Silander 2013). In order to quantify the relationship between the relative rate of occurrence and habitat suitability in Wales, a novel metric was used, described here as the relative occurrence ratio (ROR). The output from the MaxEnt model was classified into five suitability categories; Very low (0-0.2), Low (0.2-0.4), Neutral (0.4-0.6), High (0.6-0.8) and Very high (0.8-1). For each suitability category, the ROR was calculated using the following equation:

$$\frac{proportion\ of\ observed\ presences\ (\%)}{proportion\ of\ study\ region\ (\%)}$$ equation 1.

Values of the habitat suitability map were scaled according to the relationship observed between habitat suitability and ROR (see Appendix 3).

Observing the ROR also facilitated an additional comparison of the predictive performance between the two regions. Assuming that the model performed well, it was predicted that the representation of presences would be disproportionately high ($> 1$) on very highly suitable and highly suitable habitat, commensurate with area ($\cong 1$) for neutral habitat and disproportionately low ($< 1$) on unsuitable and very unsuitable habitat. AUC scores are typically lower if areas predicted to be highly suitable are poorly represented by the distribution data (Jiménez-Valverde 2012). It was predicted that the AUC scores for Wales would we be lower than those obtained for England and Scotland as it is a novel region and it was likely that a large proportion of the area would be unoccupied due to factors other than habitat suitability. Observing the relationship between the ROR and habitat suitability provides an intuitive method for describing and understanding differences in predictive performance between the two regions.

### 2.5 (2) Mechanistic individual-based model

#### 2.5.1 *RangeShifter*

The expansion of roe deer across Wales was modelled using a recently developed spatially explicit individual-based modelling platform, RangeShifter v1.1 (Bocedi *et al.* 2014). Spatial dynamics in RangeShifter are simulated through an integration of population dynamics and dispersal, which is modelled in three stages: emigration, transfer and settlement (Bocedi *et al.* 2014). Emergent patterns of colonization are influenced by landscape characteristics and a set of behavioural rules defined by the user (Bocedi *et al.* 2014).

Demographic parameters of the model were mostly estimated from long-term studies by Gaillard *et al.* (1993), Wäber *et al.* (2013), Hewison (1996) and Wahlström and Kjellander (1995) (Table 3, see Appendix 4). Simulations in RangeShifter are based on landscapes units which can be individual cells, or aggregations of cells, defined as patches (Bocedi *et al.* 2014). The continuous surfaces used in cell-based models more accurately simulate the

environmental gradients that influence spatial dynamics (Brudvig *et al.* 2017). However, the units of the landscape are finer than the scale at which density dependency is likely to act (Bocedi *et al.* 2014). Models that use discrete patches of suitable habitat as landscape units are simpler and better at simulating population dynamics but omit important detail of the surrounding matrix (Brudvig *et al.* 2017). In this study, a hybrid approach was adopted. All cells of the habitat suitability map were aggregated into units to create a landscape of contiguous patches. Voronoi polygons were used to divide the cell grid of the suitability map into contiguous patches from generated points. A total of 20,000 points were created to produce patches of an ecologically relevant size (mean = 0.58 $km^2$), representative of the home range of an individual (see Appendix 5).

There is evidence to suggest that home range size varies in response to habitat suitability (Kjellander *et al.* 2004). The points defining the centre of each patch were distributed according to the probability distribution described by the habitat suitability map. The density of points was higher on more suitable habitat and so the resultant patches were smaller than less suitable patches (see Appendix 5) (Tufto, Andersen & Linnell 1996). Patches with a predicted suitability of ≤ 0.2 were removed as they were considered very unlikely to be occupied. Habitat suitability values did not change for the duration of each simulation and were linearly related to carrying capacity.

Table 3. Summary of parameters used in the mechanistic model estimated from published sources (see Appendix 4). (DD) indicates that a parameter is constrained by density dependence, * that juvenile mortality was accounted for by reducing fecundity as recommended in the RangeShifter user manual (Bocedi *et al.* 2014, RangeShifter user manual, see Appendix 4) and † that the study region was divided longitudinally into four sectors. Only suitable patches within the easternmost sector (closest to the Wales-England border) were initially populated.

| Model parameter | Symbol | Estimate |
|---|---|---|
| Density dependence coefficient $1/b$ (individuals per ha) | | 0.2 |
| Age classes (minimum age) | | Juveniles (0)<br>Yearlings (1)<br>Adults (2)<br>Senescents (8) |
| Maximum age | | 14 |
| Probability of reproduction | | 0.98 |
| *Mean fecundity* | | |
| Adults (DD) | $\phi_2$ | 0.64 (DD) |
| Senescents (DD) | $\phi_3$ | 0.43 (DD) |
| *Survival rates* | | |
| Juveniles (DD) | $\sigma_0$ | 1* (DD) |
| Yearlings | $\sigma_1$ | 0.77 |
| Adults | $\sigma_2$ | 0.96 |
| Senescents | $\sigma_3$ | 0.73 |
| Emigration probability | $D_0$ | 1 |
| *Movement parameters* | | |
| Perceptual range | PR | 40 cells (4 km) |
| Directional persistence | DP | 5 |
| *Settlement probability* | | |
| α | $P_s$ | -100 |
| β | | 1 |
| Maximum no. of steps | | 200 (20 km) |
| Initial density (individuals per ha) in all suitable eastern† patches | | 0.025 |
| *Proportion of individuals in each age class at initialisation* | | |
| Yearlings | | 0.2 |
| Adults | | 0.7 |
| Senescents | | 0.1 |

A stage-structured, female-only model was implemented. This is justified because there is little variation in life history between sexes (Gaillard *et al.* 1993, 2008; Coulon *et al.* 2006; Van Moorter *et al.* 2008). Dispersal characteristics are also similar, although the size of female home ranges is typically less variable than males which have been shown to vary in response to local population density (Vincent *et al.* 1995; Putman *et al.* 1996; Kjellander *et al.* 2004; Saïd *et al.* 2005). Currently, RangeShifter does not have the capacity to simulate density dependence for each sex independently which can lead to inaccurate simulations when both sexes are modelled. For example, the simulated probability of a male emigrating may be higher for a patch that is unpopulated by males but has a high density of females. Considering only one sex was considered a more reasonable approach and also improved the functionality of the model, which was necessary given the large study area and fine scale individuals were simulated at. It was assumed that mate availability was constant and the sex ratio for the entire population was 1:1. Four stage classes were considered; juveniles (<1 year), yearlings (1-2 years), adults (2-7 years) and senescents (>7 years). Survival, fecundity and development were modelled using the following transition matrices:

$$\text{(a)} \quad \begin{bmatrix} 0 & \phi_1 & \phi_2 & \phi_3 \\ 1.0 & \sigma_1(1-\gamma) & 0 & 0 \\ 0 & \sigma_1\gamma & \sigma_2(1-\gamma) & 0 \\ 0 & 0 & \sigma_2\gamma & \sigma_3\gamma \end{bmatrix}$$

$$\text{(b)} \quad \begin{bmatrix} 0 & 0 & 0.64 & 0.43 \\ 1.0 & 0 & 0 & 0 \\ 0 & 0.77 & 0 & 0 \\ 0 & 0 & 0.96 & 0.73 \end{bmatrix}$$

where $\phi$ is fecundity given as the number of female offspring per female, $\sigma$ is the survival probability at each stage and $\gamma$ is the probability of development (Bocedi *et al.* 2014, RangeShifter user manual Appendix 4). Juvenile survival was accounted for by reducing fecundity and it was assumed that all surviving individuals of sufficient age developed to the next stage. Adults reproduced every year with a probability of 0.98 (Hewison 1996; Hewison & Gaillard 2001).

### 2.5.2 *Population dynamics*

Density dependence is modelled in RangeShifter as an exponential decay:

$$x_i = x_{i,0} * e^{-bN_t} \qquad \text{equation 2.}$$

where $x_i$ is a parameter for survival probability, fecundity or development, $x_{i,0}$ is the maximum value of the parameter at low densities, $b$ is the strength of density dependence and $N_t$ is the total number of individuals in the local population at time $t$ (derived from Bocedi *et al.* 2014, RangeShifter user manual, Appendix 4). The parameter $1/b$ in RangeShifter is analogous to carrying capacity and was estimated at 0.2 individuals $ha^{-1}$ such that emergent populations reflected observed free-ranging densities (Wahlström & Liberg 1995; Andersen & Linnell 2000; Focardi *et al.* 2002; Cobben *et al.* 2009; Wäber & Dolman 2015).

Fecundity and juvenile survival probability were constrained by density dependence. Several studies have documented reduced fecundity in response to increasing local population densities, most likely mediated through resource availability and the subsequent effects on body mass (Gaillard *et al.* 1993; Vincent *et al.* 1995; Hewison & Gaillard 1996; Putman *et al.* 1996). Density dependency in juvenile survival was implemented to reflect observed patterns of resource-dependent emigration (Wahlström & Liberg 1995; Pettorelli *et al.* 2003; Gaillard *et al.* 2008; Debeffe *et al.* 2013). When local population densities are high, and resources are limited, fewer juveniles attain the necessary body condition to emigrate as yearlings (Wahlström & Liberg 1995). Mediating emigration through juvenile survival instead of emigration probability reduced the modelling of non-dispersing individuals that were unlikely to contribute to the pattern of range expansion.

Accounting for density effects improves the ecological realism of the model but predictions are limited by a trade-off between spatial and temporal accuracy. If habitat patches in the model are small, local population densities may approach carrying capacity following occupation by only a few individuals. The model is likely to under-predict the number of surviving juveniles each year which will substantially lower the predicted rate of expansion.

Increasing the size of patches could resolve this issue but would reduce the spatial resolution of the output. As spatial accuracy was essential to this study, each patch was constrained to an ecologically relevant size representative of the home range of an individual. The temporal accuracy of the model was consequently reduced, so the time scale for the predicted expansion was inferred using estimates of observed expansion rates from 1972-2002 (Ward 2005) and 2002-2007 (Palmer 2014).

### 2.5.3 *Dispersal*

Emigration was modelled as the probability that an individual will leave its natal patch. Only yearlings could emigrate and it was assumed that all individuals that survived as juveniles attained sufficient fitness to disperse. Movement during the transfer phase was modelled using an embedded stochastic movement simulator (SMS; Palmer, Coulon & Travis 2011). The SMS models individual movements as a series of discrete nearest-neighbour steps across a cost surface similar to a Least Cost Path approach (Palmer, Coulon & Travis 2011). Decisions are influenced by the perceptual range of the individual and their tendency to follow a correlated random walk, defined as directional persistence (Palmer, Coulon & Travis 2011; Bocedi *et al.* 2014, RangeShifter user manual Appendix 4). The inverse of the habitat suitability map was used as a 100 m resolution cost surface. Perceptual range was assumed to be 4 km, which corresponds to the mean distance travelled during explorative pre-dispersal movements (Debeffe *et al.* 2013). Although directional persistence has not been formally quantified, Debeffe *et al.* (2013) has shown that the direction of dispersal movements is similar to that of pre-dispersal exploration. Consequently, a value of five was used for directional persistence following visual assessment of trial simulations (see Appendix 4). Dispersers could move a maximum number of 200 steps which equates to a Euclidean distance of 20 km (Wahlström & Liberg 1995; Debeffe *et al.* 2013). Individuals were assumed to settle in the first non-natal patch encountered, provided that the local population density was below carrying capacity. Failure to reach a suitable patch after the maximum number of steps resulted in death.

### 2.5.4 *Model initialisation*

The study region (Wales) was divided longitudinally into four sectors of equal width. Suitable patches within the easternmost sector, closest to the Wales-England border, were initialized at a density of 0.025 individuals $ha^{-1}$. Population dynamics were simulated for 300 years for five replicates. The mean number of occupied patches, total area of occupied patches and probability of presence for each patch was estimated at six, 50-year time steps.

## 3. RESULTS

### 3.1 (1) Correlative model

#### 3.1.1 *Habitat suitability*

The area under the receiver-operating-characteristic curve (AUC) values of $0.794 \pm 0.015$ (mean $\pm$ standard deviation) for England and Scotland and $0.664 \pm 0.073$ for Wales from 10-fold cross validation and $0.803 \pm 0.208$ for England and Scotland and $0.672 \pm 0.243$ for Wales from the $n - 1$ cross validation indicate that the correlative model predicted suitability better than chance (AUC=0.5) for both regions.

Three variables: woodland distance, woodland cover and land cover, achieved the highest regularised training gain when isolated in the jackknife test (Fig. 2.) and had a combined relative contribution of 56.0% to the full model (Table 4). Altitude, forage cover and precipitation seasonality attained moderate gain (Fig. 2) and collectively contributed 37.4% (Table 4). Little gain was achieved using Temperature range and Road cover (Fig. 2) which had a combined relative contribution of 6.6% (Table 4). Results from omitting each variable showed that precipitation seasonality and altitude contained the most information not contained in the other variables (Fig. 2). Predicted suitability increased with increasing woodland cover and decreased with increasing woodland distance, altitude and road cover. Relationships between all other variables and predicted suitability were non-linear (see Appendix 6).

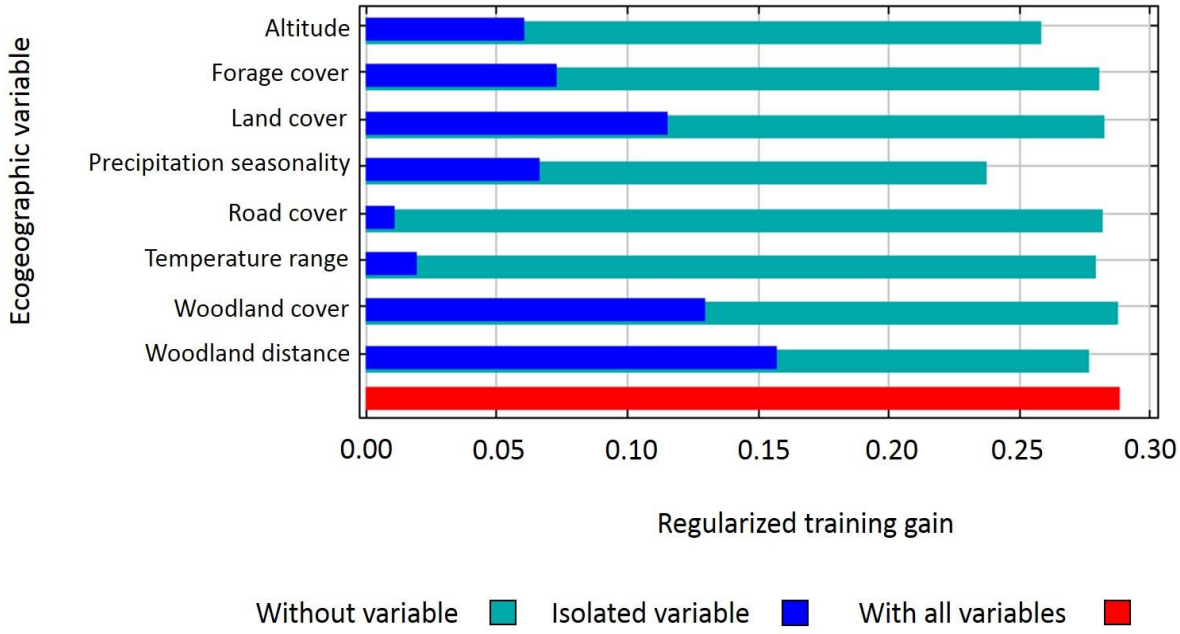


Figure 2. Importance of ecogeographic variables to the predictions of habitat suitability derived from the MaxEnt model for England and Scotland assessed using a jackknife test (Phillips, Anderson & Schapire 2006).

Table 4. Relative contribution of ecogeographic variables in the MaxEnt model of roe deer in England and Scotland.

| Ecogeographic variable | Relative contribution (%) |
|---|---|
| Woodland distance | 34.0 |
| Precipitation seasonality | 16.3 |
| Woodland cover | 15.4 |
| Altitude | 14.9 |
| Land cover | 6.6 |
| Forage cover | 6.2 |
| Temperature range | 4.4 |
| Road cover | 2.2 |

### 3.1.2 *Relative occurrence ratio*

Very high and high suitability was predicted for approximately 13.2% (25,980 $km^2$) and 4.7% (9,303 $km^2$) of the England and Scotland study region, respectively (Fig. 3a). Suitability of the remaining area was predicted as neutral, 27.4% (53,815 $km^2$); low; 33.0% (64,956 $km^2$) and very low, 21.6% (42,497 $km^2$) (Fig. 3a). As predicted, the representation of presences was disproportionately high (>1) on very high and high suitability habitat, commensurate with area (≅ 1) for neutral habitat and disproportionately low (<1) on low and very low suitability habitat (Fig. 3b, 4a & 4b).

In Wales, very high suitability and high suitability was predicted for approximately 10.0% (2,001 $km^2$) and 2.3% (459 $km^2$) of the region, respectively (Fig. 3c). Suitability of the remaining area was predicted as neutral, 24.3% (4,852 $km^2$); Low; 41.3% (8,254 $km^2$) and very low, 22.1% (4,428 $km^2$) (Fig. 3c). As for England and Scotland, a positive curvilinear relationship between proportionate representation and habitat suitability was observed (Fig. 4d). However, representation was higher for low and neutral suitability habitat and lower for high and very high suitability habitat when compared with results from England and Scotland (Fig. 4).

