Biological Conservation 164 (2013) 170–176

Contents lists available at SciVerse ScienceDirect

Biological Conservation

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

The relative contribution of terrain, land cover, and vegetation structure indices to species distribution models

John W. Wilson [a,*], Joseph O. Sexton [b], R. Todd Jobe [c], Nick M. Haddad [a]

[a] Department of Biology, North Carolina State University, Campus Box 7617, Raleigh, NC 27695, USA
[b] Global Land Cover Facility, University of Maryland, Department of Geographical Sciences, 2181 LeFrak Hall, College Park, MD 20782, USA
[c] Department of Geography, University of North Carolina, Campus Box 3220, Chapel Hill, NC 27599, USA

ARTICLE INFO

Article history:
Received 9 February 2013
Received in revised form 1 April 2013
Accepted 20 April 2013
Available online 20 June 2013

Keywords:
Butterflies
Distributions
Habitat assessment
Niche models
Remote sensing
Threatened species

ABSTRACT

Habitat assessments for biodiversity conservation are often complicated by the lack of detailed knowledge of a study species' distribution. As an alternative to resource-intensive field-based methods to obtain such information, remotely sensed products can be utilized in species distribution models to infer a species' distribution and ecological needs. Here we demonstrate how to arbitrate among a variety of remotely sensed predictor variables to estimate the distribution and ecological needs of an endangered butterfly species occurring mainly in inaccessible areas. We classified 19 continuous environmental predictor variables into three conceptually independent predictor classes, terrain, land cover, and vertical vegetation structure, and compared the accuracy of competing Maxent habitat models consisting of different combinations of each class. Each class contributed, though disproportionately, to our most reliable model that considered all 19 variables. We confirm that variables obtained from remote sensors can effectively estimate the distribution and ecological needs of a relatively unknown imperiled species occurring in inaccessible locations. Importantly, increasing the variety of predictor classes through multi-sensor fusion resulted in greater model accuracy than increasing the absolute number of predictor variables.

1. Introduction

Considering that habitat loss is a primary driver of species extinctions, detailed habitat assessments are among the most important first steps guiding conservation efforts for imperiled species (Mace and Lande, 1991). Thorough habitat assessments are, however, often complicated by the lack of detailed knowledge of a threatened species' distribution, habitat status, and ecological needs (Anderson and Martinez-Meyer, 2004). Obtaining such information is not trivial. Threatened species are often sparsely distributed, hard to detect, and – due to biotic interactions, historical legacies, and dispersal barriers – not found in all suitable habitat patches (Pulliam, 2000). These qualities make it hard to separate unsuitable habitat from unoccupied suitable habitat (Gu and Swihart, 2004). In addition, time and monetary constraints typically prevent detailed bio-assessments that involve extensive surveys, experiments, and long-term demographic studies. Here we estimate the distribution and ecological needs of a relatively unknown imperiled species occurring in inaccessible locations, and, in doing so, develop methods to evaluate the contribution of a variety of readily accessible, continuous remotely sensed predictor variables that may be incorporated into species distribution models.

To overcome the challenges associated with imperiled species' habitat assessments, ecologists employ species distribution models (SDMs) to estimate imperiled species' distributions (Elith and Leathwick, 2009). Using spatial data describing distributions and environmental characteristics, SDMs estimate the relationship between the study species' occurrences and the underlying environment. These approximations of the target species' environmental niche are then used to map suitable ecological conditions over an entire study region (Elith and Leathwick, 2009). Because they enable researchers to overcome the challenges associated with resource-intensive bio-assessments, and because of improved model reliability, SDMs have become increasingly popular among ecologists and conservationists (Elith and Leathwick, 2009).

Progress in remote sensing technologies has strongly complemented advances in SDMs. As an alternative to resource-intensive field-based methods, air- and space-borne sensors enable researchers to acquire reliable environmental data at scales relevant to SDMs in a consistent and repeatable way (Gillespie et al., 2008), even from poorly known and inaccessible areas (Raxworthy et al., 2003). Despite their utility, remotely sensed predictor variables remain underutilized in SDMs, possibly because the literature offers little guidance on appropriate datasets (Buermann

* Corresponding author. Tel.: +1 919 272 3522.
E-mail addresses: johnnybirder@gmail.com (J.W. Wilson), jsexton@umd.edu (J.O. Sexton), toddjobe@unc.edu (R. Todd Jobe), nick_haddad@ncsu.edu (N.M. Haddad).

0006-3207/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.biocon.2013.04.021

et al., 2008) and interpretation of results obtained from remotely sensed data (Turner et al., 2003). Since the scale at which organisms perceive and interact with their environment is often much smaller than the scale at which many remotely sensed variables are obtained, concerns have also been raised as to whether remotely sensed data can be used to detect environmental variation at scales relevant to SDMs (Bistrat et al., 2011; Laurent et al., 2005).

