

Review

Species detection vs. habitat suitability: Are we biasing habitat suitability models with remotely sensed data?

Bethany A. Bradley^{a,*}, Aaryn D. Olsson^b, Ophelia Wang^b, Brett G. Dickson^b, Lori Pelech^a, Steven E. Sesnie^b, Luke J. Zachmann^b

^a Department of Environmental Conservation, University of Massachusetts, Amherst, MA 01003, United States

^b Lab of Landscape Ecology and Conservation Biology, School of Earth Sciences and Environmental Sustainability, Northern Arizona University, Flagstaff, AZ 86011, United States

ARTICLE INFO

Article history:

Received 2 April 2012

Received in revised form 13 June 2012

Accepted 18 June 2012

Keywords:

Bioclimatic envelope model

Ecological niche model

Habitat suitability model

Land cover classification

Remote sensing

Species distribution model

Vegetation

ABSTRACT

Remotely sensed datasets are increasingly being used to model habitat suitability for a variety of taxa. We review habitat suitability models (HSMs) developed for both plants and animals that include remote sensing predictor variables to determine how these variables could affect model projections. For models focused on plant species habitat, we find several instances of unintentional bias in HSMs of vegetation due to the inclusion of remote sensing variables. Notably, studies that include continuous remote sensing variables could be inadvertently mapping actual species distribution instead of potential habitat due to unique spectral or temporal characteristics of the target species. Additionally, HSMs including categorical classifications are rarely explicit about assumptions of habitat suitability related to land cover, which could lead to unintended exclusion of potential habitat due to current land use. Although we support the broader application of remote sensing in general, we caution developers of HSMs to be aware of introduced model bias. These biases are more likely to arise when remote sensing variables are added to models simply because they improve accuracy, rather than considering how they affect the model results and interpretation. When including land cover classifications as predictors, we recommend that modellers provide more explicit descriptions of how habitat is defined (e.g., is deforested land considered suitable for trees?). Further, we suggest that continuous remote sensing variables should only be included in habitat models if authors can demonstrate that their inclusion characterizes potential habitat rather than actual species distribution. Use of the term 'habitat suitability model' rather than 'species distribution model' could reduce confusion about modelling goals and improve communication between the remote sensing and ecological modelling communities.

© 2012 Elsevier B.V. All rights reserved.

Contents

1. Introduction.....	58
2. Terminology.....	58
3. Remotely sensed data as predictors of animal habitat suitability.....	59
4. Remotely sensed data as predictors of plant habitat suitability.....	60
4.1. The chicken and egg problem: do trees only grow in forests?.....	60
4.2. Mistaking actual species distribution for potential species distribution.....	60
4.3. Sacrificing at the altar of accuracy.....	62
5. Recommendations for remote sensing in habitat suitability models.....	62
6. Conclusions.....	62
Acknowledgment.....	63
References.....	63

* Corresponding author at: 160 Holdsworth Way, Amherst, MA 01003, United States. Tel.: +1 413 545 1764; fax: +1 413 545 4358.

E-mail address: bbradley@eco.umass.edu (B.A. Bradley).

1. Introduction

Remotely sensed data are widely recognized for their applicability to ecological research (Kerr and Ostrovsky, 2003; Pettorelli et al., 2005). Ecological applications of remotely sensed data include classification and quantification of land features, modelling ecosystem function (e.g., to predict net primary productivity), and mapping land cover change (e.g., to identify habitat loss or afforestation) (Kerr and Ostrovsky, 2003). Remotely sensed data have also been used as a proxy for species richness and biodiversity (Gillespie et al., 2008; Turner et al., 2003). As remotely sensed datasets are more frequently and readily available, including the Landsat image archive (Woodcock et al., 2008) and moderate resolution imaging spectroradiometer (MODIS) phenology products (Tan et al., 2011), it is likely that ecological modellers and biogeographers will continue to expand their use of remotely sensed data.

Habitat suitability modelling in particular has seen considerable recent growth in the application of remotely sensed data. Habitat suitability models (HSMs; also termed species distribution models, ecological niche models, or bioclimatic envelope models) use empirical relationships between a species' distribution and environmental variables (e.g., climate, topography, and soils) to predict potential suitable habitats across a landscape or region (Franklin, 1995; Guisan and Zimmermann, 2000). Applications of HSMs include ecological reserve planning (e.g., Kremen et al., 2008), prediction of non-native species invasions (e.g., Thuiller et al., 2005), and risk assessments for native species (e.g., Thomas et al., 2004). Remotely sensed data can directly measure, or serve as a proxy for variables that affect habitat suitability. As a result, including remotely sensed data as variables can improve the overall accuracy of predictive models, making them attractive for use in HSMs (Bradley and Fleishman, 2008; Leyequien et al., 2007; Pettorelli et al., 2011).