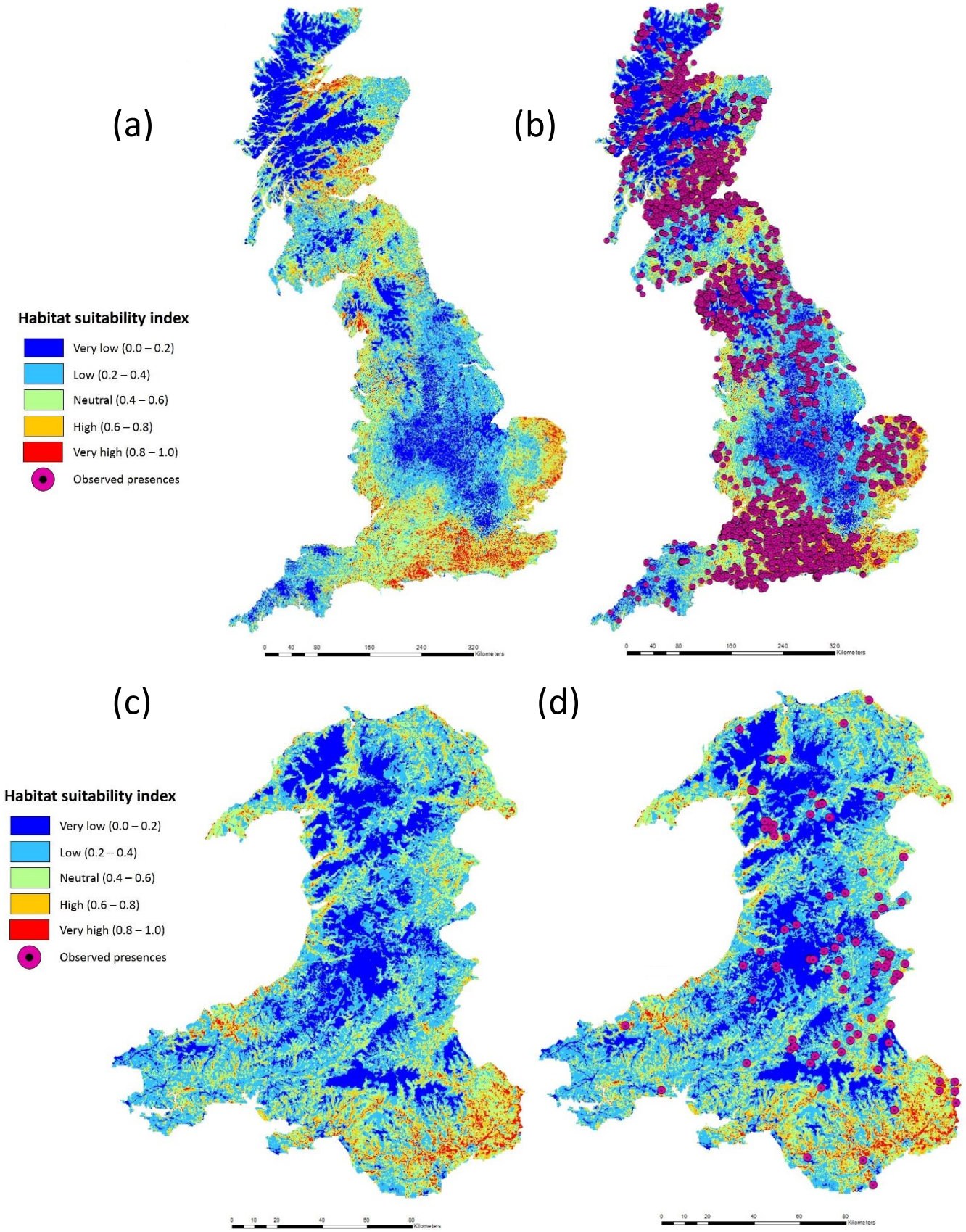


Figure. 3. Predicted suitability of habitat for roe deer from the MaxEnt model in; (a) England and Scotland and (c) Wales. Maps (b) and (d) show the locations of observed presences in England and Scotland (n=3,843) and Wales (n = 106), respectively. Observations of presence were taken from the National Biodiversity Network (NBN) Gateway, Biodiversity Information Service, Wildlife Trust of South West Wales and the North Wales Environmental Information Service.

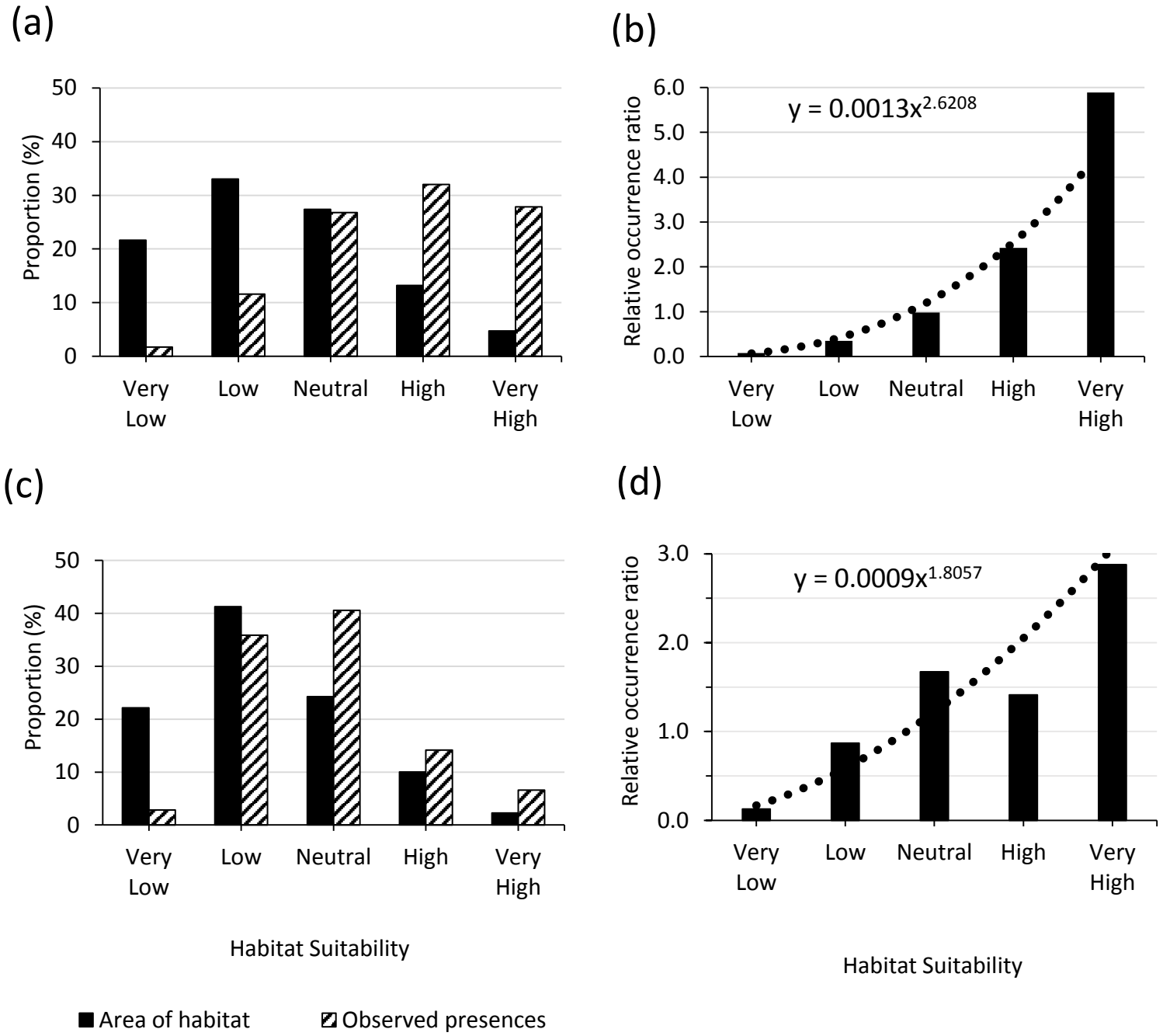


Figure 4. Relationship between the proportionate area of habitat and representation of observed presences of roe deer (a) for England and Scotland and (c) for Wales. Relationship between relative occurrence ratio, as calculated from eqn. 1, and habitat suitability (b) for England and Scotland (n = 3,843) and (d) for Wales (n = 106).Categories are based on suitability values predicted by the MaxEnt model: Very low (0-0.2); Low (0.2-0.4); Neutral (0.4-0.6); High (0.6-0.8) and Very high (0.8-1). Equations for the trend lines are shown at the top of panels (b) and (d).

### 3.2 (2) Mechanistic individual-based model

Contiguous patches of suitable habitat derived from the correlative model covered 35% (7,025 km$^2$) of the Wales study region. A total of 3,180 patches were initialised with a starting population of 5,549 individuals (Table 5). The average population size increased to 27,394 $\pm$ 1101 (mean $\pm$ standard deviation) and approximately 92% (6,474 km$^2$) of available patches were occupied by the end of the simulation after 300 years (Table 5). Assuming a sex ratio of 1:1, the predicted population density was 2.71 individuals/km$^2$ which is consistent with the current estimate for England and Scotland (4.03 individuals/km$^2$)(POST 2009). As anticipated, the compound rate of expansion predicted by the model was much lower (0.36%) than expected from observations of free ranging populations. Based on estimates of expansion rates from 1972-2002 (2.3%; Ward 2005) and 2002-2007 (5.2%; Palmer 2014), the patterns of recolonization, shown in Fig. 5, are predicted to occur within 47 and 21 years, respectively (Table 5).

Table 5. Summary of results from five replicate simulations of a mechanistic individual-based model for roe deer in Wales using six 50-year time steps.

| | | | | | Time elapsed (years) by compound annual expansion rate | | |
|---|---|---|---|---|---|---|---|
| Time step | Population size ($\pm$ standard deviation) | Area of occupied patches (km$^2$) | Colonisation of the region (%) | Colonisation of suitable habitat (%) | 0.36% (model) | 2.30% (Ward, 2005) | 5.20% (Palmer, 2014) |
| 0 | 5,549 | 2220 | 11 | 32 | 0 | 0 | 0 |
| 1 | 5,285 ($\pm$51) | 2254 | 11 | 32 | 50 | 8 | 3 |
| 2 | 16,130 ($\pm$143) | 3668 | 18 | 52 | 100 | 16 | 7 |
| 3 | 18,486 ($\pm$454) | 4434 | 22 | 63 | 150 | 23 | 10 |
| 4 | 21,619 ($\pm$672) | 5282 | 26 | 75 | 200 | 31 | 14 |
| 5 | 24,734 ($\pm$752) | 6001 | 30 | 85 | 250 | 39 | 17 |
| 6 | 27,394 ($\pm$1101) | 6474 | 32 | 92 | 300 | 47 | 21 |

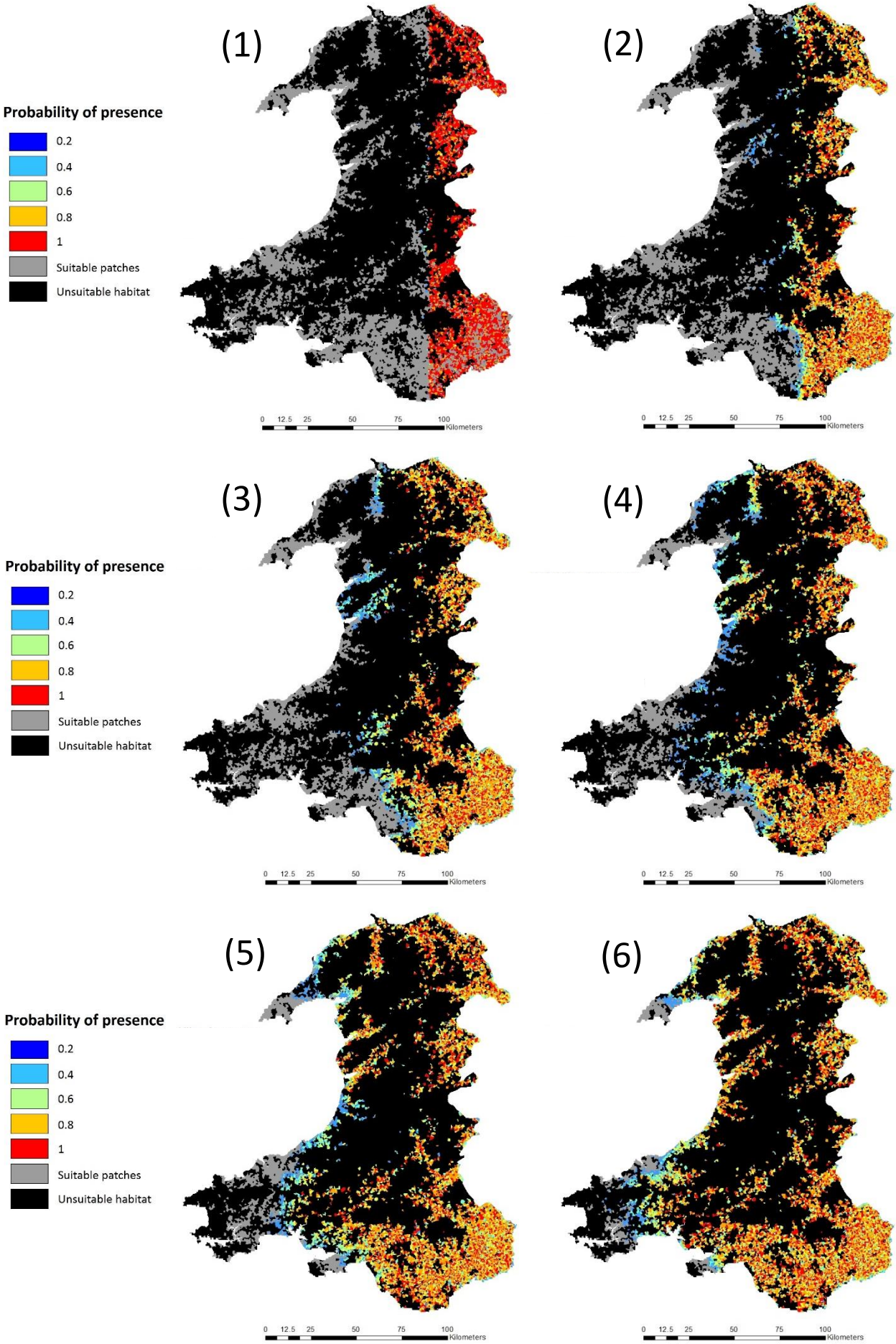


Figure 5. Predicted pattern of roe deer range expansion in Wales derived from five replicate simulations of a mechanistic individual-based model, RangeShifter (Bocedi *et al.* 2014). Panels (1)-(6) show the six 50-year time steps

## 4. DISCUSSION

This study introduces the first spatially explicit species distribution model developed to estimate roe deer habitat suitability at a 100 $m^2$ resolution using presence-only data on a national scale. Although roe deer are native to Britain, the relatively recent expansion of populations from reintroduction sites in England is similar to the spread of an invasive species (Ward 2005; Carey et al. 2012; Morelle et al. 2016). Presented here, is the first application of a distribution model to the range expansion of roe deer in a novel area. Results from the mechanistic model describe the expected pattern of range expansion across Wales. Based on the estimated expansion rate of free-ranging populations from 1972 to 2002 (Ward 2005), it is likely that over 90% of suitable habitat will be occupied within 50 years (Table 5). A more recent estimate, from 2002 to 2007 (Palmer 2014), suggests that near-total colonization of Wales will occur within 21 years.

In Europe, roe deer are one of several ungulate species that are increasing in numbers and geographic distribution in response to climate and land-use modification (Ward 2005; Acevedo *et al.* 2011; Apollonio *et al.* 2017). The spread of roe deer in Wales could have several ecological and socio-economic impacts. Biodiversity of conservation-priority habitats may be at risk as selective foraging can alter nutrient cycling, change community structure and influence vegetation dynamics (Côté et al. 2004). Browsing damage can also significantly reduce the economic value of forestry and agricultural crops (Putman & Moore 1998; White & Smart 2004; Ward *et al.* 2008). Human welfare and economic damage can occur from an increase in deer-vehicle collisions which are a major source of road traffic accidents (POST 2009; Langbein 2011). Species distribution modelling was used to forecast the expansion of roe deer in Wales and support the development of a proactive management strategy. Each stage of development was informed by (1) the type and quality of occurrence data available, (2) the ecology of the focal species and (3) the intended application of the output (Guisan and Thuiller 2005, Elith and Leathwick 2009). Model performance was assessed on the grounds of statistical robustness and biological legitimacy.

### 4.1 (1) Modelling presence-only data at a high spatial resolution

#### 4.1.1 *Identifying the drivers of range expansion*

The high resolution of the model allowed for the identification of the fine-scale drivers of population expansion. Response curves describing the relationship between each predictor variable and estimated suitability were assessed throughout the model development to ensure that predictions were consistent with habitat requirements estimated from free ranging populations. The strong trend observed in the distribution of incidental sightings towards woodlands is consistent with patterns of resource use shown in other studies (Tufto, Andersen & Linnell 1996; Mysterud *et al.* 1999; Lovari, Serrao & Mori 2017). Suitability was negatively correlated with altitude and road cover which also aligns with existing knowledge of habitat preferences (Coulon *et al.* 2008; Acevedo *et al.* 2010b; Loro *et al.* 2016). The minimal contribution of temperature range to the model was expected given that conditions in the UK are well within the limits assumed from the wide distribution of roe deer in Europe (Torres *et al.* 2012a). The moderate influence of precipitation seasonality is most likely linked with primary productivity and the availability of food resources (Morellet *et al.* 2013).

Finer detail could be incorporated in future models by accounting for habitat structure. In this study, landscape features were characterised using two-dimensional land cover data. Information was grouped into aggregate classes and all habitats within the same class were considered to be homogeneous. The three-dimensional structure of ecosystems is likely to influence the abundance and distribution of many species (Hinsley *et al.* 2006; Palminteri *et al.* 2012; Müller *et al.* 2014). Three-dimensional habitat features such as canopy cover, understorey composition and snow depth have been shown to affect resource use by roe deer (Ewald *et al.* 2014). Information on habitat structure can be obtained from light detection and ranging (lidar) remote sensing technology, which uses laser scanning to produce three-dimensional maps of surface structure (Lefsky *et al.* 2002; Vierling *et al.* 2008; Davies & Asner 2014). This enables maps to be produced at a very high resolution for a large spatial extent (Lefsky *et al.* 2002; Vierling *et al.* 2008). Models have been developed using lidar data to study the foraging behaviour of moose (*Alces alces*) (Lone *et al.* 2014), assess habitat selection by white-tailed deer (*Ococoileus virginianus*) (Wiemers *et al.* 2014) and investigate the spatial dynamics of Asian crested ibis (*Nipponia nippon*) (Sun *et al.*

2016). Lidar data from aircraft-mounted sensors is available at a 25 cm resolution for Wales but datasets are currently incomplete for the region (https://data.gov.uk/dataset/lidar-terrainand-surfaces-models-wales). Collecting lidar data will provide a valuable resource to the field of distribution modelling and allow the identification of the drivers of spatial dynamics at increasingly finer scales (Lefsky *et al.* 2002; Davies & Asner 2014).

#### 4.1.2 *Assumptions of the correlative model*

It was assumed in this study that roe deer populations are stable across England and Scotland and the species is, therefore, in equilibrium with its environment (Elith & Leathwick 2009; Elith, Kearney & Phillips 2010). However, it is likely that not all suitable environments have been reached as populations may still be expanding in some parts of the region such as in Kent and the English Midlands (Ward 2005). Estimating the current known range of a population can be challenging, particularly for a species, like roe deer, which is often difficult to detect (Putman, Watson & Langbein 2011). If range extent is over-estimated, areas that are poorly represented in the observation data may be erroneously predicted as unsuitable (Elith, Kearney & Phillips 2010). If a species range is under-estimated, habitat that is genuinely unsuitable may not be considered in analyses, which introduces a bias towards occupied areas and an over-prediction of suitability (Elith & Leathwick 2009). In this study, a conservative approach was taken, but improved knowledge of the extent of current ranges would improve the accuracy of future models.