The accelerating availability of diverse, remotely sensed products has generated questions about which and how many parameters to incorporate into model building. These parameters can be categorized into four conceptually independent remotely sensed predictor classes – terrain, (horizontal) land cover, (vertical) vegetation structure, and climate. Building on a previous effort that only considered land cover variables to track temporal habitat changes (Bartel and Sexton, 2009), we develop SDMs using a range of continuous remotely sensed predictor variables within three of these four remotely sensed predictor classes for an endangered butterfly, the St. Francis' satyr *Neonympha mitchellii francisci*. From these, we developed seven SDMs based on each predictor class independently, and in combination with one another. Four of our SDMs thereby utilized data from more than one sensor simultaneously, termed "multi-sensor fusion" (Hall and Llinas, 1997). Using our SDM results, we compared the performance of each SDM, blocked by data source, in predicting St. Francis' satyr presences. We also evaluated the relative contribution of each predictor variable to St. Francis' satyr distribution. In conducting our investigation, we developed an approach that tests significance of different classes of remotely sensed variables that should be generally applicable to arbitrate among competing models that could include various data inputs.

2. Material and methods

2.1. Study species

St Francis' satyr, globally restricted to early-successional wetlands situated on United States Department of Defense lands at Ft. Bragg, NC (35°07′S, 79°08′W, 65,032 ha), is an ideal species for a case study on SDMs utilizing remotely sensed data for a number of reasons. First, the species is listed as Endangered under the United States' Endangered Species Act because of its low population size and limited geographical range. Second, some previously healthy St. Francis' satyr subpopulations are currently in decline as once-suitable habitat transitions toward late-successional stages (Kuefler et al., 2008; Bartel and Sexton, 2009), creating an urgent need to assess the status of suitable habitat to determine the likelihood of population recovery. Third, our study area offers what we believe to be several suitable early-successional wetlands that support St. Francis' satyr's one known host plant, *Carex mitchelliana* sedges, which itself has very limited distributions. Yet, many of these patches remain unoccupied, raising questions about whether we can truly separate unsuitable from unoccupied suitable habitat. Fourth, much of the distribution of St. Francis' satyr falls within the restricted artillery impact zones at Ft. Bragg, where very limited and irregular access complicates efforts to confirm presences of this cryptic species with a short flight period (Kuefler et al., 2008). St. Francis' satyr is thereby representative of many other species whose life history is poorly described, and/or that live in inaccessible areas.

2.2. St. Francis' satyr occurrence

During 2008 we extensively (i.e. daily, during both month-long flight periods, Kuefler et al., 2008) searched for St. Francis' satyr butterflies in all known and accessible colonies ($n = 17$). For each butterfly observed, we obtained Universal Transverse Mercator (UTM) coordinates using a WAAS-enabled Trimble Nomad 900GL Global Positioning System (GPS) unit (1–3 m accuracy). In total, 138 GPS points were obtained, all within 3 m of butterfly observations to maximize locational accuracy (Graham et al., 2008). Because of the temporary, successional nature of St. Francis' satyr habitat, we based habitat suitability models on locations where we saw St. Francis' satyr during one focal year, 2008 (the year for which we obtained Landsat data, see below).

2.3. Predictor variables

We tested the relative importance of three conceptually independent predictor classes of remotely sensed predictor variables – terrain, (horizontal) land cover, and (vertical) vegetation structure – in explaining St. Francis' satyr distributions (Table 1). We omitted a fourth class consisting of climate measures because such data are usually coarsely scaled (Turner et al., 2003) and thus more appropriate for regional or continental SDMs (Gillespie et al., 2008; Elith and Leathwick, 2009). While some interpolated (e.g. Thornton et al., 1997) and combined (e.g. Herman et al., 1997) climate measures exist, remotely sensed climatic predictor variables are rare, especially for terrestrial surfaces.

Terrain variables, derived from Digital Elevation Models (DEMs) (Li et al., 2005), play an important, though indirect, role in SDMs through their influence on climate (Moore et al., 1990) and vegetation (Franklin, 1995). Five continuous terrain predictor variables were used in this study, which included proxies for moisture (flow accumulation and slope), solar radiation (aspect), and topography (relative slope position and terrain shape, Moore et al., 1990). All terrain variables in this study were derived from the USGS National Elevation Dataset (Gesch et al., 2002), which we obtained at 1/3 arcsec resolution, resampled to 10 m resolution, and processed using tools contained in the ArcGIS Spatial Analyst and TauDEM v. 4.0 (Tarboton, 2009) packages.

Land cover predictor variables, obtained through passive optical multispectral sensors, are used to describe a study area's physiographic and physiognomic characteristics. Most often, land cover

Table 1
Estimates of variable importance of terrain, land cover and vegetation structure variables used to predict St. Francis satyr presences using the Maxent software package (Phillips and Dudík, 2008).

Remotely sensed class	Variable	Variable importance (%)	Permutation importance (%)
Terrain	Slope	19.4	25.3
	Relative Slope Position	6.3	10
	Terrain Shape	0.5	0.6
	Aspect	0.6	0.2
	Flow Accumulation	0.5	0.2
Land cover	Deciduousness	24.1	41.3
	Summer brightness	9.3	2.3
	Wetness seasonality	1.6	2.2
	Brightness seasonality	1.3	1.2
	Winter greenness	2.0	1.1
	Summer wetness	5.4	0.6
	Winter brightness	0.8	0.4
	Winter wetness	0.9	0.4
	Summer greenness	15.7	0.3
Vegetation structure	Canopy density	2.0	8.5
	Understory density	3.0	2
	Shrub density	1.1	1.6
	Subcanopy density	3.6	1.4
	Midstory density	1.8	0.3

Variable importance is calculated heuristically and thus sensitive to collinearity and the order of variable importance. Permutation importance provides an alternative measure that is calculated from the AUC of the final model, and thus robust to the path of input variables.