Species habitat is affected by a range of environmental variables at varying scales. At regional to continental scales, suitability is most influenced by climate while, at landscape scales, climate suitability is modified by land use, land cover and topography. Suitability is further modified at local scales by soil conditions and micro-topography (Pearson and Dawson, 2003). Unfortunately, continuous spatial measurements of these environmental variables can be difficult to acquire, and many environmental variables are increasingly derived from remotely sensed data. Examples of these products include the normalized difference vegetation index (NDVI) as a proxy for ecosystem greenness (Fig. 1), the shuttle radar topography mission (SRTM) for high resolution topographic data, and emerging light detection and ranging (LIDAR) data as a proxy for vegetation community structure (for more details on these and other remote sensing products, see Gillespie et al., 2008; Pettorelli et al., 2005).

Remotely sensed measures of vegetation productivity (e.g., NDVI, enhanced vegetation index—EVI) and biophysical parameters (leaf area index—LAI) have been used extensively as predictors of habitat characteristics for animals (Pettorelli et al., 2011). However, a number of recent studies have applied vegetation proxies or land cover data to models of plant species habitat (Table 1). Unfortunately, the potential benefit of these remotely sensed data in HSMs constructed for plants is less clear. Plant habitat models could be biased by remotely sensed measures of vegetation like NDVI, resulting in outcomes that underestimate the species' potential distribution due to correlations between species distribution and spatial patterns identified by remote sensing. Remotely sensed proxies for vegetation could also highlight disturbed areas or other spatially heterogeneous patterns (e.g., recent precipitation history in arid lands; Fig. 2) that are poorly linked to habitat suitability.

In this context, we review remote sensing applications for habitat suitability modelling, briefly covering novel applications

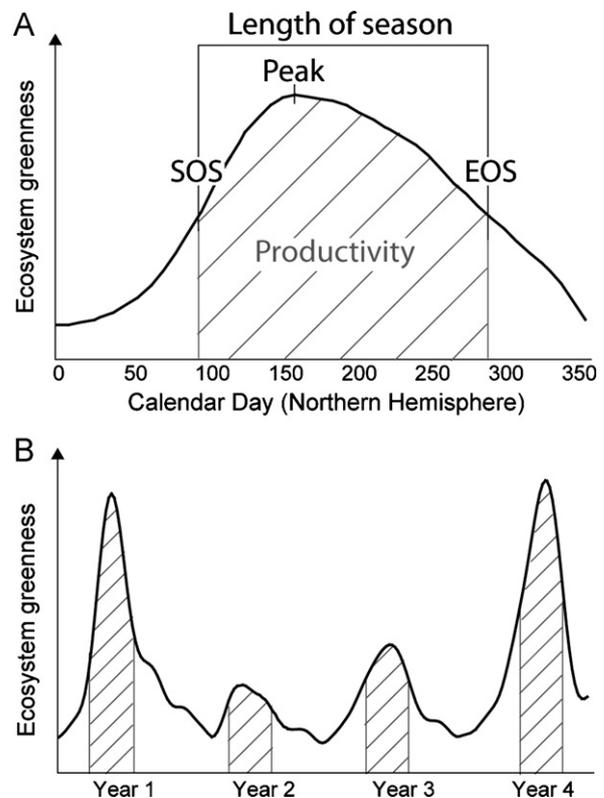


Fig. 1. Phenological metrics derived from time series of remotely sensed vegetation indices (e.g., NDVI) could provide novel predictor variables for animal and plant habitat models (Tan et al., 2011). (A) Example phenology metrics derived from single year or annual average vegetation phenology. Start of season (SOS) and end of season (EOS) are in this case based on the time when overall ecosystem greenness reaches half of its maximum value. (B) Example time series showing high inter-annual variability in vegetation productivity, potentially a predictor of habitat.

for animal habitat modelling (reviewed comprehensively by Leyequien et al., 2007; Pettorelli et al., 2011), and then focusing on the use (and possible misuse) of remotely sensed variables for plant habitat modelling.

2. Terminology

The science of habitat modelling based on empirical relationships between species distribution and spatially explicit environmental variables suffers from numerous near-synonymous terms (see Franklin, 2009 for a full discussion). One of the most widely used terms is species distribution model (SDM), which is generally understood by the ecological modelling community to mean a model of potential species distribution.

Unfortunately, this term in particular can cause confusion amongst the remote sensing community because remote sensing typically focuses on modelling the actual species distribution (e.g., Kerr and Ostrovsky, 2003; Xie et al., 2008). Remote sensing studies might aim to model the distribution of individual plant species (termed 'species mapping', Nagendra, 2001), while others model the distribution of dominant vegetation types (termed 'land cover classification', Kerr and Ostrovsky, 2003; Xie et al., 2008). Hence, there is a strong potential for confusion when terminology is poorly defined. Note, however, that there is less potential for confusion in animal studies because remotely sensed data are never used to detect actual animal distribution, only plant distribution. In this review, we will instead use the term habitat suitability model (HSM) to describe empirical models of potential species

Table 1
Reference list of habitat suitability models for plants that include remotely sensed variables.