Assumptions were also made with regard to sampling effort, which is indeterminable for data collected through non-systematic processes (Elith & Leathwick 2009; Guillera-Arroita *et al.* 2015). It was necessary to assume that the species data represented a random sample from the entire study area (Phillips 2009; Royle *et al.* 2012). However, it is possible that some areas will have been surveyed more comprehensively than others if the probability of detecting the animal was higher or the area was more accessible to observers (Phillips 2009; Kramer-Schadt *et al.* 2013). Spatial biases in sampling can inflate estimates of model accuracy due to spatial autocorrelation (Dormann *et al.* 2007c; Kramer-Schadt *et al.* 2013). There are two common methods used to account for sampling bias (Kramer-Schadt *et al.* 2013). Occurrence data can be spatially filtered to remove records from oversampled

regions (Dormann *et al.* 2007; Kramer-Schadt *et al.* 2013). Alternatively, background data can be manipulated to reflect the same bias as the occurrence data (Phillips 2009; Kramer-Schadt *et al.* 2013). Spatial filtering is typically suited to large datasets as it reduces the number of occurrence records that are available for the development of a model (Kramer-Schadt *et al.* 2013). As species data are limited for roe deer, manipulation of background data is likely to be a more appropriate approach. Predictive accuracy of future models could be improved by observing the distribution of sympatric species in order to distribute background points to areas that are likely to have been sampled more thoroughly (Phillips 2009; Smith 2013).

Several studies have adopted an ensemble approach to modelling species range expansions (Araújo *et al.* 2005; Marmion *et al.* 2009; Roura-Pascual *et al.* 2009; Millington *et al.* 2010). Model performance is based on the agreement of predictions made by multiple types of model, such as boosted regression trees, random forests, artificial neural networks and regression models (Marmion *et al.* 2009; Elith, Kearney & Phillips 2010). The ensemble approach facilitates the quantification of model uncertainty with greater confidence inferred for more congruent predictions (Marmion *et al.* 2009; Roura-Pascual *et al.* 2009). However, evaluation in the context of range expansions is challenging as true validation of predictions is only possible after the species has colonised the novel area (Elith, Kearney & Phillips 2010). Differences in performance may result from model error if the unique assumptions of each type of model are not sufficiently validated (Elith, Kearney & Phillips 2010). Alternatively, while predictive accuracy may be equally poor for all models good overall performance may be interpreted if predictions are in strong agreement (Elith, Kearney & Phillips 2010). In this study, a single model approach was taken and a stronger focus was placed on achieving biologically relevant predictions.

#### 4.1.3 *Model evaluation*

Performance of the model was evaluated based on observations of the area under the receiver operating curve (AUC) from $k$-fold and $n - 1$ cross validation. AUC scores are often misinterpreted as definitive measures of model accuracy, with values closer to 1 considered an indication of stronger predictive performance (Pearce & Ferrier 2000; Phillips, Anderson

& Schapire 2006; Lobo, Jiménez-Valverde & Real 2008). However, such inferences have been criticised following closer scrutiny of the AUC method (Lobo, Jiménez-Valverde & Real 2008; Jiménez-Valverde 2012). AUC scores describe the ability of the model to correctly rank presences higher than background points (Lobo, Jiménez-Valverde & Real 2008). Measurements are insensitive to the probability values estimated and it is possible to achieve a high AUC score from a poorly fitted model (Lobo, Jiménez-Valverde & Real 2008; Jiménez-Valverde 2012). No information is provided on the spatial distribution of model errors (Lobo, Jiménez-Valverde & Real 2008). If a low AUC score is estimated, it is impossible to ascertain if predictive performance is poor across the entire study area or just for a specific region (Lobo, Jiménez-Valverde & Real 2008). Additionally, scores are highly sensitive to the extent of the model, which reduces the reliability of comparisons of performance between regions of different sizes (Lobo, Jiménez-Valverde & Real 2008; Jiménez-Valverde 2012).

In this study, the AUC scores attained from the cross validation analysis were interpreted only as an indication that the model correctly predicted suitability better than random (Lobo, Jiménez-Valverde & Real 2008; Kearney, Wintle & Porter 2010; Jiménez-Valverde 2012). AUC scores were also used to compare predictive performance between the two study regions and sources of discrepancies were identified using a novel, qualitative technique. Suitability values from the model were classified into five categories. Performance was assessed by observing the proportion of presences in relation to the area of the study region in different ranges of habitat suitability, described as the relative occurrence ratio (ROR). The transparency of the qualitative method enabled an intuitive comparison of model performance between the two study regions. As the ROR was a relative measure of both; the total area of the study region and the total number of occurrence records, the same technique could be applied to regions of unequal size using different numbers of occurrence records.

As predicted, model performance described by the AUC scores was lower in Wales compared to England and Scotland. In both regions, very highly suitable habitat was disproportionately better represented than neutral and unsuitable habitat. However, in Wales, neutral and low-suitability habitat was comparably over-represented and highly suitable habitat under-represented. This supports the results of the cross validation analysis

and indicates under-prediction by the model in Wales which would suggest that predictions of suitability for the region are conservative. It should be recognised that the reliability of results obtained for each region is likely to be influenced by the number of occurrence records that were used in the analysis. The results from Wales were more sensitive to outliers than those from England and Scotland, as the Welsh analysis was performed using considerably fewer occurrence records (n = 106 & n = 3,843, respectively) (Moudrý & Šímová 2012). A more reliable assessment will only be possible as more sightings are recorded. It is recommended that the effect of sample size is acknowledged in future inferences from observations of the relative occurrence ratio.

### 4.2 (2) Integration of a mechanistic model

#### 4.2.1 *Capturing ecological realism*

Species distribution modelling has received considerable criticism in the past for not sufficiently accounting for ecological theory (Austin 2002; Jiménez-Valverde, Lobo & Hortal 2008; Elith & Leathwick 2009). A wide variety of modelling techniques have been developed in the last two decades, and although methods have advanced in terms of statistical robustness, the ecological realism of models is often contested (Austin 2002; Jiménez-Valverde, Lobo & Hortal 2008; Guillera-Arroita *et al.* 2015). In this study, a correlative species distribution model was integrated with a mechanistic model (RangeShifter) to improve the robustness of predictions and estimate the patterns of range expansion using ecological knowledge (Kearney & Porter 2009; Elith, Kearney & Phillips 2010; Bocedi *et al.* 2014). Dispersal was modelled through context-dependent decisions informed by landscape characteristics and local population densities (Bocedi *et al.* 2014). Demographic data were used to simulate the discrete phases of emigration, transfer and settlement. Patterns of expansion emerged as a function of the interactions between individuals and the simulated environment (Bocedi *et al.* 2014). Conventional models typically describe dispersal probabilities in relation to the Euclidean distance between patches (Aben *et al.* 2016). This study demonstrates how important landscape features, such as terrain and land cover, can influence patterns of expansion and how more realistic dispersal simulations better account for the key drivers of spatial dynamics (Aben *et al.* 2016). Sensitivity analysis on mechanistic

models can be performed to estimate the effect of certain parameters on predictions (Tews, Ferguson & Fahrig 2007; Sun *et al.* 2016). However, such analyses were infeasible in this study due to the complexity of the model (Dormann *et al.* 2012).

### 4.2.2 *Improving the RangeShifter model*

Although the RangeShifter software has become increasingly popular, it has only recently been developed and there are still some features of the model that could be improved. Probabilities of emigration and settlement are modelled in RangeShifter as a function of local population densities (Bocedi *et al.* 2014). For many species it is estimated that the probability of dispersal will be greater as competition for resources increases (Gaston *et al.* 2000). However, dispersal behaviour can also be influenced by factors independent of population density such the availability of resources in neighbouring regions (Sæther & Heim 1993; Nilsen, Linnell & Andersen 2004). In the case of roe deer, it has been shown that the probability of emigration by yearlings is related to resource quality and the amount of cover that is available in the surrounding area (Nilsen, Linnell & Andersen 2004). In the absence of competition, dispersing individuals are also more likely to be selective in their choice of where to settle, preferring areas that are richest in resources to maximise their gain in fitness (Nilsen, Linnell & Andersen 2004).

Currently, RangeShifter does not have the capacity to directly simulate resource dependency in dispersal. It is possible to simulate dispersal using fixed probabilities of emigration and settlement but models can easily become too computationally demanding. High numbers of non-dispersing individuals are simulated, which due to density dependency in fecundity also do not reproduce and so contribute very little to the overall pattern of range expansion. In this study, the effects of landscape features on dispersal were modelled indirectly by introducing density dependencies in juvenile survival and settlement probability, which had a profound effect on predictions of expansion rate. The predicted rate of expansion in the model (0.36%) was much lower than estimates for free-ranging populations (2.3%, Ward 2005; 5.2%, Palmer 2014). A trade off emerged between attaining spatial accuracy and a reasonable rate of expansion. Efforts to increase the computational

capacity of RangeShifter and the inclusion of resource-dependent emigration could help to reduce this trade-off and, ultimately, improve the functionality of the model.

Some limitations of RangeShifter were identified when it was introduced and still need to be formally addressed (Bocedi *et al.* 2014). Populations are simulated, in the current framework, on static landscapes (Bocedi *et al.* 2014). Climate and the availability of resources are assumed to be constant, which is unlikely to reflect a natural system. Trials have been conducted to incorporate details of interspecific interactions in the model but a firm methodology has not yet been established (Bocedi *et al.* 2014). Movement behaviour is modelled based on costs defined by the user, typically using an inverted habitat suitability map (Bocedi *et al.* 2014). This largely reflects the selection of resources and avoidance of highly unsuitable habitat. However, the direction and speed of movement is also likely to be influenced by key variables such as topography and landscape structure (Kie, Ager & Bowyer 2005; Coulon *et al.* 2008; Semeniuk *et al.* 2012). Movement data from global positioning system (GPS) devices could be incorporated to improve the accuracy of modelled movement behaviour in a heterogeneous landscape (Cagnacci *et al.* 2010; Bocedi *et al.* 2014). Initiatives, such as the recent EURODEER project (http://eurodeer.org/), established to consolidate GPS data for roe deer across Europe, could provide the information necessary to estimate habitat and terrain-specific movement parameters (Cagnacci *et al.* 2010). Priority areas for the future development of RangeShifter are well documented by the developers of the model (Bocedi *et al.* 2014). The platform is already very complex and it is recognised that advancing the software is as much a computational challenge, as it is methodological.

### 4.3 (3) Model application

#### 4.3.1 *Confidence in modelling techniques*

A review by Guisan *et al.* (2013) revealed that species distribution modelling has been explicitly applied to support conservation decisions in < 1% of published papers. Alternative approaches are frequently favoured by conservation planners due to the paucity of species data required to develop a reliable model (Tulloch *et al.* 2016). The functionality of modelling techniques is usually constrained by a trade-off between the resolution and scale at which species-environment relationships can be estimated (Pearce & Boyce 2006).

Concern over uncertainties in model predictions and scepticism about the capacity of methods to account for important ecological and evolutionary processes have also limited their use by conservation planners (Addison *et al.* 2013; Tulloch *et al.* 2016).

### 4.3.2 *Developing a management plan for roe deer*

This study demonstrates the utility of population modelling for making statistically robust and ecologically relevant predictions of species distributions from limited data. The methodology developed here could be applied to inform a broad range of management decisions. Modelling can be used for adaptive management of biological invasions (With 2002), identification and protection of vulnerable habitat (Heinrichs *et al.* 2010), selection of reserves (Margules & Pressey 2000; Wilson *et al.* 2010) and viability assessments of translocations (Bar-David *et al.* 2008; Guisan *et al.* 2013). The spread of roe deer in Wales provides a good example of a management scenario that could be informed by predictive distribution modelling. The Welsh Government has acknowledged the implications of expanding deer populations and established The Wild Deer Management Strategy to improve the resilience of biodiversity in Wales (Welsh Government 2017). Actions detailed in the supporting plan, developed by the Wales Deer Forum, are intended to help deliver other government strategies such as the LIFE Natura 2000 programme, which has designated 92 Special Areas of Conservation (SACs) and 20 Special Protection Areas (SPAs) in Wales (Welsh Government 2017). A strong focus is placed on the conservation of woodland habitats and some strategies, such as Woodlands for Wales (2009), specifically list the creation of new woodland as a key objective (Welsh Government 2017). Results from the individual-based model indicate that the future range of roe deer in Wales is likely to include 777 $km^2$ of SPAs, 1091 $km^2$ of SACs, 1822 $km^2$ of sites of special scientific interest (SSSIs) and 710 $km^2$ of ancient woodlands. As the availability of woodland has been identified in this study as the principle driver of roe deer range expansion, it is likely that populations will have to be managed rigorously to prevent increasing numbers from jeopardizing conservation efforts in these key designated areas.

Reliable forecasting of population expansions should facilitate the implementation of a proactive management strategy. The high-resolution suitability map produced by the

correlative model in the present study highlights areas on local and landscape scales that are most vulnerable to increased browsing pressure. If sightings are recorded in a novel area, the patterns of range expansion predicted by the mechanistic model, reveal where in the neighbouring region populations are most likely to spread to. The output from both models developed in this study can be used to target susceptible habitats for investment in protective fencing and to coordinate collaborative culls at the appropriate scale to limit expansion from recently populated areas (Putman, Watson & Langbein 2011). Diligent monitoring of populations through physical techniques is, however, still essential for effective management. The models developed in this study could be used to direct surveillance towards areas of potential occupancy thereby making such actions more efficient and cost-effective (Guisan *et al.* 2013). As more data are collected, the correlative model can be evaluated more reliably and the mechanistic model can be adapted to predict the performance of various mitigation strategies.

#### 4.3.3 *Developing a pan-European ungulate management strategy*

Ungulate populations have been expanding in Europe for several decades (Apollonio *et al.* 2017). Recognition of the value of species data to conservation has prompted the development of online databases in many European countries to support a multi-national approach to management (Zaragozí *et al.* 2015; Bubnicki, Churski & Kuijper 2016; Helle & Ikonen 2016; Apollonio *et al.* 2017). Improvements to modelling techniques should help to make more effective use of spatial data and ensure that management decisions are informed by robust scientific evidence (Tulloch *et al.* 2016; Apollonio *et al.* 2017).

The methods used in this study could be adapted to predict the expansion patterns of a wide range of species and contribute to the development of a multi-national ungulate management strategy (Apollonio *et al.* 2017). However, the spatial resolution of models is largely governed by the movement behaviour of the focal species (Guisan & Thuiller 2005; Guisan *et al.* 2013). Because roe deer show high fidelity to small home ranges, it was reasonable to assume that environmental conditions at locations where the species is present described its habitat requirements or preferences (Vincent et al. 1995; Wahlström & Liberg 1995; Guisan & Thuiller 2005). High-resolution modelling using presence-only data

is less feasible for more mobile or migratory species as the area that could be considered as occupied by the species is much larger and it is difficult to conclude that features of the local area are being utilised at the time of recording (Guisan & Thuiller 2005). For species that migrate, such as caribou (*Rangifer tarandus*), or have large seasonal home ranges, such as red deer (*Cervus elaphus*), tracking data obtained from GPS devices may be required to estimate the behavioural mode of the species and identify the fine scale drivers of movement (Gautestad, Loe & Mysterud 2013; Dalziel *et al.* 2015). Mechanistic models are also often limited by the availability of demographic data (Elith, Kearney & Phillips 2010). Extensive research on roe deer ecology provided evidence for every parameter in the RangeShifter model in the present study. For rare or endangered species, however, it may be more difficult to obtain the information on life history traits and dispersal behaviour required for parameterization (Kearney & Porter 2009; Elith, Kearney & Phillips 2010). Better application of modelling to management objectives should promote the collection of more distribution and demographic data for a wider variety of species which will positively feedback to method development.

### 4.4 Conclusions

This study demonstrates that it's possible to produce robust, high-resolution predictions from models developed with presence-only data. Habitat suitability and patterns of range expansion were estimated for a large mammal at a national scale using incidental sightings and demographic knowledge. Details of the decisions made during model development were described to highlight the importance of considering the caveats of techniques used, ecological theory and the intended application of the output. Relationships between suitability predictions and ecogeographic variables were explicitly considered to critically assess the ecological validity of the correlative model. An integrated approach using a mechanistic model provided a flexible tool that can be modified to inform a variety of conservation decisions. Both models are adaptable and can be refined as data collection and computational power improve.

Predictions for the expansion of roe deer in Wales support the objectives of the Wales Deer Forum and can be used to manage populations as a valuable and sustainable natural

resource (Welsh Government 2017). The techniques demonstrated could be applied to a wide range of species and help to achieve the vision of pan-European ungulate management (Apollonio *et al.* 2017). The transparency of methods used will hopefully reinforce confidence in distribution modelling among conservation planners and promote their wider use for assessing species-environment relationships.

## 5. REFERENCES


Aben, J., Bocedi, G., Palmer, S.C.F., Pellikka, P., Strubbe, D., Hallmann, C., Travis, J.M.J., Lens, L. & Matthysen, E. (2016) The importance of realistic dispersal models in conservation planning: application of a novel modelling platform to evaluate management scenarios in an Afrotropical biodiversity hotspot. *Journal of Applied Ecology*, **53**, 1055–1065.

Acevedo, P., Cassinello, J., Hortal, J. & Gort??zar, C. (2007) Invasive exotic aoudad (Ammotragus lervia) as a major threat to native Iberian ibex (Capra pyrenaica): A habitat suitability model approach. *Diversity and Distributions*, **13**, 587–597.

Acevedo, P., Delibes-Mateos, M., Escudero, M.A., Vicente, J., Marco, J. & Gortazar, C. (2005) Environmental constraints in the colonization sequence of roe deer (Capreolus capreolus Linnaeus, 1758) across the Iberian Mountains, Spain. *Journal of Biogeography*, **32**, 1671–1680.

Acevedo, P., Farfán, M.Á., Márquez, A.L., Delibes-Mateos, M., Real, R. & Vargas, J.M. (2011) Past, present and future of wild ungulates in relation to changes in land use. *Landscape Ecology*, **26**, 19–31.

Acevedo, P., Ferreres, J., Escudero, M.A., Jimenez, J., Boadella, M. & Marco, J. (2017) Population dynamics affect the capacity of species distribution models to predict species abundance on a local scale. *Diversity and Distributions*, **23**, 1008–1017.

Acevedo, P., Ferreres, J., Jaroso, R., Durán, M., Escudero, M.A., Marco, J. & Gortázar, C. (2010a) Estimating roe deer abundance from pellet group counts in Spain: An assessment of methods suitable for Mediterranean woodlands. *Ecological Indicators*, **10**, 1226–1230.

Acevedo, P., Ward, A.I., Real, R. & Smith, G.C. (2010b) Assessing biogeographical relationships of ecologically related species using favourability functions: A case study on British deer. *Diversity and Distributions*, **16**, 515–528.

