Biological Conservation 217 (2018) 251–258

Contents lists available at ScienceDirect

Biological Conservation

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

Adding small species to the big picture: Species distribution modelling in an age of landscape scale conservation

Sally Eaton[a,*], Christopher Ellis[b], David Genney[c], Richard Thompson[d], Rebecca Yahr[b], Daniel T. Haydon[e]

[a] Royal Botanic Garden Edinburgh, 20A Inverleith Row, Edinburgh EH3 5LR, UK
[b] Royal Botanic Garden Edinburgh, UK
[c] Scottish Natural Heritage, Leachkin Road, Inverness IV3 8NW, UK
[d] Forest Enterprise Scotland, Glentress Peel, Peebles EH45 8NB, UK
[e] Institute of Biodiversity Animal Health and Comparative Medicine, University of Glasgow, G12 8QQ, UK

ARTICLE INFO

Keywords:
Compound variables
Generalised Additive Model (GAM)
Lichen epiphytes
Native Woodland Survey Scotland (NWSS)
Temperate rainforests

ABSTRACT

A recent shift in conservation policy from the site scale to the ecosystem or landscape scale requires underpinning by large-scale species distribution data. This poses a significant challenge in conserving small/less charismatic species (SLCS's) whose often cryptic nature can result in spatially restricted sampling, thus preventing landscape scale conservation projects from being realised for these ecologically important groups.

Species distribution models (SDMs) can provide a powerful tool to bridge this gap. However, in the case of SLCS's (here lichen epiphytes in temperate rainforests of western Scotland are used as a model system), direct predictor variables exist at micro-scales (millimetres to centimetres), which are not extensively available in landscape-scale datasets. Here we identify a group of well-mapped larger-scale ‘compound variables’ which capture the effect of multiple direct predictors (such as bark pH and topography), and test whether they can be successfully used to predict species distributions at the landscape scale, circumventing the need for direct (micro-scale) predictor data.

By testing the SDMs more widely within western Scotland, accurate predictions of species presence/absence could be made throughout the region for 5 of the 9 lichen epiphytes, making these SDMs extremely valuable as a conservation planning tool.

Species distribution models utilising compound variables as predictors offer a solution to the paucity of species distributional data for SLCS's, and present a valuable resource in conservation planning for such species. The importance of testing the SDMs outside of a training region to prevent prediction error is highlighted however.

1. Introduction

Charismatic megaflora/fauna make up just 2.8% of all life on earth (loosely all vascular plants and vertebrates, Jefferies, 2006) yet command the highest share of media exposure, research attention and conservation action (Clark and May, 2002). Because the remaining 97.2% of species are small and/or considered less charismatic (comprising for example nematodes, mollusks, arthropods, fungi and other microorganisms), they receive far less attention from the public, researchers, and policy makers (Jefferies, 2006). Nevertheless, many of these overlooked species are understood to be key providers of regulating and supporting ecosystem services (MEA, 2005). Obtaining a good understanding of the distribution of ecologically important small/less charismatic species is therefore essential for continued ecosystem service provision (Liu et al., 2013; Polce et al., 2013) and robust conservation planning (Whittaker et al., 2005).

Existing datasets for small/less charismatic species fall into two contrasting categories, being either (i) small scale and highly detailed (Rondinini et al., 2006), often concentrated on protected areas or monitoring plots, or (ii) extensive but coarse grained (Hartley et al., 2004) including standard 10 km mapping schemes (e.g. Seaward, 1995). This leads to a trade-off between spatially restricted (high

---

Abbreviations: AUC, area under the receiver operating curve; CT, Classification Tree; GAM, Generalised Additive Model; NPMR, nonparametric multiplicative regression; NWSS, Native Woodland Survey Scotland; SLSC's, small/less charismatic species; SDM, species distribution model; TSS, true skill statistic

* Corresponding author.
E-mail addresses: seaton@rbge.ac.uk (S. Eaton), c.ellis@rbge.ac.uk (C. Ellis), David.Genney@snh.gov.uk (D. Genney), Richard.Thompson@forestry.gsi.gov.uk (R. Thompson), R.Yahr@rbge.ac.uk (R. Yahr), Daniel.Haydon@glasgow.ac.uk (D.T. Haydon).

https://doi.org/10.1016/j.biocon.2017.11.012
Received 22 June 2017; Received in revised form 30 October 2017; Accepted 10 November 2017

resolution) accuracy and spatially extensive (low resolution) generality (McPherson et al., 2006), and thus fails to reflect a new conservation paradigm focused at the landscape-scale. Over the last decade, the species extinction crisis has led to a renewed focus on conserving entire ecosystems at landscape-scales, rather than individual site or single species strategies (Watson et al., 2011). This new approach has demanded species distribution data mapped over correspondingly large geographic areas, though at a resolution that reflects local variability in a species' habitat. This is most challenging for small/less charismatic species which require specialist taxonomic skills, or costly molecular diagnostics for identification. The surveying methods required to gather high resolution data for these species are therefore too limiting to be applied over entire landscapes (Britton et al., 2013; Vanderpoorten et al., 2005). As a result, ecologists are faced with a major challenge in generating species distribution data at a high resolution (at the granularity of the species' habitat) that is extensive (mapped at the landscape scale).

This study draws on species distribution modelling (SDM) as a widely-used technique to derive spatially explicit projections of habitat suitability (Guisan et al., 2013). This approach has made valuable contributions to conservation management and planning (Guisan and Zimmermann, 2000; Nicholls, 1989; Vaughan and Ormerod, 2003) across a wide range of different contexts (see review in Peterson, 2006 and Rodríguez et al., 2007). In the case of small/less charismatic species in particular, SDM provides a potentially cost-effective approach to mapping habitat suitability across large spatial areas (Nicholls, 1989).