Remotely sensed predictor variables	Target species	References
Spectrally homogenous land cover classification based on Landsat	Dyers woad (<i>Isatis tinctoria</i>)	Dewey et al. (1991)
European land cover (PELCOM) from AVHRR aggregated to 50 km	European tree species	Thuiller et al. (2004)
Land cover of Britain (2000) & automated land cover map derived from Landsat	Four plant species: <i>Rhynchospora alba</i> , <i>Erica tetralix</i> , <i>Salix herbacea</i> & <i>Geranium sylvaticum</i> .	Pearson et al. (2004)
ASTER-based classification of snow cover	Arctic dwarf shrub (<i>Dryas octopetala</i>)	Beck et al. (2005)
MODIS-derived phenology metrics based on NDVI and EVI	Tamarisk (<i>Tamarix</i> spp.)	Morisette et al. (2006)
MODIS NDVI for summer–fall of 2000	Purple loosestrife (<i>Lythrum salicaria</i>)	Anderson et al. (2006)
NDVI, wetness index, temperature, and soil brightness derived from Landsat	19 trees in the Great Basin, USA	Zimmermann et al. (2007)
MODIS LAI, veg moisture from Qscat and MODIS NDVI	Three Amazon trees: <i>Calophyllum brasiliensis</i> , <i>Carapa guianensis</i> , <i>Virola surinamensis</i>	Prates-Clark et al. (2008)
QSCAT backscatter for canopy roughness and MODIS LAI phenology metrics	Five widespread commercial timber trees	Saatchi et al. (2008)
MODIS phenology metrics derived in house from 2001 to 2007 time series	Pine (<i>Pinus</i> spp.) and white oak (<i>Quercus</i> spp.)	Cord et al. (2009)
MODIS phenology metrics, tree cover	Dalmatian toadflax, Musk thistle, Cheatgrass, White sweet clover	Stohlgren et al. (2010)
Forest type classes based on visual interpretation of satellite imagery	Rare tree: <i>Pittosporum eriocarpum</i>	Padalia et al. (2010)
Global land cover 2000, % tree cover in 2001, AVHRR average NDVI 1985–1988	Twenty-nine palm species in Africa	Blach-Overgaard et al. (2010)
MODIS EVI and land surface temperature	Tamarisk (<i>Tamarix</i> spp.)	Cord et al. (2010)
MODIS phenology metrics	Two central American trees: <i>Broumum alicastrum</i> and <i>Liquidambar macrophylla</i>	Cord et al. (2011)
MODIS mean EVI and annual range of EVI	Tamarisk (<i>Tamarix</i> spp.)	Jarnevich et al. (2011)
Landsat NDVI, greenness and brightness indices	<i>Rosa rubiginosa</i> in Argentina	Zimmermann et al. (2011)
Landsat spectral indices and NDVI	Generalist shrub species in Spain	Morán-Ordóñez et al. (2012)

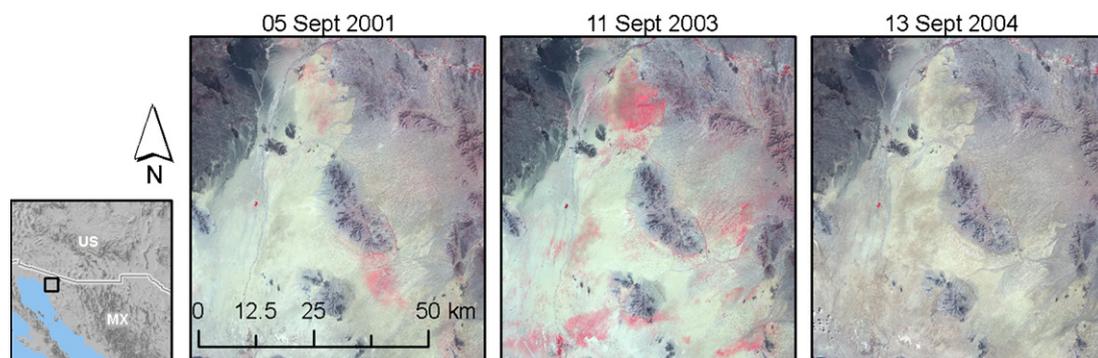


Fig. 2. Three Landsat TM colour composite images (bands 4, 3 and 2) of the Sierra Pinacate in northern Sonora, Mexico, each acquired in early September. Red colours identify areas with actively growing vegetation (similar to high NDVI) in a hot desert environment where precipitation drives phenology. The spatially heterogeneous patterns are caused by asynchronous plant growth due to isolated rainfall events. If used in a habitat suitability model, the model prediction could vary considerably depending on inter-annual variability (Fig. 1B). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)

distribution based on environmental correlates (which might include remotely sensed data).

3. Remotely sensed data as predictors of animal habitat suitability

There are numerous examples of remote sensing variables being tested and applied to animal habitat modelling (Leyequien et al., 2007; Pettorelli et al., 2011; Vierling et al., 2008). In many (but not all) cases, the addition of remotely sensed data improves the accuracy of the habitat model (Pettorelli et al., 2011). In most of these studies, remote sensing variables serve as proxies for vegetation (e.g., structure, composition, and land cover change) or other attributes of habitat quality. For example, Willems et al. (2009) used NDVI to model habitat of the vervet monkey in Africa, with areas of high NDVI acting as a proxy for food availability and low visibility for predators. Bergen et al. (2007) used biomass measurements derived from RADAR to model North American bird habitat, with RADAR-derived biomass acting as a proxy for forest structure.

Peterson et al. (2006) used time series of land cover classifications to model corvid habitat in Mexico, with changes in land cover acting as a proxy for habitat loss.