Addison, P.F.E., Rumpff, L., Bau, S.S., Carey, J.M., Chee, Y.E., Jarrad, F.C., McBride, M.F. & Burgman, M.A. (2013) Practical solutions for making models indispensable in conservation decision-making. *Diversity and Distributions*, **19**, 490–502.

Akashi, N., Unno, A. & Terazawa, K. (2011) Effects of deer abundance on broad-leaf tree seedling establishment in the understory of Abies sachalinensis plantations. *Journal of Forest Research*, **16**, 500–508.

Allombert, S., Stockton, S. & Martin, J.L. (2005) A natural experiment on the impact of overabundant deer on forest invertebrates. *Conservation Biology*, **19**, 1917–1929.

Andanje, S.A. & Ottichilo, W.K. (1999) Population status and feeding habits of the translocated subpopulation of Hunter's antelope or hirola (Beatragus hunteri, Sclater, 1889) in Tsavo East National Park, Kenya. *African Journal of Ecology*, **37**, 38–48.

Andersen, R., Gaillard, J.M., Linnell, J.D.C. & Duncan, P. (2000) Factors affecting maternal care in an income breeder, the European roe deer. *Journal of Animal Ecology*, **69**, 672–682.

Andersen, R. & Linnell, J.D.C. (2000) Irruptive Potential in Roe Deer : Density-Dependent Effects on Body Mass and Fertility. *The Journal of Wildlife Management*, **64**, 698–706.

Apollonio, M., Belkin, V. V., Borkowski, J., Borodin, O.I., Borowik, T., Cagnacci, F., Danilkin, A.A., Danilov, P.I., Faybich, A., Ferretti, F., Gaillard, J.M., Hayward, M., Heshtaut, P., Heurich, M., Hurynovich, A., Kashtalyan, A., Kerley, G.I.H., Kjellander, P., Kowalczyk, R., Kozorez, A., Matveytchuk, S., Milner, J.M., Mysterud, A., Ozoliņš, J., Panchenko, D. V., Peters, W., Podgórski, T., Pokorny, B., Rolandsen, C.M., Ruusila, V., Schmidt, K., Sipko, T.P., Veeroja, R., Velihurau, P. & Yanuta, G. (2017) Challenges and science-based implications for modern management and conservation of European ungulate populations. *Mammal Research*, **62**, 209–217.

Araújo, M.B., Pearson, R.G., Thuillers, W. & Erhard, M. (2005) Validation of species–climate impact models under climate change. *Global Change Biology*, **11**, 1504–1513.

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

Austin, Z., Smart, J.C.R., Yearley, S., Irvine, R.J. & White, P.C.L. (2010) Identifying conflicts and opportunities for collaboration in the management of a wildlife resource: A mixed-methods approach. *Wildlife Research*, **37**, 647–657.

Avgar, T., Mosser, A., Brown, G.S. & Fryxell, J.M. (2013) Environmental and individual drivers of animal movement patterns across a wide geographical gradient. *Journal of Animal Ecology*, **82**, 96–106.

Baker, K.H. & Rus Hoelzel, A. (2013) Evolution of population genetic structure of the British roe deer by natural and anthropogenic processes (Capreolus capreolus). *Ecology and Evolution*, **3**, 89–102.

Bar-David, S., Saltz, D., Dayan, T. & Shkedy, Y. (2008) Using spatially expanding populations as a tool for evaluating landscape planning: The reintroduced Persian fallow deer as a case study. *Journal for Nature Conservation*, **16**, 164–174.

Bennett, A.F., Haslem, A., Cheal, D.C., Clarke, M.F., Jones, R.N., Koehn, J.D., Lake, P.S., Lumsden, L.F., Lunt, I.D., MacKey, B.G., Mac Nally, R., Menkhorst, P.W., New, T.R., Newell, G.R., O'Hara, T., Quinn, G.P., Radford, J.Q., Robinson, D., Watson, J.E.M. & Yen, A.L. (2009) Ecological processes: A key element in strategies for nature conservation. *Ecological Management and Restoration*, **10**, 192–199.

Bocedi, G., Palmer, S.C.F., Pe'er, G., Heikkinen, R.K., Matsinos, Y.G., Watts, K. & Travis, J.M.J. (2014) RangeShifter: A platform for modelling spatial eco-evolutionary dynamics and species' responses to environmental changes. *Methods in Ecology and Evolution*, **5**, 388–396.

Borowik, T., Cornulier, T. & Jedrzejewska, B. (2013) Environmental factors shaping ungulate abundances in Poland. *Acta Theriologica*, **58**, 403–413.

Bouyer, Y., Rigot, T., Panzacchi, M., Van Moorter, B., Poncin, P., Beudels-Jamar, R., Odden, J. & Linnell, J.D.C. (2015) Using zero-inflated models to predict the relative distribution and abundance of roe deer over very large spatial scales. *Annales Zoologici Fennici*, **52**, 66–76.

Boyce, M.S., Vernier, P.R., Nielsen, S.E. & Schmiegelow, F.K.A. (2002) Evaluating resource selection functions. *Ecological Modelling*, **157**, 281–300.

Bressette, J.W., Beck, H. & Beauchamp, V.B. (2012) Beyond the browse line: Complex cascade effects mediated by white-tailed deer. *Oikos*, **121**, 1749–1760.

Brudvig, L.A., Leroux, S.J., Albert, C.H., Bruna, E.M., Davies, K.F., Ewers, R.M., Levey, D.J., Pardini, R. & Resasco, J. (2017) Evaluating conceptual models of landscape change. *Ecography*, **40**, 74–84.

Bubnicki, J.W., Churski, M. & Kuijper, D.P.J. (2016) Trapper: an Open Source Web-Based Application To Manage Camera Trapping Projects. *Methods in Ecology and Evolution*, **7**, 1209–1216.

Burbaite, L. & Csányi, S. (2009) Roe deer population and harvest changes in Europe. *Estonian Journal of Ecology*, **58**, 169–180.

Cagnacci, F., Boitani, L., Powell, R. a & Boyce, M.S. (2010) Animal ecology meets GPS-based radiotelemetry: a perfect storm of opportunities and challenges. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, **365**, 2157–62.

Cassinello, J., Acevedo, P. & Hortal, J. (2006) Prospects for population expansion of the exotic aoudad (Ammotragus lervia; Bovidae) in the Iberian Peninsula: Clues from habitat suitability modelling. *Diversity and Distributions*, **12**, 666–678.

Cawley, G.C. & Talbot, N.L.C. (2003) Efficient leave-one-out cross-validation of kernel fisher discriminant classifiers. *Pattern Recognition*, **36**, 2585–2592.

Chitwood, M.C., Lashley, M.A., Kilgo, J.C., Moorman, C.E. & Deperno, C.S. (2015) White-tailed deer population dynamics and adult female survival in the presence of a novel predator. *Journal of Wildlife Management*, **79**, 211–219.

Cobben, M.M.P., Linnell, J.D.C., Solberg, E.J. & Andersen, R. (2009) Who wants to live forever? roe deer survival in a favourable environment. *Ecological Research*, **24**, 1197–1205.

Cooke, A.S. & Lakhani, K.H. (1996) Damage to coppice regrowth by muntjac deer Muntiacus reevesi and protection with electric fencing. *Biological Conservation*, **75**, 231–238.

Le Corre, M., Pellerin, M., Pinaud, D., Van Laere, G., Fritz, H. & Saïd, S. (2008) A multi-patch use of the habitat: testing the First-Passage Time analysis on roe deer *Capreolus capreolus* paths. *Wildlife Biology*, **14**, 339–349.

Costanza, R., D'Arge, R., de Groot, R., Farber, S., Grasso, M., Hannon, B., Limburg, K., Naeem, S., O'Neill, R. V., Paruelo, J., Raskin, R.G., Sutton, P. & van den Belt, M. (1998) The value of the world's ecosystem services and natural capital. *Nature*, **387**, 253–260.

Côté, S.D., Rooney, T.P., Tremblay, J., Dussault, C. & Waller, D.M. (2004) Ecological Impacts of Deer Overabundance. *Annual Review of Ecology, Evolution, and Systematics*, **35**, 113–147.

Coulon, A., Cosson, J.-F., Morellet, N., Angibault, J.-M., Cargnelutti, B., Galan, M., Aulagnier, S. & Hewison, A.J.M. (2006) Dispersal is not female biased in a resource-defence

mating ungulate, the European roe deer. *Proceedings. Biological sciences / The Royal Society*, **273**, 341–348.

Coulon, A., Morellet, N., Goulard, M., Cargnelutti, B., Angibault, J.M. & Hewison, A.J.M. (2008) Inferring the effects of landscape structure on roe deer (Capreolus capreolus) movements using a step selection function. *Landscape Ecology*, **23**, 603–614.

Dalziel, B.D., Le Corre, M., Côté, S.D. & Ellner, S.P. (2015) Detecting collective behavior in animal relocation data, with application to migrating caribou. *Methods in Ecology and Evolution*, n/a-n/a.

Davies, A.B. & Asner, G.P. (2014) Advances in animal ecology from 3D-LiDAR ecosystem mapping. *Trends in Ecology and Evolution*, **29**, 681–691.

Davis, M.L., Stephens, P.A. & Kjellander, P. (2016) Beyond climate envelope projections: Roe deer survival and environmental change. *Journal of Wildlife Management*, **80**, 452–464.

Dawe, K.L., Bayne, E.M. & Boutin, S. (2014) Influence of climate and human land use on the distribution of white-tailed deer (*Odocoileus virginianus*) in the western boreal forest. , **363**, 353–363.

DeAngelis, D.L. & Mooij, W.M. (2005) Individual-Based Modeling of Ecological and Evolutionary Processes. *Annual Review of Ecology, Evolution, and Systematics*, **36**, 147–168.

Debeffe, L., Morellet, N., Cargnelutti, B., Lourtet, B., Coulon, A., Gaillard, J.M., Bon, R. & Hewison, A.J.M. (2013) Exploration as a key component of natal dispersal: Dispersers explore more than philopatric individuals in roe deer. *Animal Behaviour*, **86**, 143–151.

DeVore, R.M., Butler, M.J., Wallace, M.C., Liley, S.L., Mertz, A.A., Sesnie, S.E. & Gipson, P.S. (2016) Elk resource selection patterns in a semiarid riparian corridor. *Journal of Wildlife Management*, **80**, 479–489.

Dormann, C.F. (2007) Promising the future? Global change projections of species distributions. *Basic and Applied Ecology*, **8**, 387–397.

Dormann, C.F., McPherson, J.M., Araújo, M.B., Bivand, R., Bolliger, J., Carl, G., Davies, R.G., Hirzel, A., Jetz, W., Daniel Kissling, W., Kühn, I., Ohlemüller, R., R. Peres-Neto, P., Reineking, B., Schröder, B., Schurr, F.M. & Wilson, R. (2007) Methods to account for spatial autocorrelation in the analysis of species distributional data: A review. *Ecography*, **30**, 609–628.

Dormann, C.F., Schymanski, S.J., Cabral, J., Chuine, I., Graham, C., Hartig, F., Kearney, M., Morin, X., Römermann, C., Schröder, B. & Singer, A. (2012) Correlation and process in species distribution models: Bridging a dichotomy. *Journal of Biogeography*, **39**, 2119–2131.

Elith, J. & 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.

Elith, J., Graham, C.H., Anderson, R.P., Dudík, M., Guisan, A., Hijmans, R.J., Huettmann, F., Leathwick, J.R., Lehmann, A., Li, J., Lohmann, L.G., Loiselle, B.A., Manion, G., Nakamura,

M., Nakazawa, Y., Overton, J.M., Townsend, A., Phillips, S.J., Richardson, K., Scachetti-pereira, R., Robert, E., Soberón, J., Williams, S., Elith, J., Graham, C.H., Anderson, R.P., Dudik, M., Ferrier, S., Guisan, A., 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, A.T., Phillips, S.J., Richardson, K., Scachetti-pereira, R., Schapire, R.E., Sobern, J., Williams, S., Wisz, M.S. & Zimmermann, N.E. (2006) Novel Methods Improve Prediction of Species Distributions from Occurrence Data.

Elith, J., Kearney, M. & Phillips, S. (2010) The art of modelling range-shifting species. *Methods in Ecology and Evolution*, **1**, 330–342.

Elith, J. & Leathwick, J.R. (2009) Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. *Annual Review of Ecology Evolution and Systematics*, **40**, 677–697.

Ewald, M., Dupke, C., Heurich, M., Müller, J. & Reineking, B. (2014) LiDAR remote sensing of forest structure and GPS telemetry data provide insights on winter habitat selection of european roe deer. *Forests*, **5**, 1374–1390.

Feber, R.E., Brereton, T.M., Warren, M.S. & Oates, M. (2001) The impacts of deer on woodland butterflies: The good, the bad and the complex. *Forestry*, **74**, 271–276.

Felix, A.B., Walsh, D.P., Hughey, B.D., Campa, H. & Winterstein, S.R. (2007) Applying Landscape-Scale Habitat-Potential Models to Understand Deer Spatial Structure and Movement Patterns. *Journal of Wildlife Management*, **71**, 804–810.

Ferrier, S., Drielsma, M., Manion, G. & Watson, G. (2002) Extended statistical approaches to modelling spatial pattern in biodiversity in northeast New South Wales. II. *Biodiversity and Conservation*, **11**, 2309–2338.

Festa-Bianchet, M., Gaillard, J.M. & Côté, S.D. (2003) Variable age structure and apparent density dependence in survival of adult ungulates. *Journal of Animal Ecology*, **72**, 640–649.

Flowerdew, J.R. & Ellwood, S.A. (2001) Impacts of woodland deer on small mammal ecology. *Forestry*, **74**, 277–287.

Focardi, S., Raganella Pelliccioni, E., Petrucco, R. & Toso, S. (2002) Spatial patterns and density dependence in the dynamics of a roe deer (Capreolus capreolus) population in central Italy. *Oecologia*, **130**, 411–419.

Gaillard, J., Delormet, D., Boutint, J. & Laeret, G.U.Y.V.A.N. (1993) Roe Deer Survival Patterns : A Comparative Analysis of Contrasting Populations Author ( s ): Gaillard Jean-Michel , Daniel Delorme , Boutin Jean-Marie , Guy Van Laere , Bernard Boisaubert and Roger Pradel Source : Journal of Animal Ecology , Vol . 62 , No. , **62**, 778–791.

Gaillard, J.M., Festa-Bianchet, M. & Yoccoz, N.G. (1998) Population dynamics of large herbivores: variable recruitment with constant adult survival (vol 13, pg 58, 1998). *Trends in Ecology & Evolution*, **13**, 170.

Gaillard, J.-M., Hewison, A.J.M., Kjellander, P., PettorelliI, N., Bonenfant, C., Van Moorter, B., Liberg, O., Andrén, H., Van Laere, G., Klein, F., Angibault, J.M., Coulon, A. & Vanpé, C.

(2008) Population density and sex do not influence fine-scale natal dispersal in roe deer. *Proceedings Of The Royal Society B: Biological Sciences*, **275**, 2025–2030.

Gallien, L., Münkemüller, T., Albert, C.H., Boulangeat, I. & Thuiller, W. (2010) Predicting potential distributions of invasive species: Where to go from here? *Diversity and Distributions*, **16**, 331–342.

Gaston, K.J., Blackburn, T.M., Greenwood, J.J.D., Gregory, R.D., Quinn, R.M. & Lawton, J.H. (2000) Abundance-occupancy relationships. *Journal of Applied Ecology*, **37**, 39–59.

Gautestad, A.O., Loe, L.E. & Mysterud, A. (2013) Inferring spatial memory and spatiotemporal scaling from GPS data: Comparing red deer Cervus elaphus movements with simulation models. *Journal of Animal Ecology*, **82**, 572–586.

Gerard, J., Le Pendu, Y., Maublanc, M., Vincent, J.P., Poulle, M.L. & Cibien, C. (1995) Large group formation in European roe deer: An adaptive feature? *Revue D EcologieLa Terre Et La Vie*, **50**, 391–401.

Gill, R.M.A. & Beardall, V. (2001) The impact of deer on woodlands: The effects of browsing and seed dispersal on vegetation structure and composition. *Forestry*, **74**, 209–218.

Girardet, X., Conruyt-Rogeon, G. & Foltête, J.C. (2015) Does regional landscape connectivity influence the location of roe deer roadkill hotspots? *European Journal of Wildlife Research*, **61**, 731–742.

Greenwald, K.R., Petit, L.J. & Waite, T. (2008) Indirect effects of a keystone herbivore elevate local animal diversity. *Journal of Wildlife Management*, **72**, 1318–1321.

Guillera-Arroita, G., Lahoz-Monfort, J.J., Elith, J., Gordon, A., Kujala, H., Lentini, P.E., Mccarthy, M.A., Tingley, R. & Wintle, B.A. (2015) Is my species distribution model fit for purpose? Matching data and models to applications. *Global Ecology and Biogeography*, **24**, 276–292.

Guisan, A. & Thuiller, W. (2005) Predicting species distribution: Offering more than simple habitat models. *Ecology Letters*, **8**, 993–1009.

Guisan, A., Tingley, R., Baumgartner, J.B., Naujokaitis-Lewis, I., Sutcliffe, P.R., Tulloch, A.I.T., Regan, T.J., Brotons, L., Mcdonald-Madden, E., Mantyka-Pringle, C., Martin, T.G., Rhodes, J.R., Maggini, R., Setterfield, S.A., Elith, J., Schwartz, M.W., Wintle, B.A., Broennimann, O., Austin, M., Ferrier, S., Kearney, M.R., Possingham, H.P. & Buckley, Y.M. (2013) Predicting species distributions for conservation decisions. *Ecology Letters*, **16**, 1424–1435.

Guisan, A. & Zimmermann, N.E. (2000) Predictive habitat distribution models in ecology. *Ecological Modelling*, **135**, 147–186.

Heikkinen, R.K., Luoto, M., Araújo, M.B., Virkkala, R., Thuiller, W. & Sykes, M.T. (2006) Methods and uncertainties in bioclimatic envelope modelling under climate change. *Progress in Physical Geography*, **30**, 751–777.

Heinrichs, J.A., Bender, D.J., Gummer, D.L. & Schumaker, N.H. (2010) Assessing critical habitat: Evaluating the relative contribution of habitats to population persistence. *Biological Conservation*, **143**, 2229–2237.

Helle, P. & Ikonen, K. (2016) Wildlife monitoring in Finland - on-line information to game administration, hunters and wider public. *Canadian Journal of Forest Research*, **46**, 1491–1496.

Hewison, A.J.M. (1995) Isozyme Variation in Roe Deer in Relation to Their Population History in Britain. *Journal of Zoology*, **235**, 279–288.

Hewison. (1996) Variation in the fecundity of roe deer in Britain: Effects of age and body weight. *Acta Theriologica*, **41**, 187–198.