The study focused on epiphytic lichens, a group of small cryptic species that lack comprehensive high-resolution distributional data at a landscape scale due to the specialist skills and intensive survey methods required to map them. Previous studies have successfully applied SDMs to project the large-scale epiphyte response to effects such as climate change, at a regional 10 km grid-scale (Ellis et al., 2007), but these approaches have not been applied continuously across a landscape at the resolution of the species' habitat (for epiphytic lichens these habitats are individual trees). Furthermore, the use of SDM in predicting the high-resolution occurrence of these species may be severely limited by the micro-scale over which environmental predictors operate, and which would demand environmental layers at centimetre-scales. For example, micro-niche factors associated with bark e.g. pH (Gauslaa, 1995; Jüriado et al., 2009; Kuusinen, 1996; Lewis and Ellis, 2010) and bark topography/texture (Bates, 1992; Ranius et al., 2008; Fritz et al., 2009) are known to affect the diversity and distribution of lichen epiphytes on individual trees (Ellis et al., 2015). These micro-niche factors will interact with local micro-climatic effects such as humidity/vapour pressure deficit (Hosokawa and Odani, 1957; Rambo, 2010) and light levels (Gustafsson and Eriksson, 1995; Kuusinen, 1994; Uliczka and Angelstam, 1999), to further influence epiphytic lichen distribution giving rise to environmental complexity within individual tree boles. Such fine-scale micro-niche and micro-climatic factors are impossible to map in a resource efficient way that would enable projections of species distribution across entire landscapes. To address this problem, and ultimately to enable the use of SDM to predict the distribution of small/less charismatic species, we test a series of compound variables for their ability to accurately predict species distributions. For example, tree species and size are widely available as digitised layers, and have the potential to capture differences among micro-niche effects such as pH and bark topography. Additionally, the distance to a water source and canopy cover can be mapped remotely, and may be used to represent micro-climatic factors such as humidity and light levels.

The method is tested for a range of lichen epiphyte species (exhibiting a variety of niche specialisms and reproductive modes), in order to understand whether there is variability in the applicability of the method to species exhibiting different ecological traits.

The success of SDMs can be measured by their ability to accurately predict species presence/absence within a region in which the model was fitted, and beyond this 'training region' into spatially removed but analogous habitat space. Considering the difficulties involved in gathering distributional data for small/less charismatic species, the extent to which a model may be used predictively outside the training region is critical. This extension to the use of a model is referred to as 'generality' (Fielding and Haworth, 1995), or 'transferability' (Kleyeer, 2002), and can be highly variable (Randin et al., 2006) depending on, for example, landscape scale population processes such as mass effects (Pulliam, 2000). Testing of model generality/transferability is therefore required to ensure that SDMs are not applied to environments in which predictive error is unacceptable, but conversely, practical application could be unnecessarily limited by avoiding the wider application of effective SDMs.

This study:

1. Tested the ability of SDMs to accurately assign small/less charismatic species to their suitable micro-environments using compound variables as predictors of species presence/absence. It is based on a training data set of 600 sub-sampled tree stems (representing an approximate basal trunk cross-sectional area of 54 m^2);
2. Applies the sub-sampled SDMs to a spatially extensive area relevant to conservation planning (a wooded landscape of 15 km in length) using a widely available digital dataset containing the compound variables;
3. Uses independent data to test the extent to which species distributions may be predicted beyond the wooded landscape, to characterize an entire region (representing a distance of over 200 km), through a comparison with ten widely sampled but comparable sites.

2. Methods

2.1. Study system

The species studied here are oceanic epiphytic lichens occurring in temperate rainforests in Scotland.

A total of nine epiphytic lichen species were selected for field survey according to three criteria: (i) their ability to be identified under field conditions using gross-morphology, (ii) their contrasting niche specialism (niche specialism was derived from local expert opinion), and (iii) their contrasting dispersal modes (derived from Purvis, 1992). Based on prior expectation, two niche generalists were chosen, frequently found in temperate rainforests throughout Scotland (the sexually reproducing *Graphis scripta* sensu lato, (L.) Ach., and the asexually reproducing *Parmelia saxatilis* sensu lato. (L.) Ach), five niche specialists, frequently-occasionally found in temperate rainforests throughout Scotland (the sexually/asexually reproducing *Lobaria pulmonaria* (L.) Hoffm., the asexually reproducing *Pannaria conoplea* (Ach.) Bory, the sexually reproducing *Pannaria rubiginosa* (Ach.) Bory, the sexually reproducing *Pectenia cyanoloma* (Schaerer) P. M. Jørg., L. Lindblom, Wedin & S. Ekman and the asexually reproducing *Pectenia atlantica* (Degel.) P.M.Jørg., L. Lindblom, Wedin & S. Ekman), and two 'super-specialists', occasionally-rarely found in temperate rainforests throughout Scotland (the asexually reproducing *Nephroma parile* (Ach.) Ach., and the sexually reproducing *Nephroma laevigatum*, Ach.).

The study area comprised a total of 11 sites lying within the temperate rainforest bioclimatic zone of Western Scotland, including the training site and ten test sites (see Fig. 1). The training site was located within Glen Creran, which is a north easterly aligned valley, comprising a sea loch with steep wooded slopes of native broadleaved semi natural woodland. The ten test sites included analogous areas of semi-natural ancient woodland within the Scottish National Forest Estate.

2.2. Collecting the training dataset

Data from the Native Woodland Survey of Scotland (NWSS, Forestry Commission Scotland, 2014) were used to stratify 181 individual