The use of NDVI or land cover classifications in conjunction with climatic variables often does not improve habitat models because these variables can be highly collinear, particularly at regional scales (e.g., Thuiller et al., 2004; Zimmermann et al., 2007). Remote sensing variables have the greatest potential benefit when they provide information distinct from climate. This may occur in areas of broadly similar climate where other features such as soils or disturbance alter vegetation characteristics. In other cases, remotely sensed vegetation indices could be more reflective of climate conditions than interpolated climate because weather stations are hundreds of kilometres apart (e.g., the Amazon; Saatchi et al., 2008). In addition to vegetation indices, other novel remote sensing products and techniques are promising for habitat modelling.

Technological and scientific advances have created new proxies for vegetation characteristics that are quite distinct from climate: RADAR/LIDAR-based vegetation structure and temporal patterns of

vegetation phenology. High-resolution (e.g., <5 m pixel) RADAR and LIDAR measurements are increasingly being used to characterize three-dimensional vegetation community structure (Bergen et al., 2009; Vierling et al., 2008). RADAR and LIDAR are active sensors, whereby a long-wavelength or laser pulse is directed at the Earth's surface. The returned waveform or coordinates can be used to estimate detailed horizontal and vertical vegetation attributes, such as forest canopy height and aboveground biomass (Lefsky et al., 2002). In one example application, Buermann et al. (2008) used RADAR sensitivity to estimate canopy roughness at 1 km resolution as a predictor of bird habitat in the Amazon. The National Aeronautics and Space Administration (NASA) is likely to launch a RADAR satellite aimed at measuring ecosystem structure in 2016 (DESDynI), while the European Space Agency (ESA) will launch the RADAR mission Sentinel-1 in 2013. Meanwhile, state-wide airborne LIDAR data are becoming more readily available (e.g., Asner et al., 2011).

A second novel predictor of vegetation characteristics that contribute to animal habitat suitability is phenology, or the annual and inter-annual timing of biological events. Average annual phenology is broadly correlated with climate, particularly start of season with temperature (e.g., Stöckli and Vidale, 2004), and may serve as a climate proxy in poorly gaged areas. Spatial variation in phenology can identify differential responses, even within similar ecosystems, that could affect habitat quality (e.g., Fisher and Mustard, 2007; Morissette et al., 2009), while inter-annual phenology can serve as an indicator of vegetation variability, which may be important for long-lived animals. Phenological metrics, typically derived from NDVI time series, include start of season, length of growing season, and inter-annual variability (Tan et al., 2011) (Fig. 1). As a predictor of animal habitat, vegetation phenology has been, for example, correlated with mosquito life cycles resulting in malaria outbreaks in Africa (Rogers et al., 2002), habitat of Great Bustards in Spain (Osborne et al., 2001) and moose body mass in Norway (Herfindal et al., 2006). Newly available phenology metrics derived from high temporal resolution imagery such as MODIS (Tan et al., 2011) at 250 m to 1 km resolution and advanced very high resolution radiometer (AVHRR; <http://phenology.cr.usgs.gov/index.php>) at 1–8 km resolution should increase the applications of this potential predictor in animal habitat modelling.

4. Remotely sensed data as predictors of plant habitat suitability

Although remotely sensed variables are widely used to model animal habitat (Leyequien et al., 2007; Pettorelli et al., 2011), the practice has been less common in models of plant habitat. However, applications for plant habitat modelling are on the rise (Table 1). Modelling plant habitat requires more care because remotely sensed variables and land cover classifications often directly measure the attributes of the same plant species whose habitat the model aims to predict. This point seems to be underappreciated in plant habitat modelling. Below, we review potential sources of bias that remotely sensed data can introduce into HSMS, and how they have been handled (or mishandled) in recent research.

4.1. The chicken and egg problem: do trees only grow in forests?

Are forests the only locations that should be defined as potential tree habitat? The answer to this question lies at the heart of whether and how land cover classifications and continuous variables should act as proxies for vegetation cover (e.g., NDVI) and be used to predict plant and animal habitat suitability. If distribution data occur only in currently forested areas (recently collected data are more likely to reflect current land use/land cover patterns), then including these variables will effectively exclude presently

'non-forested' land from potential habitat. For example, Prates-Clark et al. (2008) use leaf area index (LAI) and RADAR as predictor variables for habitat of three rare Amazon trees. The use of LAI in this instance excluded deforested lands, which had low LAI values. Similarly, by including NDVI in habitat suitability models, Cord et al. (2009) exclude agriculture, urban, and degraded lands from potential habitat for pine and oak species in Mexico. In both of these cases, the model of suitable tree habitat excludes human modified landscapes. This assumes that potential tree habitat is defined relative to current land use, but the assumption is never made explicit. If currently deforested or non-forested areas are considered to be unsuitable for tree species, then these areas should be excluded prior to further analysis.

Conversely, Zimmermann et al. (2011) use remote sensing-derived brightness and greenness indices to detect forest clearance as a predictor of the invasive shrub *Rosa rubiginosa* in Argentina. In this case, the authors are explicit about their definition of habitat suitability (or invasibility) as stemming directly from existing land clearing and disturbance. The specific interpretation of remote sensing variables in Zimmermann et al. (2011) makes the resulting model more readily understandable and applicable to invasive plant management.