Hewison, A.J.M. & Gaillard, J.M. (1996) Birth-sex ratios and local resource competition in roe deer, Capreolus capreolus. *Behav. Ecol.*, **7**, 461–464.

Hewison, A.J.M. & Gaillard, J.M. (2001) Phenotypic quality and senescence affect different components of reproductive output in roe deer. *Journal of Animal Ecology*, **70**, 600–608.

Hinsley, S.A., Hill, R.A., Bellamy, P.E. & Balzter, H. (2006) The application of lidar in woodland bird ecology: climate, canopy structure, and habitat quality. *Photogrammetric Engineering and Remote Sensing*, **72**, 1399–1406.

Holeski, L.M., Mckenzie, S.C., Kruger, E.L., Couture, J.J., Rubert-Nason, K. & Lindroth, R.L. (2016) Phytochemical traits underlie genotypic variation in susceptibility of quaking aspen (Populus tremuloides) to browsing by a keystone forest ungulate. *Journal of Ecology*, **104**, 850–863.

Hothorn, T., M??ller, J., Held, L., M??st, L. & Mysterud, A. (2015) Temporal patterns of deer-vehicle collisions consistent with deer activity pattern and density increase but not general accident risk. *Accident Analysis and Prevention*, **81**, 143–152.

Jepsen, J.U. & Topping, C.J. (2004) Modelling roe deer (Capreolus capreolus) in a gradient of forest fragmentation: behavioural plasticity and choice of cover. *Can. J. Zool*, **82**, 1528–1541.

Jepsen, J.U., Topping, C.J., Odderskær, P. & Andersen, P.N. (2005) Evaluating consequences of land-use strategies on wildlife populations using multiple-species predictive scenarios. *Agriculture, Ecosystems and Environment*, **105**, 581–594.

Jiménez-Valverde, A. (2012) Insights into the area under the receiver operating characteristic curve (AUC) as a discrimination measure in species distribution modelling. *Global Ecology and Biogeography*, **21**, 498–507.

Jiménez-Valverde, A., Lobo, J.M. & Hortal, J. (2008) Not as good as they seem: The importance of concepts in species distribution modelling. *Diversity and Distributions*, **14**, 885–890.

Johnson, D.H. (1980) The Comparison of Usage and Availability Measurements for Evaluating Resource Preference. *Ecology*, **61**, 65–71.

Johnson, C.J. & Gillingham, M.P. (2005) An evaluation of mapped species distribution models used for conservation planning. *Environmental Conservation*, **32**, 117–128.

Kearney, M. & Porter, W. (2009) Mechanistic niche modelling: Combining physiological and

spatial data to predict species' ranges. *Ecology Letters*, **12**, 334–350.

Kearney, M.R., Wintle, B.A. & Porter, W.P. (2010) Correlative and mechanistic models of species distribution provide congruent forecasts under climate change. *Conservation Letters*, **3**, 203–213.

Kie, J.G., Ager, A.A. & Bowyer, R.T. (2005) Landscape-level movements of North American elk (Cervus elaphus): Effects of habitat patch structure and topography. *Landscape Ecology*, **20**, 289–300.

Kjellander, P., Hewison, A.J.M., Liberg, O., Angibault, J.M., Bideau, E. & Cargnelutti, B. (2004) Experimental evidence for density-dependence of home-range size in roe deer (Capreolus capreolus L.): A comparison of two long-term studies. *Oecologia*, **139**, 478–485.

Kramer-Schadt, S., Niedballa, J., Pilgrim, J.D., Schröder, B., Lindenborn, J., Reinfelder, V., Stillfried, M., Heckmann, I., Scharf, A.K., Augeri, D.M., Cheyne, S.M., Hearn, A.J., Ross, J., Macdonald, D.W., Mathai, J., Eaton, J., Marshall, A.J., Semiadi, G., Rustam, R., Bernard, H., Alfred, R., Samejima, H., Duckworth, J.W., Breitenmoser-Wuersten, C., Belant, J.L., Hofer, H. & Wilting, A. (2013) The importance of correcting for sampling bias in MaxEnt species distribution models. *Diversity and Distributions*, **19**, 1366–1379.

Kuemmerle, T., Baskin, L., Leit??o, P.J., Prishchepov, A. V., Thonicke, K. & Radeloff, V.C. (2014) Potential impacts of oil and gas development and climate change on migratory reindeer calving grounds across the Russian Arctic. *Diversity and Distributions*, **20**, 416–429.

Lahoz-Monfort, J.J., Guillera-Arroita, G. & Wintle, B.A. (2014) Imperfect detection impacts the performance of species distribution models. *Global Ecology and Biogeography*, **23**, 504–515.

Langbein, J. (2011) Monitoring reported deer road casualties and related accidents in England to 2010. , 72.

Leblond, M., Dussault, C. & Ouellet, J.P. (2010) What drives fine-scale movements of large herbivores? A case study using moose. *Ecography*, **33**, 1102–1112.

Lefsky, M.A., Cohen, W.B., Parker, G.G. & Harding, D.J. (2002) Lidar Remote Sensing for Ecosystem Studies. *Biogeosciences*, **52**, 19–30.

Lobo, J.M., Jiménez-Valverde, A. & Real, R. (2008) AUC: A misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, **17**, 145–151.

Loison, A., Festa-Bianchet, M., Gaillard, J.-M., Jorgenson, J.T. & Jullien, J.-M. (1999) Age-specific survival in five populations of ungulates:evidence of senescence. *Ecology*, **80**, 2539–2554.

Lone, K., Van Beest, F.M., Mysterud, A., Gobakken, T., Milner, J.M., Ruud, H.-P. & Loe, L.E. (2014) Improving broad scale forage mapping and habitat selection analyses with airborne laser scanning: the case of moose. *Ecosphere*, **5**, 144.

López-Alfaro, C., Estades, C.F., Aldridge, D.K. & Gill, R.M.A. (2012) Individual-based modeling

as a decision tool for the conservation of the endangered huemul deer (Hippocamelus bisulcus) in southern Chile. *Ecological Modelling*, **244**, 104–116.

Lorenzini, R., Garofalo, L., Qin, X., Voloshina, I. & Lovari, S. (2014) Global phylogeography of the genus Capreolus (Artiodactyla: Cervidae), a Palaearctic meso-mammal. *Zoological Journal of the Linnean Society*, **170**, 209–221.

Loro, M., Ortega, E., Arce, R.M. & Geneletti, D. (2016) Assessing landscape resistance to roe deer dispersal using fuzzy set theory and multicriteria analysis: a case study in Central Spain. *Landscape and Ecological Engineering*, **12**, 41–60.

Lovari, S., Serrao, G. & Mori, E. (2017) Woodland features determining home range size of roe deer. *Behavioural Processes*, **140**, 115–120.

Mackenzie, D.I., Nichols, J.D., Lachman, G.B., Droege, S., Andrew, J. & Langtimm, C. a. (2002) Estimating Site Occupancy Rates When Detection Probabilities Are Less Than One. *Ecology*, **83**, 2248–2255.

Margules, C.R. & Pressey, R.L. (2000) Systematic conservation planning. *Nature*, **405**, 243–253.

Marmion, M., Parviainen, M., Luoto, M., Heikkinen, R.K. & Thuiller, W. (2009) Evaluation of Consensus Methods in Predictive Species Distribution Modelling. *Diversity and Distributions*, **15**, 59–69.

Martin, J.L., Stockton, S.A., Allombert, S. & Gaston, A.J. (2010) Top-down and bottom-up consequences of unchecked ungulate browsing on plant and animal diversity in temperate forests: Lessons from a deer introduction. *Biological Invasions*, **12**, 353–371.

McShea, W.J., Underwood, B.H. & Rappole, J.H. (1997) *The Science of Overabundance: Deer Ecology and Population Management*. Smithsonian Institution Press, Washington, DC.

Menke, S.B., Holway, D.A., Fisher, R.N. & Jetz, W. (2009) Characterizing and predicting species distributions across environments and scales: Argentine ant occurrences in the eye of the beholder. *Global Ecology and Biogeography*, **18**, 50–63.

Merow, C., Smith, M.J. & Silander, J.A. (2013) A practical guide to MaxEnt for modeling species' distributions: What it does, and why inputs and settings matter. *Ecography*, **36**, 1058–1069.

Miller, J.R., Turner, M.G., Smithwick, E. a. H., Dent, C.L. & Stanley, E.H. (2004) Spatial Extrapolation: The Science of Predicting Ecological Patterns and Processes. *BioScience*, **54**, 310.

Millington, J.D.A., Walters, M.B., Matonis, M.S. & Liu, J. (2010) Effects of local and regional landscape characteristics on wildlife distribution across managed forests. *Forest Ecology and Management*, **259**, 1102–1110.

Mitchell, F.J.G. & Kirby, K.J. (1990) The impact of large herbivores on the conservation of semi natural woods in the British uplands. *Forestry*, **63**, 333–353.

Van Moorter, B., Gaillard, J.-M., Hewison, A.J.M., Said, S., Coulon, A., Delorme, D., Widmer, O. & Cargnelutti, B. (2008) Evidence for exploration behaviour in young roe deer

(Capreolus capreolus) prior to dispersal. *Ethology Ecology & Evolution*, **20**, 1–15.

Morellet, N., Bonenfant, C., Börger, L., Ossi, F., Cagnacci, F., Heurich, M., Kjellander, P., Linnell, J.D.C., Nicoloso, S., Sustr, P., Urbano, F. & Mysterud, A. (2013) Seasonality, weather and climate affect home range size in roe deer across a wide latitudinal gradient within Europe. *Journal of Animal Ecology*, **82**, 1326–1339.

Morgia, V. La, Bona, F. & Badino, G. (2008) Bayesian modelling procedures for the evaluation of changes in wildlife habitat suitability: A case study of roe deer in the Italian Alps. *Journal of Applied Ecology*, **45**, 863–872.

Moudrý, V. & Šímová, P. (2012) Influence of positional accuracy, sample size and scale on modelling species distributions: a review. *International Journal of Geographical Information Science*, **8816**, 1–13.

Müller, J., Bae, S., Röder, J., Chao, A. & Didham, R.K. (2014) Airborne LiDAR reveals context dependence in the effects of canopy architecture on arthropod diversity. *Forest Ecology and Management*, **312**, 129–137.

Mysterud, A., Larsen, P.K., Ims, R.A. & Østbye, E. (1999) Habitat selection by roe deer and sheep: does habitat ranking reflect resource availability? *Canadian Journal of Zoology*, **77**, 776–783.

Nilsen, E.B., Gaillard, J.M., Andersen, R., Odden, J., Delorme, D., Van Laere, G. & Linnell, J.D.C. (2009) A slow life in hell or a fast life in heaven: Demographic analyses of contrasting roe deer populations. *Journal of Animal Ecology*, **78**, 585–594.

Nilsen, E.B., Linnell, J.D.C. & Andersen, R. (2004) Individual access to preferred habitat affects fitness components in female roe deer Capreolus capreolus. *Journal of Animal Ecology*, **73**, 44–50.

Nishizawa, K., Tatsumi, S., Kitagawa, R. & Mori, A.S. (2016) Deer herbivory affects the functional diversity of forest floor plants via changes in competition-mediated assembly rules. *Ecological Research*, **31**, 569–578.

Noss, R.F., Dobson, A.P., Baldwin, R., Beier, P., Davis, C.R., Dellasala, D.A., Francis, J., Locke, H., Nowak, K., Lopez, R., Reining, C., Trombulak, S.C. & Tabor, G. (2012) Bolder Thinking for Conservation. *Conservation Biology*, **26**, 1–4.

Ossi, F., Gaillard, J.M., Hebblewhite, M. & Cagnacci, F. (2015) Snow sinking depth and forest canopy drive winter resource selection more than supplemental feeding in an alpine population of roe deer. *European Journal of Wildlife Research*, **61**, 111–124.

Palmer, G. (2014) Deer in Britain: population spread and the implications for biodiversity.

Palmer, S.C.F., Broadhead, J.E., Ross, I. & Smith, D.E. (2007) Long-term habitat use and browsing by deer in a Caledonian pinewood. *Forest Ecology and Management*, **242**, 273–280.

Palmer, S.C.F., Coulon, A. & Travis, J.M.J. (2011) Introducing a 'stochastic movement simulator' for estimating habitat connectivity. *Methods in Ecology and Evolution*, **2**, 258–268.

Palmer, S.C.F., Hester, A.J., Elston, D.A., Gordon, I.J. & Hartley, S.E. (2003) the Perils of Having Tasty Neighbors: Grazing Impacts of Large Herbivores At Vegetation Boundaries. *Ecology*, **84**, 2877–2890.

Palminteri, S., Powell, G.V.N., Asner, G.P. & Peres, C.A. (2012) LiDAR measurements of canopy structure predict spatial distribution of a tropical mature forest primate. *Remote Sensing of Environment*, **127**, 98–105.

Pearce, J.L. & Boyce, M.S. (2006) Modelling distribution and abundance with presence-only data. *Journal of Applied Ecology*, **43**, 405–412.

Pearce, J. & Ferrier, S. (2000) Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling*, **133**, 225–245.

Pettorelli, N., Gaillard, J.-M., Duncan, P., Maillard, D., G., van L. & Delorme, D. (2003) Age and density modify the effects of habitat quality on survival and movement of roe deer. *Ecology*, **84**, 3307–3316.

Pettorelli, N., Gaillard, J.M., Duncan, P., Ouellet, J.P. & Van Laere, G. (2001) Population density and small-scale variation in habitat quality affect phenotypic quality in roe deer. *Oecologia*, **128**, 400–405.

Phillips, S.J. (2009) Sample selection bias and presence-only distribution models : implications for background and pseudo-absence data Reference Sample selection bias and presence-only distribution models : implications for background and pseudo-absence data. , **19**, 181–197.

Phillips, S.J., Anderson, R.P., Dudík, M., Schapire, R.E. & Blair, M.E. (2017) Opening the black box: an open-source release of Maxent. *Ecography*, **40**, 887–893.

Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 213–259.

Pompilio, L. & Meriggi, A. (2001) Modelling wild ungulate distribution in alpine habitat: A case study. *Italian Journal of Zoology*, **68**, 281–289.

POST (Parliamentary Office of Science and Technology). (2009) Wild deer. *Wild deer*, 1–4.

Putman, R., Langbein, J., Green, P. & Watson, P. (2011) Identifying threshold densities for wild deer in the UK above which negative impacts may occur. *Mammal Review*, **41**, 175–196.

Putman, R.J., Langbein, J., Hewison, M.A.J. & Sharma, S.K. (1996) Relative roles of density-dependent and density-independent factors in population dynamics of British deer. *Mammal Review*, **26**, 81–101.

Putman, R.J. & Moore, N.P. (1998) Impact of deer in lowland Britain on agriculture, forestry and conservation habitats. *Mammal Review*, **28**, 141–164.

Putman, R., Watson, P. & Langbein, J. (2011) Assessing deer densities and impacts at the appropriate level for management: A review of methodologies for use beyond the site scale. *Mammal Review*, **41**, 197–219.

Radeloff, V. (1996) Dynamic modelling of a roe deer population in a GIS. *Zeitschrift fur*

*Jagdwissenschaft*, **42**, 203–213.

Radeloff, V.C., Pidgeon, A.M. & Hostert, P. (1999) Habitat and population modelling of roe deer using an interactive geographic information system. *Ecological Modelling*, **114**, 287–304.

Reimoser, S., Partl, E., Reimoser, F. & Vospernik, S. (2009) Roe-deer habitat suitability and predisposition of forest to browsing damage in its dependence on forest growth-Model sensitivity in an alpine forest region. *Ecological Modelling*, **220**, 2231–2243.

Richard, E., Calenge, C., Saïd, S., Hamann, J.-L. & Gaillard, J.-M. (2013) Studying spatial interactions between sympatric populations of large herbivores: a null model approach. *Ecography*, **36**, 157–165.

Rodríguez-Rey, M., Jiménez-Valverde, A. & Acevedo, P. (2013) Species distribution models predict range expansion better than chance but not better than a simple dispersal model. *Ecological Modelling*, **256**, 1–5.

Roura-Pascual, N., Brotons, L., Peterson, A.T. & Thuiller, W. (2009) Consensual predictions of potential distributional areas for invasive species: A case study of Argentine ants in the Iberian Peninsula. *Biological Invasions*, **11**, 1017–1031.

Royle, J.A., Chandler, R.B., Yackulic, C. & Nichols, J.D. (2012) Likelihood analysis of species occurrence probability from presence-only data for modelling species distributions. *Methods in Ecology and Evolution*, **3**, 545–554.

Ruckstuhl, K.E. & Festa-Bianchet, M. (2001) Group Choice by Subadult Bighorn Rams : Trade-offs between Foraging Efficiency and Predator Avoidance. *Ethology*, **107**, 161–172.

Sæther, B.-E. & Heim, M. (1993) Ecological correlates of individual variation in age at maturity in female moose (Alces alces): the effects of environmental variability. *The Journal of Animal Ecology*, **62**, 482–489.

Saïd, S., Gaillard, J.-M., Duncan, P., Guillon, N., Guillon, N., Servanty, S., Pellerin, M., Lefeuvre, K., Martin, C. & Van Laere, G. (2005) Ecological correlates of home-range size in spring–summer for female roe deer ( Capreolus capreolus ) in a deciduous woodland. *Journal of Zoology*, **267**, 301.

Semeniuk, C.A.D., Musiani, M., Hebblewhite, M., Grindal, S. & Marceau, D.J. (2012) Incorporating behavioral-ecological strategies in pattern-oriented modeling of caribou habitat use in a highly industrialized landscape. *Ecological Modelling*, **243**, 18–32.

Smith, A.B. (2013) On evaluating species distribution models with random background sites in place of absences when test presences disproportionately sample suitable habitat. *Diversity and Distributions*, **19**, 867–872.

Sonnichsen, L., Borowik, T., Podgorski, T., Plis, K., Berger, A. & Jodrzejewska, B. (2017) Survival rates and causes of mortality of roe deer Capreolus capreolus in a rural landscape, eastern Poland. *Mammal Research*, **62**, 141–147.

Stache, A., Heller, E., Hothorn, T. & Heurich, M. (2013) Activity patterns of European roe deer (Capreolus capreolus) are strongly influenced by individual behaviour. *Folia Zoologica*, **62**, 67–75.

Stephens, P.A., Zaumyslova, O.Y., Miquelle, D.G., Myslenkov, A.I. & Hayward, G.D. (2006) Estimating population density from indirect sign: Track counts and the Formozov-Malyshev-Pereleshin formula. *Animal Conservation*, **9**, 339–348.

Sun, Y., Wang, T., Skidmore, A.K., Palmer, S.C.F., Ye, X., Ding, C. & Wang, Q. (2016) Predicting and understanding spatio-temporal dynamics of species recovery: implications for Asian crested ibis Nipponia nippon conservation in China. *Diversity and Distributions*, **22**, 893–904.

Tablado, Z., Revilla, E., Dubray, D., Saïd, S., Maillard, D., Loison, A. & McArthur, C. (2016) From steps to home range formation: species-specific movement upscaling among sympatric ungulates. *Functional Ecology*, **30**, 1384–1396.

Tews, J., Ferguson, M.A.D. & Fahrig, L. (2007) Potential net effects of climate change on High Arctic Peary caribou: Lessons from a spatially explicit simulation model. *Ecological Modelling*, **207**, 85–98.

The Deer Initiative. (2007) Deer on our roads: counting the cost.

Tittensor, D.P., Walpole, M., Hill, S.L.L., Boyce, D.G., Britten, G.L., Burgess, N.D., Butchart, S.H.M., Leadley, P.W., Regan, E.C., Alkemade, R., Baumung, R., Bellard, C., Bouwman, L., Bowles-newark, N.J., Chenery, A.M. & Cheung, W.W.L. (2014) Biodiversity Targets. *Science*, **346**, 241–245.

Toigo, C., Gaillard, J., Laere, G. Van, Hewison, M. & Morellet, N. (2006) How Does Environmental Variation Influence Body Mass , Body Size , and Body Condition? Roe Deer as a Case Study. *Ecography*, **29**, 301–308.

Torres, R.T.., Carvalho, J.., Fonseca, C.., Serrano, E.. & López-Martín, J.M. (2016) Longterm assessment of roe deer reintroductions in North-East Spain: A case of success. *Mammalian Biology*, **81**, 415–422.

Torres, R.T., Virgós, E., Panzacchi, M., Linnell, J.D.C. & Fonseca, C. (2012a) Life at the edge: Roe deer occurrence at the opposite ends of their geographical distribution, Norway and Portugal. *Mammalian Biology*, **77**, 140–146.

Torres, R.T., Virgós, E., Santos, J., Linnell, J.D.C. & Fonseca, C. (2012b) Habitat use by sympatric red and roe deer in a Mediterranean ecosystem. *Animal Biology*, **62**, 351–366.

Tsaparis, D., Katsanevakis, S., Stamouli, C. & Legakis, A. (2008) Estimation of roe deerCapreolus capreolus and mouflonOvis aries densities, abundance and habitat use in a mountainous Mediterranean area. *Acta theriologica*, **53**, 87–94.

Tufto, J., Andersen, R. & Linnell, J. (1996) Habitat Use and Ecological Correlates of Home Range Size in a Small Cervid: The Roe Deer. *The Journal of Animal Ecology*, **65**, 715.

Tulloch, A.I.T., Sutcliffe, P., Naujokaitis-Lewis, I., Tingley, R., Brotons, L., Ferraz, K.M.P.M.B., Possingham, H., Guisan, A. & Rhodes, J.R. (2016) Conservation planners tend to ignore improved accuracy of modelled species distributions to focus on multiple threats and ecological processes. *Biological Conservation*, **199**, 157–171.

Turner, M.G., Arthaud, G.J., Engstrom, R.T., Hejl, S.J., Applications, S.E., Feb, N., Turner,

M.G., Arthaud, G.J. & Engstrom, R.T. (1995) Usefulness of Spatially Explicit Population Models in Land Management Liu , Susan Loeb and Kevin McKelvey Published by : Wiley Stable URL : http://www.jstor.org/stable/1942046 REFERENCES Linked references are available on JSTOR for this article : You may. , **5**, 12–16.

Valente, A.M., Marques, T.A., Fonseca, C. & Torres, R.T. (2016a) A new insight for monitoring ungulates: density surface modelling of roe deer in a Mediterranean habitat. *European Journal of Wildlife Research*, **62**, 577–587.

Valente, A.M., Marques, T.A., Fonseca, C. & Torres, R.T. (2016b) A new insight for monitoring ungulates: density surface modelling of roe deer in a Mediterranean habitat. *European Journal of Wildlife Research*, **62**, 577–587.

Vanpé, C., Kjellander, P., Galan, M., Cosson, J.F., Aulagnier, S., Liberg, O. & Hewison, A.J.M. (2008) Mating system, sexual dimorphism, and the opportunity for sexual selection in a territorial ungulate. *Behavioral Ecology*, **19**, 309–316.

Venables, W.N. & Dichmont, C.M. (2004) GLMs, GAMs and GLMMs: An overview of theory for applications in fisheries research. *Fisheries Research*, **70**, 319–337.

Vierling, K.T., Vierling, L.A., Gould, W.A., Martinuzzi, S. & Clawges, R.M. (2008) Lidar: Shedding new light on habitat characterization and modeling. *Frontiers in Ecology and the Environment*, **6**, 90–98.

Vincent, J.P., Bideau, E., Hewison, A.J.M. & Angibault, J.M. (1995) The influence of increasing density on body weight, kid production, home range and winter grouping in roe deer (Capreolus capreolus). *Journal of Zoology*, **236**, 371–382.

Wäber, K. & Dolman, P.M. (2015) Deer abundance estimation at landscape-scales in heterogeneous forests. *Basic and Applied Ecology*, **16**, 610–620.

Wäber, K., Spencer, J. & Dolman, P.M. (2013a) Achieving landscape-scale deer management for biodiversity conservation: The need to consider sources and sinks. *Journal of Wildlife Management*, **77**, 726–736.

Wäber, K., Spencer, J. & Dolman, P.M. (2013b) Achieving landscape-scale deer management for biodiversity conservation: The need to consider sources and sinks. *Journal of Wildlife Management*, **77**, 726–736.

Wahlström, L.K. (1994) The significance of male-male aggression for yearling dispersal in roe deer (Capreolus capreolus). *Behavioral Ecology and Sociobiology*, **35**, 409–412.

Wahlström, L.K. & Kjellander, P. (1995) Ideal free distribution and natal dispersal in female roe deer. *Oecologia*, **103**, 302–308.

Wahlström, L.K. & Liberg, O. (1995) Patterns of dispersal and seasonal migration in roe deer (Capreolus capreolus). *Journal of Zoology*, **235**, 455–467.

Ward, A.I. (2005) Expanding ranges of wild and feral deer in Great Britain. *Mammal Review*, **35**, 165–173.

Ward, A.I., White, P.C.L., Smith, A. & Critchley, C.H. (2004) Modelling the cost of roe deer browsing damage to forestry. *Forest Ecology and Management*, **191**, 301–310.

Ward, A.I., White, P.C.L., Walker, N.J. & Critchley, C.H. (2008) Conifer leader browsing by roe deer in English upland forests: Effects of deer density and understorey vegetation. *Forest Ecology and Management*, **256**, 1333–1338.

Weisberg, P.J., Thompson Hobbs, N., Ellis, J.E. & Coughenour, M.B. (2002) An ecosystem approach to population management of ungulates. *Journal of Environmental Management*, **65**, 181–197.

Welsh Government. (2017) Action Plan: For wild deer management in Wales.

West, N.E. (1993) Biodiversity of Rangelands. *Journal of Range Management*, **46**, 2–13.

White, M.A. (2012) Long-term effects of deer browsing: Composition, structure and productivity in a northeastern Minnesota old-growth forest. *Forest Ecology and Management*, **269**, 222–228.

White, P. & Smart, J. (2004) Economic impacts of wild deer in the east of England. *… Park, Tinkley Lane …*, 1–17.

Wiemers, D.W., Fulbright, T.E., Wester, D.B., Ortega-S, J.A., Rasmussen, G.A., Hewitt, D.G. & Hellickson, M.W. (2014) Role of Thermal Environment in Habitat Selection by Male White-Tailed Deer during Summer in Texas, USA. *Wildlife Biology*, **20**, 47–56.

Williams, D.M., Dechen Quinn, A.C. & Porter, W.F. (2014) Informing disease models with temporal and spatial contact structure among GPS-collared individuals in wild populations. *PLoS ONE*, **9**.

Wilson, K.A., Meuaard, E., Drummond, S., Grantham, H.S., Boitani, L., Catullo, G., Christie, L., Dennis, R., Dutton, I., Falcucci, A., Maiorano, L., Possingham, H.P., Rondinini, C., Turner, W.R., Venter, O. & Watts, M. (2010) Conserving biodiversity in production landscapes. *Ecological Applications*, **20**, 1721–1732.

With, K.A. (2002) The landscape ecology of invasive spread. *Conservation Biology*, **16**, 1192–1203.

Wu, W., Li, Y. & Hu, Y. (2016) Simulation of potential habitat overlap between red deer (Cervus elaphus) and roe deer (Capreolus capreolus ) in northeastern China. *PeerJ*, **4**, e1756.

Zaragozí, B., Belda, A., Giménez, P., Navarro, J.T. & Bonet, A. (2015) Advances in camera trap data management tools: Towards collaborative development and integration with GIS. *Ecological Informatics*, **30**, 6–11.

Zhang, W.-Q. & Zhang, M.-H. (2012) Phylogeny and evolution of Cervidae based on complete mitochondrial genomes. *Genetics and molecular research : GMR*, **11**, 628–35.

**APPENDIX**

**(1) Systematic literature review**

***Search terms (Web of Science)***: ("roe deer" OR "capreolus capreolus) AND ("habitat suitability" OR "resource selection" OR "habitat selection" OR "resource use" OR "habitat use" OR "land use" OR "species distribution" OR "population dynamics" OR "spatial dynamics") AND (model*)

***Document type***: Article

***Inclusion criteria***: Studies must use a spatially explicit empirical model to;

(i) describe and/or predict the distribution of roe deer (*Capreolus capreolus*)

AND/OR

(ii) quantify AND map habitat suitability or probability of occurrence for roe deer

***Number of studies identified using search terms:*** 162

***Number of studies that met all inclusion criteria:*** 23 (plus one unpublished thesis: Palmer 2014)

Table 1A. Classification of model types in the systematic review of roe deer (*Capreolus capreolus*) species distribution modelling literature.

| Model classification | Model types |
|---|---|
| Regression | Generalized linear models (GLMs), generalized linear mixed models (GLMMs), generalized addictive models (GAMs), generalized additive mixed models (GAMMs) |
| Individual | Simulation based, mechanistic models |
| Machine learning | Boosted regression trees (BRTs), genetic algorithm for rule set production (GARP), maximum entropy (MaxEnt), artificial neural networks (ANNs) |
| Selection function | Step selection functions, resource selection functions |

Table 2A. Classification of data types in the systematic review of roe deer (*Capreolus capreolus*) species distribution modelling literature.

| Data classification | Data source |
|---|---|
| Presence-only (PO) | Incidental sights, museum records, hunting bag records |
| Presence-absence (PA) | Systematic surveys, pellet counts, camera trap surveys, distance sampling |
| Detection (DET) | Repeated surveys that allowed for the calculation of detection probability |
| Movement (MOV) | Global Positioning System (GPS) or Very High Frequency (VHF) radio collars |

Table 3A. Summary of the outcomes of a systematic review of roe deer (*Capreolus capreolus*) species distribution modelling literature (n = 24, 1996-2017). DET = detection, PA = presence-absence, PO = presence-only, MOV = movement data obtained from GPS or VHF radio collars (note: details of one study, Torres *et al.* (2012), were not accessible). Occurrence resolution grade: 1, = ≤0.01 $km^2$; 2, = 0.01 – 1 $km^2$; 3, = 1 – 100 $km^2$; 4, = >100 $km^2$.

| Reference | Model type | Data Type | Study area ($km^2$) | Occurrence resolution grade (1-4) |
|---|---|---|---|---|
| (Tsaparis *et al.* 2008) | Regression | DET | 3 | 1 |
| (Torres *et al.* 2012b) | Regression | PA | 10 | 1 |
| (Le Corre *et al.* 2008) | Regression | MOV | 26 | 1 |
| (Richard *et al.* 2013) | Regression | MOV | 27 | 1 |
| (Coulon *et al.* 2008) | Selection function | MOV | 80 | 1 |
| (Palmer *et al.* 2007) | Regression | PA | 117 | 1 |
| (Stache *et al.* 2013) | Regression | MOV | 242 | 1 |
| (Ewald *et al.* 2014) | Selection function | MOV | 244 | 1 |
| (Ossi *et al.* 2015) | Selection function | MOV | 400 | 1 |
| (Morgia, Bona & Badino 2008) | Regression | PA | 900 | 1 |
| (Wu, Li & Hu 2016) | Machine learning | PA | 2058 | 1 |
| (Radeloff 1996) | Regression | PO | 7 | 2 |
| (Reimoser *et al.* 2009) | Regression | PA | 19 | 2 |
| (Mysterud *et al.* 1999) | Selection function | MOV | 250 | 2 |
| (Pompilio & Meriggi 2001) | Regression | PO | 725 | 2 |
| (Bouyer *et al.* 2015) | Regression | PA | 15550 | 2 |

| | | | | |
|---|---|---|---|---|
| (Acevedo *et al.* 2011) | Regression | PO | 87268 | 2 |
| (Torres *et al.* 2016) | Machine learning | PO | 32114 | 3 |
| (Acevedo *et al.* 2017) | Regression | PA | 47669 | 3 |
| (Acevedo *et al.* 2010a) | Regression | PO | 228754 | 3 |
| (Rodríguez-Rey, Jiménez-Valverde & Acevedo 2013) | Ensemble - Regression and Machine learning | PO | 228754 | 3 |
| (Palmer 2014) | Ensemble - Regression, Individual and machine learning | PO/PA/DET | 228754 | 3 |
| (Borowik, Cornulier & Jedrzejewska 2013) | Regression | PO | 311904 | 4 |

**(2) MaxEnt model development**

a) Data preparation

***Species***

- Roe deer records at a 100 x 100 m resolution were downloaded from the National Biodiversity Network Atlas (formerly NBN Gateway)
- Years range: 1960 to 2015
- n = 5,652
- Removed duplicates, n = 4,537
- England and Scotland, (n = 4,516)
- Removed points within 100 m of roads (n = 3,948)
- England and Scotland (n = 3,932)
- Removed points with incomplete environmental data (n = 3,843)
- Presences in Wales (n = 20) were augmented by data obtained from the Biodiversity Information Service (n = 124), Wildlife Trust of South West Wales (n = 4) and the North Wales Environmental Information Service (n = 32)
- Total presences in Wales (n = 180)
- Removed duplicates (n = 125)
- Selected points for the mainland of Wales only (n = 120)
- Removed points within 100m of roads (n = 106)

***Land cover***

- Downloaded Centre of Ecology and Hydrology (CEH) Land Cover Map 2015 (25 x 25 m resolution, digimap.ac.uk)
- Reclassified into 10 aggregate classes
- Resampled to 100 m resolution to match that of the species data (0.01 $km^2$, majority value taken from 4x4 cell rectangle)
- Used land cover map used as a base map to define the cell size and extent of all other layers (as this dataset is the highest resolution environmental data used)

***Land cover proportion***

- Original data source: CEH Land Cover Map 2015
- Calculated the proportionate cover of specific land cover types within an area
- Roe deer are mobile with home ranges approximately 0.78 $km^2$.
- As home ranges are estimated to be approximately 0.78 $km^2$ (Morellet *et al*., 2013), cover of each land type was calculated as a proportion (0-1) of the area within a 500 m radius of each cell (0.785 $km^2$ area)
- Cover types included; (1) urban and suburban (low resources, poor cover, high disturbance - likely to be avoided), (2) woodland (highest resources, highest degree of cover - most likely to be selected) and (3) non-woodland forage (mostly grassland

and arable, high resources for forage but low cover and high disturbance – likely to be weakly favoured)

***Roads***

- Shapefiles for Motorways, A roads and B roads were obtained from the University shared server (U drive, original source: Meridian Transport OS map, available from digimap.ac.uk)
- DVCs are a prominent source of roe deer fatalities
- Distance to roads produced spurious predictions from the model most likely because the species data is biased towards areas that are more accessible to observers. Additionally, roe deer are well able to occupy a fragmented landscape and are likely to have frequently been observed in close proximity to a road as coverage in the UK is so extensive
- Road *cover* was therefore used to differentiate between locations that are near to roads and those that are near to dense road networks (dense areas least likely to be occupied due to noise/light disturbance and mortality risk)
- As for percentage covers, the cover of roads ($m^2$) was calculated within a 500 m radius of each cell
- Species data points within 100 m of a road were removed from the model as it is likely that these originated from DVCs

***Terrain***

- Downloaded Ordnance Survey Terrain 50 data (0.0025 $km^2$ resolution, digimap.ac.uk)
- Altitude and terrain influence the movement of most terrestrial species
- Assumed that high altitude/mountainous areas are likely to be avoided
- Used bilinear resampling to increase the cell size to 0.01 $km^2$
- Produced additional layers for (1) aspect (2) slope and (3) rugosity

***Climate***

- Downloaded data for 19 bioclimatic variables (30 arc secs resolution (approximately 1 $km^2$, worldclim.org)
- Used bilinear resampling to decrease the cell size to 0.01 $km^2$ (necessary for the model but does not improve data resolution)
- Variables included:

BIO1 = Annual Mean Temperature
BIO2 = Mean Diurnal Range (Mean of monthly (max. temp – min. temp))
BIO3 = Isothermality (BIO2/BIO7) (* 100)
BIO4 = Temperature Seasonality (standard deviation *100)
BIO5 = Max. Temperature of Warmest Month

BIO6 = Min. Temperature of Coldest Month
BIO7 = Temperature Annual Range (BIO5-BIO6)
BIO8 = Mean Temperature of Wettest Quarter
BIO9 = Mean Temperature of Driest Quarter
BIO10 = Mean Temperature of Warmest Quarter
BIO11 = Mean Temperature of Coldest Quarter
BIO12 = Annual Precipitation
BIO13 = Precipitation of Wettest Month
BIO14 = Precipitation of Driest Month
BIO15 = Precipitation Seasonality (Coefficient of Variation)
BIO16 = Precipitation of Wettest Quarter
BIO17 = Precipitation of Driest Quarter
BIO18 = Precipitation of Warmest Quarter
BIO19 = Precipitation of Coldest Quarter

***Omitted Data***

- *Other deer species* - Resource use between different species is likely to be similar (NOTE: this could be used to produce an estimate of sampling effort and account for spatial bias in future models). A visual examination of the data and some preliminary trials yielded no evidence to suggest that roe deer are displaced by any other species. As the available data is limited to "presence-only" observations, we chose to omit data for other deer species to reduce error from inconsistent sampling. It is unlikely that any of the six deer species in the UK are truly in equilibrium with their environment and the data show a strong spatial bias towards south-east England where the species was introduced/reintroduced
- *Hunter effort* – No consistent records are available for the entire study area. The National Gamebag Census was considered but its records for Wales are very limited.