The same problem of defining habitat relative to current land cover potentially applies to any use of land cover classifications, even in the absence of anthropogenic factors. For example, Thuiller et al. (2004) note that modelled habitat suitability for the European tree *Quercus petraea* is positively associated with percentage forest cover (derived from land cover classifications). Similarly, the modelled habitat of the rare tree *Pittosporum ericarpum* in India is best described by including forest composition classes (Padalia et al., 2010). In these examples, the resulting habitat models will be biased towards current forest, and away from currently non-forested land, suggesting that climatically suitable non-forest is unsuitable habitat for trees.

To overcome this 'chicken and egg' problem, habitat suitability modelling efforts that include land cover classifications (or proxies for land cover) should be explicit about their goals and applications, and discuss how the inclusion of land cover classes influences their interpretation. For example, Pearson et al. (2004) argue that by including land cover in a habitat model of the flower *Erica tetralix*, they were able to differentiate between range contraction caused by climate and range contraction caused by land use. Beck et al. (2005) use a remote sensing based classification of snow cover to exclude habitat for the Arctic shrub *Dryas octopetala*, which does not grow if the climate is moist enough to produce snow. Morán-Ordóñez et al. (2012) show that current land cover limits a generalist species more than climate alone. In all of the above examples, the authors provide a clear interpretation of how the inclusion of remotely sensed variables alters their models of habitat.

4.2. Mistaking actual species distribution for potential species distribution

The use of continuous remote sensing variables (e.g., spectral bands, NDVI, LAI) can be a problem if these variables act as proxies for land cover (see previous section), or if they are sensitive to unique spectral or temporal properties of the target species itself (Fig. 3). This is of particular concern if the target species is common, grows in patches with high abundance that could be detected remotely, or has a unique phenology (e.g., Bradley and Mustard, 2005; Tuanmu et al., 2010). For example, Zimmermann et al. (2007) used time series of Landsat-derived NDVI to develop habitat models for 19 tree species across a Utah landscape at 90 m spatial resolution. The 19 species ranged from rare to common, but those for which habitat models were most improved using remote sensing were broadleaf, deciduous species that are