<u>b) Variable selection</u>

***Checking for collinearity***

- Pairwise Pearson's correlation coefficients were calculated and correlation plots were created in R (https://www.r-project.org/) from data extracted at points of recorded presences and from a random background sample
- Any variables highly correlated ($R^2 \geq 0.6$) were identified for potential removal
- VIF analysis was performed using a combination of the presence and background datasets
- Any variables highly correlated (VIF>2) were identified for potential removal
- VIF analysis was used to support the collinearity analysis but was not the primary tool for variable selection

***Assessing variable contribution***

- Variables were grouped into the following bins and a MaxEnt model was developed for each:
  1) All
  2) Land cover (including proportionate cover)
  3) Roads
  4) Terrain
  5) Climate (and altitude*)
- A variable contribution plot was produced for each model
- Any variable with a low (<5%) contribution in the individual (2-5) and collective (1) models was identified for potential removal

**altitude was included as it was predicted that it would covary with many of the climate variables*

***Selection***

- MaxEnt is less sensitive to high collinearity than traditional statistical models (Elith *et al*. 2011) but it is recommended that highly correlated predictors are removed to aid interpretation (Merow *et al.* 2013)
- Each variable was first assessed for collinearity
- Within each bin, if two variables were highly correlated in both the presence samples and background samples, the variable with the lowest contribution was removed
- If two variables were highly correlated in only one of either the presence or background samples then each variable was assessed for contribution
- Correlated variables were removed if they contributed < 5% to both the binned model and collective model
- Response curves and VIF analysis were also carried out to support the results of the collinearity analysis

***Land cover***

- Seven land cover variables were considered (NOTE: Density = proportion of cover within a 500 m radius of each cell, Distance = distance (m) from each land cover type):

FORAGE DENSITY

FORAGE DISTANCE

LAND COVER

URBAN DENSITY

URBAN DISTANCE

WOOD DENSITY

WOOD DISTANCE

Table 4A. Pairwise Pearson's correlation coefficients calculated for land cover variables using data extracted at points of recorded presences (n = 3,843)

| | forage_density | forage_distance | landcover | urban_density | wood_density | wood_distance |
|---|---|---|---|---|---|---|
| forage_density | 1.0 | -0.6 | 0.0 | -0.3 | -0.7 | 0.3 |
| forage_distance | **-0.6** | 1.0 | 0.0 | 0.0 | 0.4 | 0.0 |
| landcover | 0.0 | 0.0 | 1.0 | 0.3 | -0.4 | 0.3 |
| urban_density | -0.3 | 0.0 | 0.3 | 1.0 | -0.1 | 0.0 |
| wood_density | **-0.7** | 0.4 | -0.4 | -0.1 | 1.0 | -0.5 |
| wood_distance | 0.3 | 0.0 | 0.3 | 0.0 | -0.5 | 1.0 |

Table 5A. Pairwise Pearson's correlation coefficients calculated for land cover variables using data extracted from randomly generated points (n = 10,000)

| | forage_density | forage_distance | landcover | urban_density | wood_density | wood_distance |
|---|---|---|---|---|---|---|
| forage_density | 1.0 | -0.6 | -0.4 | -0.3 | -0.4 | -0.1 |
| forage_distance | **-0.6** | 1.0 | 0.2 | 0.0 | 0.1 | 0.3 |
| landcover | -0.4 | 0.2 | 1.0 | 0.6 | -0.4 | 0.3 |
| urban_density | -0.3 | 0.0 | **0.6** | 1.0 | -0.1 | 0.0 |
| wood_density | -0.4 | 0.1 | -0.4 | -0.1 | 1.0 | -0.4 |
| wood_distance | -0.1 | 0.3 | 0.3 | 0.0 | -0.4 | 1.0 |

Table 6A. Summary of variance inflation factor (VIF) analysis for land cover variables

| Variable | GVIF | Df | GVIF^(1/(2*Df)) |
|---|---|---|---|
| forage_density | 9.35 | 1 | **3.06** |
| forage_distance | 2.15 | 1 | 1.47 |
| landcover | 29.75 | 10 | 1.18 |
| urban_density | 4.30 | 1 | **2.07** |
| wood_density | 5.79 | 1 | **2.41** |
| wood_distance | 1.63 | 1 | 1.28 |

- FORAGE DENSITY was highly correlated with FORAGE DISTANCE in both samples
- FORAGE DISTANCE contributed less than FORAGE DENSITY in both models and was removed
- URBAN DENSITY and LANDCOVER was highly correlated in the background sample only
- URBAN DENSITY contributed < 5% to both models and was removed
- URBAN DISTANCE was removed following inspection of response curves. The model predicted a negative relationship between URBAN DISTANCE and relative occurrence

rate. However, the same relationship was observed for URBAN DENSITY. England and Scotland has an extensive coverage of small urban and suburban areas that are likely to be in close proximity to suitable roe deer habitat. The negative relationship observed for URBAN DENSITY, however, suggests that urban areas are not being preferentially selected. It was assumed that the relationship observed for URBAN DISTANCE was most likely due to sampling bias towards accessible locations (NOTE: this should be explored in future analysis, other deer records could be used)

***Roads***

- Four road variables were considered (for each road type= presence within a 500 m radius was modelled)

MOTORWAYS

A ROADS

B ROADS

ROAD DENSITY

Table 7A. Pairwise Pearson's correlation coefficients calculated for road variables using data extracted at points of recorded presences (n = 3,843)

| | A_roads | B_roads | Motorways | road_density |
|---|---|---|---|---|
| A_roads | 1.0 | 0.0 | 0.1 | 0.5 |
| B_roads | 0.0 | 1.0 | 0.0 | 0.5 |
| Motorways | 0.1 | 0.0 | 1.0 | 0.3 |
| road_density | 0.5 | 0.5 | 0.3 | 1.0 |

Table 8A. Pairwise Pearson's correlation coefficients calculated for road variables using data extracted from randomly generated points (n = 10,000)

| | A_roads | B_roads | Motorways | road_density |
|---|---|---|---|---|
| A_roads | 1.0 | 0.1 | 0.1 | 0.6 |
| B_roads | 0.1 | 1.0 | 0.1 | 0.5 |
| Motorways | 0.1 | 0.1 | 1.0 | 0.2 |
| road_density | **0.6** | 0.5 | 0.2 | 1.0 |

Table 9A. Summary of variance inflation factor (VIF) analysis for road variables

| Variable | GVIF |
|---|---|
| A_roads | 1.69 |
| B_roads | 1.52 |
| Motorways | 1.08 |
| road_density | **2.41** |

- ROAD DENSITY was correlated with A ROADS in the background sample only
- A ROADS contributed less than ROAD DENSITY in both models and was removed
- VIF analysis indicated high covariation for ROAD DENSITY
- ROAD DENSITY contributed >50% to the binned model and was the only variable retained

***Terrain***

- Four terrain variables were considered:

ALTITUDE

ASPECT

SLOPE

RUGOSITY

Table 10A. Pairwise Pearson's correlation coefficients calculated for terrain variables using data extracted at points of recorded presences (n = 3,843)

| | altitude | aspect | rugosity | slope |
|---|---|---|---|---|
| altitude | 1.0 | 0.0 | -0.1 | 0.3 |
| aspect | 0.0 | 1.0 | 0.0 | 0.0 |
| rugosity | -0.1 | 0.0 | 1.0 | 0.0 |
| slope | 0.3 | 0.0 | 0.0 | 1.0 |

Table 11A. Pairwise Pearson's correlation coefficients calculated for terrain variables using data extracted from randomly generated points (n = 10,000)

| | altitude | aspect | rugosity | slope |
|---|---|---|---|---|
| altitude | 1.0 | 0.0 | -0.1 | 0.6 |
| aspect | 0.0 | 1.0 | 0.0 | 0.0 |
| rugosity | -0.1 | 0.0 | 1.0 | 0.0 |
| slope | **0.6** | 0.0 | 0.0 | 1.0 |

Table 12A. Summary of variance inflation factor (VIF) analysis for terrain variables

| Variable | GVIF |
|---|---|
| altitude | 1.47 |
| aspect | 1.00 |
| rugosity | 1.01 |
| slope | 1.46 |

- ALTITUDE was correlated with SLOPE in the background sample only
- SLOPE contributed < 5% to the collective model only
- ASPECT and RUGOSITY contributed <5% to both models
- ALTITUDE contributed >80% to the binned model and was the only variable retained

***Climate***

- Eight climate variables were selected (a priori) to develop the model:

BIO1 = Annual Mean Temperature
BIO6 = Min Temperature of Coldest Month
BIO7 = Temperature Annual Range (BIO5-BIO6)
BIO11 = Mean Temperature of Coldest Quarter
BIO12 = Annual Precipitation
BIO13 = Precipitation of Wettest Month
BIO15 = Precipitation Seasonality (Coefficient of Variation)
BIO16 = Precipitation of Wettest Quarter

Table 13A. Pairwise Pearson's correlation coefficients calculated for climate variables using data extracted at points of recorded presences (n = 3,843)

| | altitude | annual_ precip | mean_ temp | precip_ seas | precip_ wetmon | precip_ wq | temp_ cm | temp_ cq | temp_ range |
|---|---|---|---|---|---|---|---|---|---|
| altitude | 1.0 | 0.3 | -0.6 | 0.1 | 0.2 | 0.2 | -0.6 | -0.7 | 0.0 |
| annual_precip | 0.3 | 1.0 | -0.4 | 0.9 | 1.0 | 1.0 | -0.2 | -0.3 | -0.5 |
| mean_temp | **-0.6** | -0.4 | 1.0 | -0.3 | -0.4 | -0.4 | 0.8 | 0.9 | 0.4 |
| precip_seas | 0.1 | **0.9** | -0.3 | 1.0 | 0.9 | 0.9 | 0.0 | -0.1 | -0.5 |
| precip_wetmon | 0.2 | **1.0** | -0.4 | **0.9** | 1.0 | 1.0 | -0.1 | -0.2 | -0.5 |
| precip_wq | 0.2 | **1.0** | -0.4 | **0.9** | **1.0** | 1.0 | -0.1 | -0.2 | -0.5 |
| temp_cm | **-0.6** | -0.2 | **0.8** | 0.0 | -0.1 | -0.1 | 1.0 | 1.0 | -0.1 |
| temp_cq | **-0.7** | -0.3 | **0.9** | -0.1 | -0.2 | -0.2 | **1.0** | 1.0 | 0.1 |
| temp_range | 0.0 | -0.5 | 0.4 | -0.5 | -0.5 | -0.5 | -0.1 | 0.1 | 1.0 |

Table 14A. Pairwise Pearson's correlation coefficients calculated for climate variables using data extracted from randomly generated points (n = 10,000)

| | altitude | annual_ precip | mean_ temp | precip_ seas | precip_ wetmon | precip_ wq | temp_ cm | temp_ cq | temp_ range |
|---|---|---|---|---|---|---|---|---|---|
| altitude | 1.0 | 0.6 | -0.9 | 0.4 | 0.6 | 0.6 | -0.8 | -0.9 | -0.1 |
| annual_precip | **0.6** | 1.0 | -0.7 | 0.8 | 1.0 | 1.0 | -0.5 | -0.6 | -0.5 |
| mean_temp | **-0.9** | **-0.7** | 1.0 | -0.5 | -0.6 | -0.6 | 0.9 | 1.0 | 0.3 |
| precip_seas | 0.4 | **0.8** | -0.5 | 1.0 | 0.9 | 0.9 | -0.3 | -0.3 | -0.6 |
| precip_wetmon | **0.6** | **1.0** | **-0.6** | **0.9** | 1.0 | 1.0 | -0.5 | -0.5 | -0.5 |
| precip_wq | **0.6** | **1.0** | **-0.6** | **0.9** | **1.0** | 1.0 | -0.5 | -0.5 | -0.5 |
| temp_cm | **-0.8** | -0.5 | **0.9** | -0.3 | -0.5 | -0.5 | 1.0 | 1.0 | -0.1 |
| temp_cq | **-0.9** | **-0.6** | **1.0** | -0.3 | -0.5 | -0.5 | **1.0** | 1.0 | 0.0 |
| temp_range | -0.1 | -0.5 | 0.3 | **-0.6** | -0.5 | -0.5 | -0.1 | 0.0 | 1.0 |

Table 15A. Summary of variance inflation factor (VIF) analysis for climate variables

| Variable | GVIF |
|---|---|
| Altitude | **5.50** |
| annual_precip | **473.43** |
| mean_temp | **122.00** |
| precip_seas | **13.13** |
| precip_wetmon | **289.69** |
| precip_wq | **1008.86** |
| temp_cm | **87.23** |
| temp_cq | **97.71** |
| temp_range | **17.98** |

- *It was predicted that ALTITUDE would be highly correlated with several climate variables and was also included for analysis
- BIO1, BIO6 and BIO11 were highly correlated with altitude in both samples and were removed. The suitability map produced by the correlative model was intended to be used as a cost surface in the individual-based model. Consequently, altitude was retained in favour of climate variables because it was considered more likely to influence movement behaviour during dispersal. Altitude was included in the model as a surrogate for the correlated climate variables as well as non-woodland cover. Non-woodland forage was included as a land cover variable but some habitat types within this category may offer varying degrees of cover (i.e. shelter from detection and harsh weather) which is likely to strongly influence suitability. It was assumed that altitude is negatively correlated with non-woodland cover and could therefore be used as a surrogate. It is recognised that such an assumption should be validated, possibly using LiDAR data, in future studies

- BIO7 and BIO15 were the only climate variables not highly correlated with altitude in either sample. They were also the two highest contributing climate variables in both models
- BIO12, BIO13 and BIO16 were highly correlated with BIO15 in both models and were removed
- BIO7 was correlated with BIO15 in the background sample only

***Final selected variables***

FORAGE DENSITY

LAND COVER

WOOD DENSITY

WOOD DISTANCE

ROAD DENSITY

ALTITUDE

BIO7 = Temperature Annual Range (BIO5-BIO6)

BIO15 = Precipitation Seasonality (Coefficient of Variation)

Table 16A. Pairwise Pearson's correlation coefficients calculated for selected variables using data extracted at points of recorded presences (n = 3,843)

| | altitude | forage_ density | landcover | precip_ seas | road_ density | temp_ range | wood_ density | wood_ distance |
|---|---|---|---|---|---|---|---|---|
| altitude | 1.0 | -0.1 | 0.0 | 0.1 | -0.1 | 0.0 | 0.1 | 0.0 |
| forage_density | -0.1 | 1.0 | 0.0 | -0.4 | -0.1 | 0.2 | -0.7 | 0.3 |
| landcover | 0.0 | 0.0 | 1.0 | 0.1 | 0.0 | -0.1 | -0.4 | 0.3 |
| precip_seas | 0.1 | -0.4 | 0.1 | 1.0 | 0.1 | -0.5 | 0.2 | -0.1 |
| road_density | -0.1 | -0.1 | 0.0 | 0.1 | 1.0 | 0.0 | 0.0 | 0.0 |
| temp_range | 0.0 | 0.2 | -0.1 | -0.5 | 0.0 | 1.0 | 0.0 | 0.1 |
| wood_density | 0.1 | **-0.7** | -0.4 | 0.2 | 0.0 | 0.0 | 1.0 | -0.5 |
| wood_distance | 0.0 | 0.3 | 0.3 | -0.1 | 0.0 | 0.1 | -0.5 | 1.0 |

Table 17A. Pairwise Pearson's correlation coefficients calculated for selected variables using data extracted from randomly generated points (n = 10,000)

| | altitude | forage_ density | landcover | precip_ seas | road_ density | temp_ range | wood_ density | wood_ distance |
|---|---|---|---|---|---|---|---|---|
| altitude | 1.0 | -0.3 | 0.2 | 0.4 | -0.2 | -0.2 | 0.0 | 0.4 |
| forage_density | -0.3 | 1.0 | -0.4 | -0.3 | 0.0 | 0.2 | -0.4 | -0.2 |
| landcover | 0.2 | -0.4 | 1.0 | 0.1 | 0.2 | -0.1 | -0.3 | 0.3 |
| precip_seas | 0.4 | -0.3 | 0.1 | 1.0 | -0.1 | -0.6 | 0.2 | 0.2 |
| road_density | -0.2 | 0.0 | 0.2 | -0.1 | 1.0 | 0.1 | -0.1 | -0.1 |
| temp_range | -0.2 | 0.2 | -0.1 | **-0.6** | 0.1 | 1.0 | -0.1 | -0.1 |
| wood_density | 0.0 | -0.4 | -0.3 | 0.2 | -0.1 | -0.1 | 1.0 | -0.4 |
| wood_distance | 0.4 | -0.2 | 0.3 | 0.2 | -0.1 | -0.1 | -0.4 | 1.0 |

Table 18A. Summary of variance inflation factor (VIF) analysis for selected variables

| Variable | GVIF | Df | GVIF^(1/(2*Df)) |
|---|---|---|---|
| altitude | 2.15 | 1 | 1.47 |
| forage_density | 5.03 | 1 | **2.24** |
| landcover | 19.88 | 9 | 1.18 |
| precip_seas | 1.81 | 1 | 1.35 |
| road_density | 1.17 | 1 | 1.08 |
| temp_range | 1.52 | 1 | 1.23 |
| wood_density | 4.23 | 1 | **2.06** |
| wood_distance | 1.61 | 1 | 1.27 |

**(3) Preparation of the landscape for the mechanistic model**

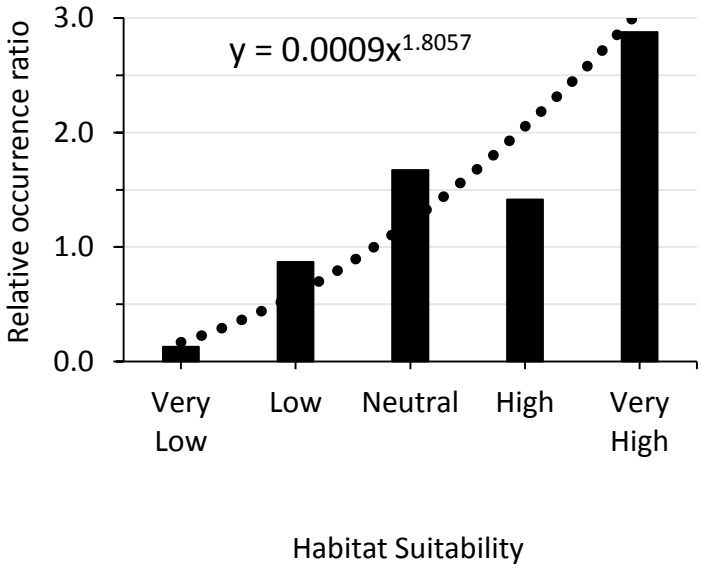


Figure 1A. Relationship between relative occurrence ratio, as calculated from eqn. 1, and habitat suitability for Wales (n=106).Categories are based on suitability values predicted by the MaxEnt model: Very low (0-0.2); Low (0.2-0.4); Neutral (0.4-0.6); High (0.6-0.8) and Very high (0.8-1). The equation for the trend line is shown at the top of the chart

The habitat suitability map produced by the MaxEnt model was scaled based on the equation for the trend line in Fig. 1A using the following equation:

$$Original\ suitability\ value^{1.8057}$$ equation 1A.