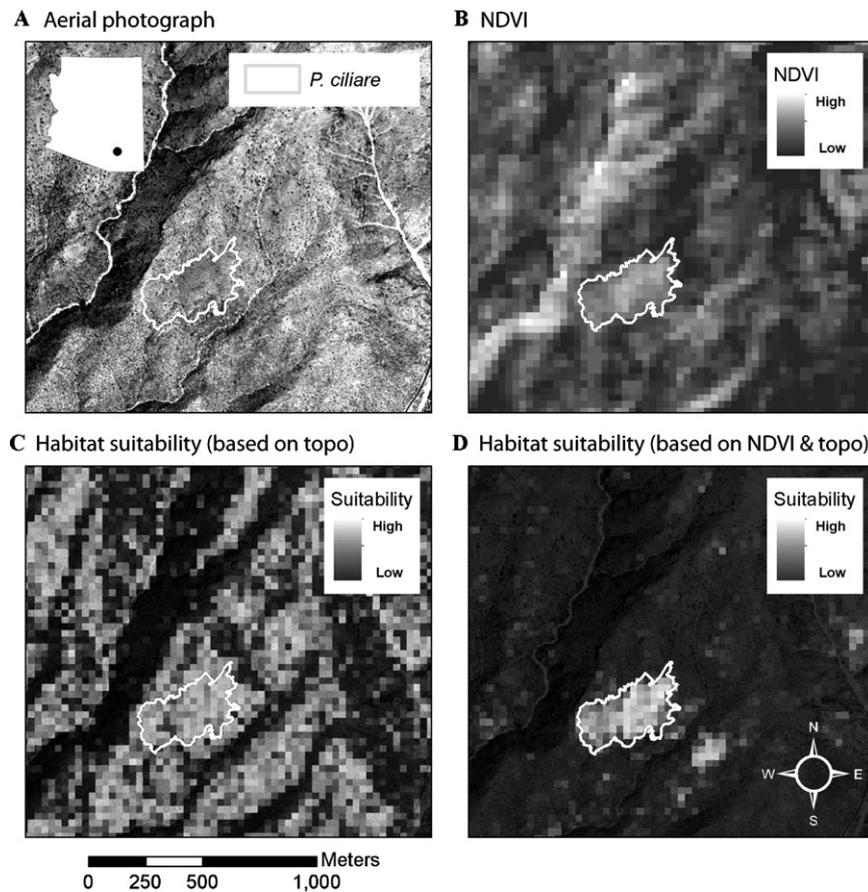


Fig. 3. (A) Aerial photograph taken near Tucson, AZ, which is undergoing invasion by *Pennisetum ciliare*. The mapped boundary of a large and dynamically expanding patch is shown in outline. Because of its cold intolerance, *P. ciliare* tends to colonize south-facing slopes in this part of its range. Inset shows this location in relation to the state of Arizona, USA. (B) NDVI of the invasion site is higher than on neighbouring south-facing slopes due to denser cover of *P. ciliare*. But, NDVI is similar to surrounding, more heavily vegetated north-facing slopes where *P. ciliare* is rare or absent. (C) Habitat suitability model of *P. ciliare* derived from topographic variables (elevation, slope, aspect) using presence and absence data collected throughout the area and based on a Random Forest regression tree model (Breiman, 2001). Note that areas of highest suitability (light shades) are associated with south-facing slopes. This model explained 2.6% of the variance. (D) Habitat suitability model of *P. ciliare* derived from topographic variables plus NDVI. Because NDVI acts as a proxy for species actual distribution, the predicted suitability model identifies only areas where *P. ciliare* is already present, yet *P. ciliare* has been doubling in area at this site every 2–5 years since 1989 (Olsson et al., 2012). This model explained 38.1% of the variance (an “improvement” of model accuracy).

easier to discriminate using multi-temporal satellite data because of their seasonal phenology. This is one example of including a predictor variable that models actual species distribution rather than potential habitat. Hence, the resulting model likely underestimated potential suitability because it essentially included a proxy for current distribution. In contrast, Saatchi et al. (2008) included RADAR and LAI in a habitat model for five widespread commercial timber trees in the Amazon at 1–2 km spatial resolution, while Cord et al. (2011) used MODIS phenology as a predictor for two tree species in Mexico at 1 km spatial resolution. The coarse spatial resolution of these studies combined with high tree diversity in both tropical regions likely prevented remote sensing bias from entering either model. However, including remotely sensed data in habitat models for common species should always be treated with caution because common or abundant species are more likely to directly influence measurements obtained from remotely sensed data.

Other examples of this phenomenon come from the invasive species literature. Invasive plants can grow as a monoculture that can be spectrally or phenologically unique (e.g., Bradley and Mustard, 2005; Casady et al., 2005; Huang and Geiger, 2008; Resasco et al., 2007), or can grow in spectrally unique areas, such as abandoned farmland (e.g., Elmore et al., 2006). Characterizing these distinguishing features in remote sensing is frequently the

most successful method for mapping invasive species at both landscape (e.g., Landsat) and regional (e.g., MODIS) scales. Hence, HSMS created for common invasive plants could be easily biased with the addition of remotely sensed data (Fig. 3).

For example, Stohlgren et al. (2010) predicted invasive plant habitat suitability for *Linaria dalmatica* in Yellowstone National Park, and for *Bromus tectorum* in Sequoia and Kings Canyon National Parks based on phenological metrics derived from MODIS NDVI at 250 m resolution. However, *L. dalmatica* has been shown to be spectrally unique in Yellowstone (Rew et al., 2005), while *B. tectorum* is phenologically unique throughout much of its range (Bradley and Mustard, 2005; Peterson, 2005), and both species often occur in extensive monocultures. Hence, it is possible that remotely sensed variables used in the habitat suitability models were biased by invasive species' actual distributions, thereby underestimating invasion risk.

In a second example, Morissette et al. (2006) used phenological metrics from MODIS NDVI and EVI to model habitat suitability for *Tamarix* spp. across the U.S. at 250 m resolution. Correlation along the 1:1 line between annual NDVI and EVI values was identified as one of the best remotely sensed predictors of tamarisk habitat suitability. However, the authors hypothesize that this relationship occurs where tamarisk canopy cover is thick enough to

Table 2
All three of the criteria below must be met in order for a target plant species to influence remotely sensed variables (NDVI or reflectance spectra) in a way that could model actual species distribution rather than model potential habitat.

1. Species is part of the overstorey	
Examples	Forest canopy trees Any plant species in single-level vegetation communities (grassland, shrubland)
2. Species is common relative to the spatial resolution of the remote sensing instrument^a	
Examples	Work around?
Dominant native species	Choose imagery dates prior to invasion
Invasive plants existing in high abundance across landscapes	Use coarser spatial resolution imagery
3. Species has a unique phenological or spectral signal	
Examples	Work around?
Species is the first to green up in the spring or is highly responsive to precipitation	Avoid using phenological metrics, avoid imagery coinciding with unique phenological stages
Flowers are prominent and a unique colour or leaves are spectrally distinct from other vegetation	Avoid flowering times, avoid unique spectral bands

^a E.g., abundant across 30 m × 30 m areas for Landsat or 1 km² areas for MODIS.

mask underlying soil. If true, the resulting habitat model was biased towards locations where tamarisk already occurs and could underestimate invasion risk in uninvaded areas suitable for establishing new populations. An update of the tamarisk habitat suitability model (Jarnevich et al., 2011) also included MODIS EVI, however, the remotely sensed variable was of low importance in the newer model and these results are unlikely to be biased.

Problems with introduced bias are less likely to occur if the target plant species is rare or if the remotely sensed variables do not measure vegetation directly (e.g., land surface temperatures; Cord et al., 2010; microtopography from LIDAR; Sellars and Jolls, 2007). If a species is rare or in low abundance, remotely sensed variables will be responsive to the habitat characteristics rather than the species itself. If the species is common or abundant, the same result could be accomplished by using remotely sensed data at a much coarser resolution than the species extents (Anderson et al., 2006; Saatchi et al., 2008; Cord et al., 2011). For invasive species, satellite data that pre-date infestations could provide unbiased information on suitable habitat (Dewey et al., 1991), although new infestations resulting from recent land use would not be captured. If detection bias continues to be a potential problem, researchers should forego using remotely sensed variables unless they can convincingly argue that those variables do not bias the habitat model.