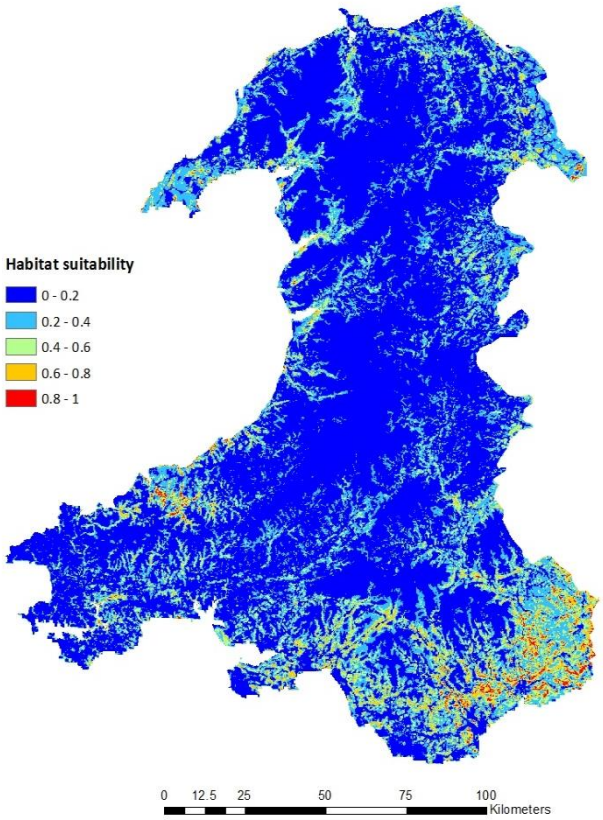


Figure 1A. Scaled suitability map for Wales used in the RangeShifter model.

**(4) Parameterization of the mechanistic model**

Table 19A. Summary of parameter settings used in the RangeShifter model

| Parameter | Setting | Rationale | References |
|---|---|---|---|
| Patch vs cell based | Patch | Dynamics should be modelled at a resolution relevant to the species (HR size) | (Bocedi *et al.* 2014) |
| Landscape | >Habitat suitability model output – contiguous patches<br>>Adjusted patch size according to suitability (using spatially biased points)<br>>Patches represent potential female home ranges within which several males may occupy territories | Resource availability likely to be the primary factor determining direction/speed of range expansion<br><br>HR range varies with habitat suitability | (Vincent *et al.* 1995; Wahlström & Liberg 1995; Tufto, Andersen & Linnell 1996; Pettorelli *et al.* 2001; Kjellander *et al.* 2004; Saïd *et al.* 2005; Gaillard *et al.* 2008) |
| Overlapping generations/stage structure | Stage structured model | Roe have discrete life stages:<br>juveniles<br>yearlings<br>adults<br>senescents | (Gaillard *et al.* 1993; Putman *et al.* 1996; Wäber, Spencer & Dolman 2013a) |
| Model type | Asexual/only females model | There is no way to explicitly simulate density dependence for each sex independently.<br><br>If we assume that patches represent male territories, males could occupy the same patch and both reproduce which is very unlikely.<br><br>Additionally, males could be less likely to settle if female density is high which is also unlikely. More reasonable to assume that, within | (Wahlström & Liberg 1995; Pettorelli *et al.* 2003; Gaillard *et al.* 2008; Debeffe *et al.* 2013; Bocedi *et al.* 2014) |

| | | | |
|---|---|---|---|
| | | occupied patches, there is sufficient number of males for a sustainable population<br><br>Dispersal probabilities and distances are estimated to be very similar between sexes so the overall pattern of expansion should not be affected.<br><br>Could increase patch size and assume multiple male territories within each patch but this will reduce the spatial resolution of the output | |
| Survival rates | Yearlings: 0.77<br>Adults: 0.96<br>Senescents: 0.73 | Taken from Gaillard *et al.* (1993) paper:<br><br>>Longest study of its kind<br>>Large sample size<br>>No predators (similar to UK)<br>>Similar habitat type<br>>Hunted population (similar to UK<br>>Use Chize only as TF experienced severe winters<br>>Widely used in other studies | (Gaillard *et al.* 1993) |
| Juvenile survival rates | Juveniles: 0.83 (accounted for by reducing fecundity) | Data not available from Gaillard *et al.* (1993) paper. Used Wäber *et al.* (2013) paper:<br><br>>Long term study (9 years)<br>>From Britain<br>>Only British study with large dataset<br>>Other results were consistent with Gaillard study | (Wäber, Spencer & Dolman 2013a) |
| Fecundity | Adults: 0.64<br>Senescents: 0.43 | 1.54 = mean no. of foetuses per female | (Hewison & Gaillard 2001; |

| | | | |
|---|---|---|---|
| | | Account for juvenile survival (83%)<br>=1.28<br>=0.64 (Females only, Wäber, Spencer & Dolman 2013a)<br><br>Hewison and Gaillard (2001) suggest 25% implantation failure in yearlings and adults and 50% implantation failure in senescents<br><br>(1.54/75)*100=2.05 (theoretical max litter size)<br>2.05*0.5=1.03 (50% failure)<br>(1.03/100)*83=0.85 (account for juvenile survival)<br>0.85/2=0.43 (Females only) | Wäber, Spencer & Dolman 2013a)<br><br>Waber et al 2013<br>Hewison et al 2001 |
| Max age | 14 | NA | (Loison *et al.* 1999) (same dataset as Gaillard study) |
| Nr. of reproductive seasons/year | 1 | NA | (Putman *et al.* 1996) |
| Probability of reproducing | 0.98 | >98% fertilization rate<br>>Density dependency affects fecundity by increasing implantation failure most likely mediated through resource availability. At higher densities females continue to reproduce (with a probability 0.98) but give birth to fewer offspring | (Hewison & Gaillard 1996, 2001) |
| Nr. of reproductive seasons before subsequent reproduction | 0 | NA | (Putman *et al.* 1996) |

| | | | |
|---|---|---|---|
| Number of stages | 4 | Juveniles (0-12 months)<br>Yearlings (12-24 months)<br>Adults (2-7 years)<br>Senescents (8-14 years | (Gaillard *et al.* 1993) |
| Scheduling of survival | Between reproductive events | Reproduction, dispersal, survival - all juveniles become yearlings within their first year – if they survive they are assumed to have sufficient fitness to disperse | (Wahlström & Liberg 1995) |
| Density dependence – Survival | YES- Juveniles only<br>Coeff: 1<br>Weights: Only Yearling, Adult and Senescent density influence juvenile survival | No consistent evidence for density dependence in adult survival rates<br>Some studies (Gaillard *et al.* 1993; Andersen & Linnell 2000; Pettorelli *et al.* 2001, 2003) suggest that density may interact with resource availability and mediate survival via body mass but a direct association has not been recorded.<br><br>Addition of density dependence in juveniles is to simulate resource dependent emigration (as suggested by *Debeffe et al.* (2013), Pettorelli *et al.* (2003), Gaillard *et al.* (2008) and Wahlstrom and Liberg (1995). Rangeshifter only has the capacity to mediate demographic parameters through density dependencies, not in response to resource availability or climate change.<br><br>However, as the landscape used is a suitability map, maximum popn. densities will be | (Gaillard *et al.* 1993; Wahlström 1994; Wahlström & Liberg 1995; Putman *et al.* 1996; Festa-Bianchet, Gaillard & Côté 2003; Cobben *et al.* 2009) |

| | | | |
|---|---|---|---|
| | | higher in more suitable, resource rich, patches.<br><br>Juveniles are likely to have highest fitness in resource rich patches when adult densities are low (i.e. resource availability is maximised).<br><br>Simply simulating emigration probability as density dependent will result in very high population densities within patches as non-dispersing individuals remain in their natal patch and may still reproduce (density dependence reduces fecundity but minimum values are >0).<br><br>More reasonable to model juvenile survival as density dependent and assume that all individuals with sufficient fitness disperse | |
| Density dependence – Development | NO | Model assumes that if individuals reach the required age they will develop to the next stage | (Bocedi *et al.* 2014) |
| Density dependence – Fecundity | YES | Several studies suggest density dependence in fecundity most likely through resource limitation so will be highly variable | (Gaillard *et al.* 1993; Vincent *et al.* 1995; Hewison & Gaillard 1996; Putman *et al.* 1996; Radeloff, Pidgeon & Hostert 1999; Focardi *et al.* 2002) |
| Strength of density dependence | 0.2 | 1/b = Max density/ha = approx.. 40 indivs/km$^2$ - modelling females only - 40/2=20<br>20/100=0.2 | (Wahlström & Liberg 1995; Andersen & Linnell 2000; Focardi *et al.* 2002; Cobben *et al.* 2009; Wäber, |

| | | | |
|---|---|---|---|
| | | | Spencer & Dolman 2013a) |
| Emigration probability | >Density-independent<br>>Stage dependent | Emigration thought to occur most at pre saturation levels<br><br>Some debate over the scale at which density dependence is evident.<br><br>Strong support for lack of a sex bias and dispersal occurring only in juveniles and yearlings. Adult roe are highly sedentary | (Wahlström & Kjellander 1995; Wahlström & Liberg 1995; Pettorelli *et al.* 2003; Gaillard *et al.* 2008) |
| Stage specific emigration probability | Juv=0<br>Yearlings=1<br>Adults=0<br>Senescents=0 | Adults and senescents highly sedentary<br><br>Dispersal most common in yearlings<br><br>Dispersal can occur in juveniles and 2 year olds but as there is only 1 dispersal event per lifetime using yearlings only is most appropriate<br><br>Assumes that all surviving juveniles emigrate from natal patch | (Wahlström & Liberg 1995; Gaillard *et al.* 2008; Debeffe *et al.* 2013) |
| Movement model | >Movement processes<br>>Stochastic Movement Simulator | Transfer is highly influenced by landscape structure<br><br>Dispersal distances are highly varied – distance alone is a poor predictor of emigration success<br><br>Evidence of pre-dispersal explorative movements | (Wahlström & Liberg 1995; Pettorelli *et al.* 2001, 2003; Nilsen, Linnell & Andersen 2004; Gaillard *et al.* 2008; Debeffe *et al.* 2013) |
| Stochastic movement simulator | >Step Mortality=0<br>>Perceptual Range=40 (cells) | No consistent evidence of dispersal mortality. It is | (Wahlström & Liberg 1995; Gaillard *et al.* |

| | | | |
|---|---|---|---|
| | >PR Method=2 (harmonic mean)<br>>Directional Persistence=5<br>>Used inverted habitat suitability map as cost map | assumed that dispersers are in good condition<br><br>PR=4 km, pre dispersal explorative movements<br><br>PR Method= Harmonic mean- likely to select good habitat rather than avoid costly habitat<br><br>Directional Persistence= Debeffe *et al.* (2013), dispersal tends to follow the direction of explorative movements. DP=5 considered a reasonable intermediate between 1 (minimum possible value, do not follow correlated paths) and 10 (quoted example, follow highly correlated paths) | 2008; Debeffe *et al.* 2013; Bocedi *et al.* 2014) |
| Settlement | >Settle if… Find a suitable patch (NOT natal patch) and density dependence<br>α = -100<br>β= 1<br><br>>If not settled, move until…Maximum nr. of steps=100<br>Max nr. of steps per year=0 (every individual disperses in 1 year)<br>>Min no. of steps=0 | Assumes mate availability is constant<br><br>Assuming that females disperse to maximise resource gain, settlement will not occur if a patch (representing a potential home range) is populated at carrying capacity.<br><br>20 km max dispersal distance (Euclidean distance) - further have been recorded but are rare. | (Wahlström & Liberg 1995; Debeffe *et al.* 2013; Bocedi *et al.* 2014)<br><br>Debeffe 2013<br>Wahlstrom 1995 |

**(5) Scaling patch size in relation to suitability for the landscape in the mechanistic model**

Patches were created using voronoi polygons from randomly generated points. 20,000 points were created to produce patches (mean area ≅ 1 $km^2$, after exclusion of very unsuitable habitat mean area = 0.58 $km^2$). Points were distributed according to the probability distribution described by the habitat suitability map (Figure 1A)

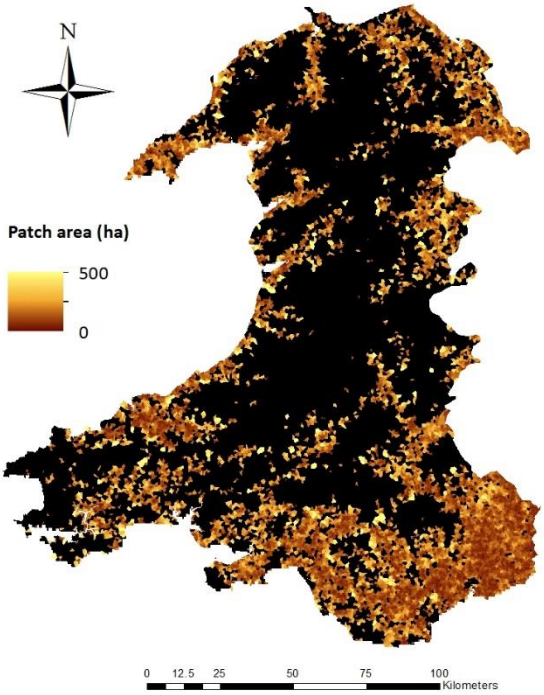


Figure 2A. Area of habitat patches in Wales used in the RangeShifter model. Black indicates areas of unsuitable habitat inaccessible to simulated individuals.

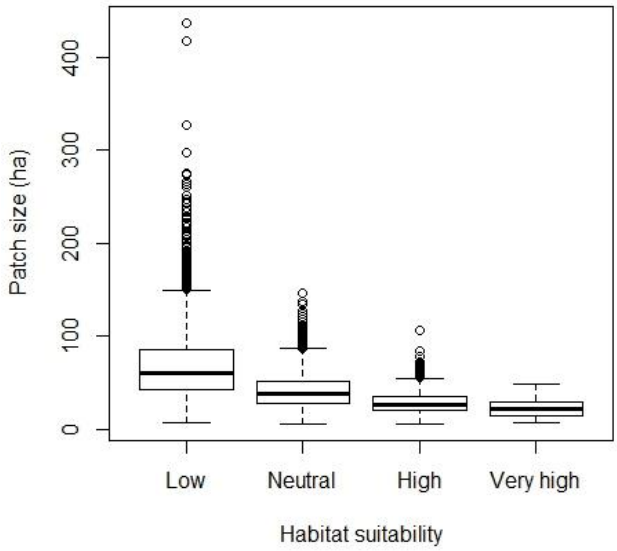


Figure 3A. Relationship between patch area (ha) and habitat suitability for the landscape used in the RangeShifter model.

## (6) Response curves produced by the MaxEnt model

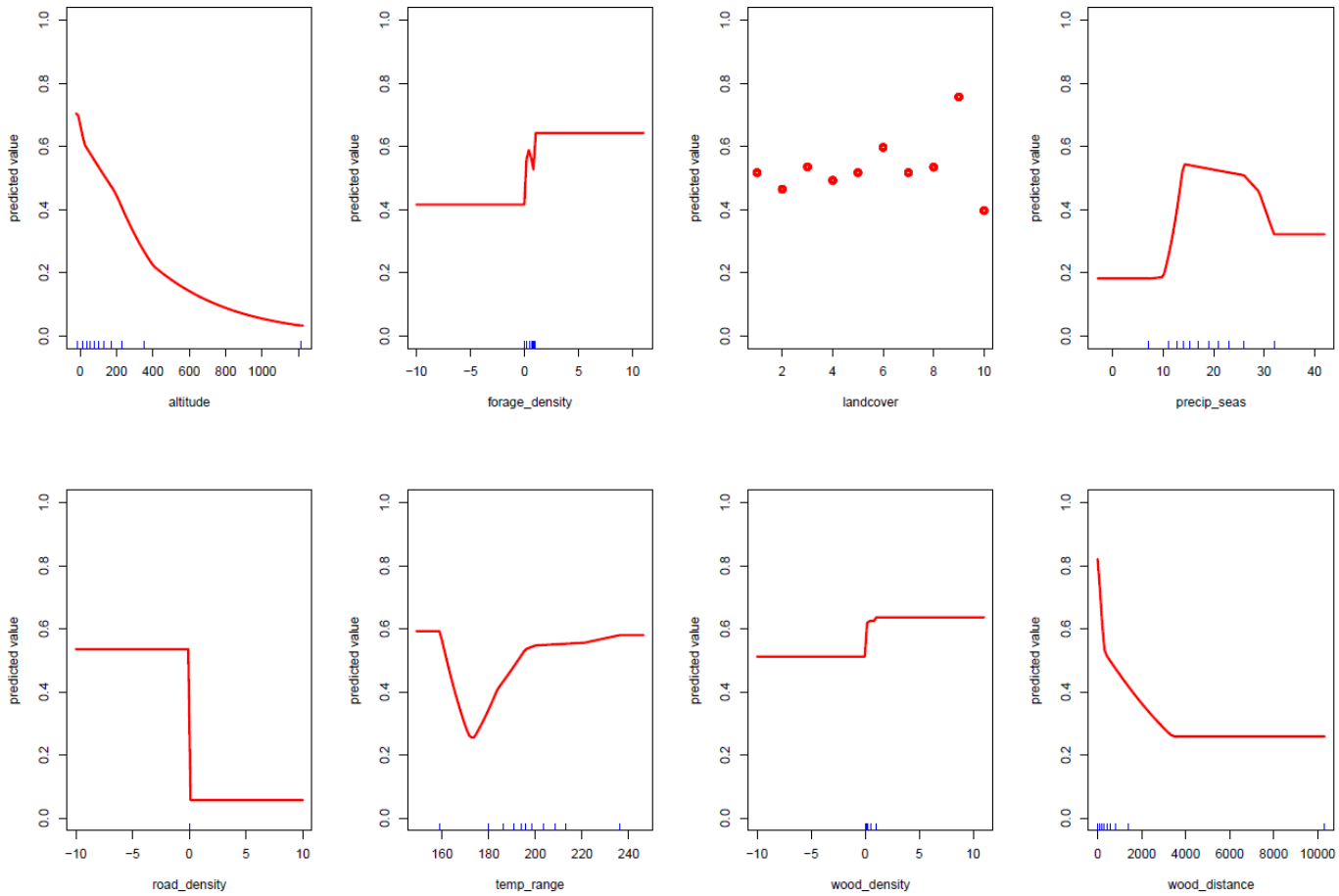


Figure 4A. Response curves produced by the MaxEnt model