4.3. Sacrificing at the altar of accuracy

Overall model accuracy is most often the measuring stick used to include or exclude predictor variables from habitat suitability models. However, unbiased HSMs will almost always have lower overall accuracy than biased HSMs because maps of species distribution have lower overprediction rates than models of habitat suitability. Thus, achieving higher overall accuracy should not be the sole criterion for evaluating inclusion of remote sensing variables in habitat suitability models. In addition, researchers must consider whether or not there is any a priori reason to suggest that remote sensing variables will improve the characterization of suitable habitat, but not of actual species distribution. If this second criterion can be justified, then the model can be considered unbiased.

5. Recommendations for remote sensing in habitat suitability models

A first important step towards reducing confusion when including remote sensing variables would be to clarify terminology. The term species distribution implies actual distribution when it should mean potential distribution. For example, Saatchi et al. (2008) use the term 'species distribution' interchangeably with potential range, potential distribution, and suitable habitat. This mixing of terms can be of particular concern when communicating with the

remote sensing community, where species distribution model is apt to be interpreted as 'mapping species distribution'. We prefer the terminology of Kearney (2006), where 'habitat' is defined as the physical characteristics of the place where an organism potentially lives, and thus a habitat suitability model projects suitable habitat for that organism. However, any terminology will suffice provided the authors explicitly define their goals relative to actual and potential distribution.

Second, climate variables used in HSMs are typically selected either a priori based on species physiology, or empirically based on the best statistical relationship. In the former case, the influence of predictor variables on species habitat is assumed, while in the latter it is often interpreted. The same approach should be taken with remotely sensed predictor variables. Rather than just selecting and including variables that improve overall accuracy, defining a priori the role of remote sensing variables, or interpreting their role after the fact would greatly help the interpretation and application of such models.

Third, we emphasize the need for caution when including continuous remotely sensed variables in models where there is reason to believe remote sensing might map actual species distribution. Confusion of potential habitat with actual distribution is most likely if the target plant species exists in high abundance within portions of its range. It is also more likely if the target plant species is itself spectrally or phenologically unique from surrounding vegetation such that remote sensing variables could identify its presence (Table 2).

Finally, habitat suitability modelling studies that include remote sensing variables need to be more explicit about their definition of habitat. Remotely sensed variables reflect the current state of anthropogenic influence on the Earth's surface. Hence, their inclusion in habitat models, both animal and plant targeted, makes the most sense when land use or land cover change is an explicit consideration in the definition of potential habitat (for example, should a deforested area be excluded from suitable habitat for a given tree species). Depending on the goals of the modelling exercise, and the timeline for which it is appropriate, excluding particular land cover types may or may not be appropriate. Better defining habitat suitability relative to climate and land use would facilitate the interpretation and application of model results.

6. Conclusions

We have shown in this review that, while the use of remote sensing variables has been widely and reasonably applied to modelling animal habitat, there is substantial potential for bias when remote sensing variables are included in models of plant habitat. This is a young field, so the literature on this topic is relatively sparse (Table 1). However, the growing availability of remotely

sensed data suggests that their application to habitat suitability models will continue to increase. We strongly encourage both plant and animal habitat modellers to interpret remote sensing predictor variables relative to the habitat they aim to model by developing hypotheses about the ecological relationships those variables might reveal. This practice, along with explicitly stating how habitat is defined relative to existing land use or land cover, will reduce the likelihood of developing biased models and improve overall model interpretation and application.

Acknowledgment

This research was supported wholly by the U.S. Department of Defense, through the Strategic Environmental Research and Development Program (SERDP).

References

- Anderson, R.P., Peterson, A.T., Egbert, S.L., 2006. Vegetation-index models predict areas vulnerable to purple loosestrife (*Lythrum salicaria*) invasion in Kansas. *The Southwestern Naturalist* 51, 471–480.
- Asner, G.P., Hughes, R.F., Mascaro, J., Uwololo, A.L., Knapp, D.E., Jacobson, J., Kennedy-Bowdoin, T., Clark, J.K., 2011. High-resolution carbon mapping on the million-hectare Island of Hawaii. *Frontiers in Ecology and the Environment*, <http://dx.doi.org/10.1890/100179>.
- Beck, P.S.A., Kalmbach, E., Joly, D., Stien, A., Nilsen, L., 2005. Modelling local distribution of an Arctic dwarf shrub indicates an important role for remote sensing of snow cover. *Remote Sensing of Environment* 98, 110–121.
- Bergen, K.M., Gilboy, A.M., Brown, D.G., 2007. Multi-dimensional vegetation structure in modeling avian habitat. *Ecological Informatics* 2, 9–22.
- Bergen, K.M., Goetz, S.J., Dubayah, R.O., Henebry, G.M., Hunsaker, C.T., Imhoff, M.L., Nelson, R.F., Parker, G.G., Radeloff, V.C., 2009. Remote sensing of vegetation 3-D structure for biodiversity and habitat: review and implications for lidar and radar spaceborne missions. *Journal of Geophysical Research, Biogeosciences* 114.
- Blach-Overgaard, A., Svenning, J.C., Dransfield, J., Greve, M., Balslev, H., 2010. Determinants of palm species distributions across Africa: the relative roles of climate, non-climatic environmental factors, and spatial constraints. *Ecography* 33, 380–391.
- Bradley, B.A., Fleishman, E., 2008. Can remote sensing of land cover improve species distribution modelling. *Journal of Biogeography* 35, 1158–1159.
- Bradley, B.A., Mustard, J.F., 2005. Identifying land cover variability distinct from land cover change: cheatgrass in the Great Basin. *Remote Sensing of Environment* 94, 204–213.
- Breiman, L., 2001. Random forests. *Machine Learning* 45, 5–32.
- Buermann, W., Saatchi, S., Smith, T., Zutta, B., Chaves, J., Mila, B., Graham, C., 2008. Predicting species distributions across the Amazonian and Andean Regions using remote sensing data. *Journal of Biogeography* 35, 1160–1176.
- Casady, G.M., Hanley, R.S., Seelan, S.K., 2005. Detection of leafy spurge (*Euphorbia esula*) using multitemporal high-resolution satellite imagery. *Weed Technology* 19, 462–467.
- Cord, A., Klein, D., Dech, S., 2010. Remote sensing time series for modeling invasive species distribution: a case study of *Tamarix* spp. in the US and Mexico. In: Swayne, D.A., Yang, W., Voinov, A.A., Rizzoli, A., Filatova, T. (Eds.), *International Congress on Environmental Modelling and Software Modelling for Environment's Sake*. International Environmental Modelling and Software Society (iEMSs), Ottawa, CA.
- Cord, A., Klein, D., Dech, S., 2011. The impact of inter-annual variability in remote sensing time series on modeling tree species distributions. In: 6th International Workshop on the Analysis of Multi-temporal Remote Sensing Images (Multi-Temp), pp. 181–184.
- Cord, A., Schmidt, M., Dech, S., 2009. Potential and limitations of multi-temporal earth observation data to improve model results of tree species distribution in Mexico. In: 33rd International Symposium on Remote Sensing of the Environment, Stresa, Lago Maggiore, Italy, p. 4.
- Dewey, S.A., Price, K.P., Ramsey, D., 1991. Satellite remote-sensing to predict potential distribution of Dyers woad (*Isatis tinctoria*). *Weed Technology* 5, 479–484.
- Elmore, A.J., Mustard, J.F., Hamburg, S.P., Manning, S.J., 2006. Agricultural legacies in the great basin alter vegetation cover, composition, and response to precipitation. *Ecosystems* 9, 1231–1241.
- Fisher, J.L., Mustard, J.F., 2007. Cross-scalar satellite phenology from ground, Landsat, and MODIS data. *Remote Sensing of Environment* 109, 261–273.
- Franklin, J., 1995. Predictive vegetation mapping: geographic modelling of biospatial patterns in relation to environmental gradients. *Progress in Physical Geography* 19, 474–499.
- Franklin, J., 2009. *Mapping Species Distributions Spatial Inference and Prediction*. Cambridge University Press, Cambridge, UK, 320 pp.
- Gillespie, T.W., Foody, G.M., Rocchini, D., Giorgi, A.P., Saatchi, S., 2008. Measuring and modelling biodiversity from space. *Progress in Physical Geography* 32, 203–221.
- Guisan, A., Zimmermann, N.E., 2000. Predictive habitat distribution models in ecology. *Ecological Modelling* 135, 147–186.
- Herfindal, I., Solberg, E.J., Saether, B.E., Hogda, K.A., Andersen, R., 2006. Environmental phenology and geographical gradients in moose body mass. *Oecologia* 150, 213–224.
- Huang, C.Y., Geiger, E.L., 2008. Climate anomalies provide opportunities for large-scale mapping of non-native plant abundance in desert grasslands. *Diversity and Distributions* 14, 875–884.
- Jarnevich, C.S., Evangelista, P., Stohlgren, T.J., Morisette, J., 2011. Improving national-scale invasion maps: tamarisk in the western United States. *Western North American Naturalist* 71, 164–175.
- Kearney, M., 2006. Habitat, environment and niche: what are we modelling? *Oikos* 115, 186–191.
- Kerr, J.T., Ostrovsky, M., 2003. From space to species: ecological applications for remote sensing. *Trends in Ecology and Evolution* 18, 299–305.
- Kremen, C., Cameron, A., Moilanen, A., Phillips, S.J., Thomas, C.D., Beentje, H., Dransfield, J., Fisher, B.L., Glaw, F., Good, T.C., Harper, G.J., Hijmans, R.J., Lees, D.C., Louis, E., Nussbaum, R.A., Raxworthy, C.J., Razafimpahanana, A., Schatz, G.E., Vences, M., Vieites, D.R., Wright, P.C., Zihra, M.L., 2008. Aligning conservation priorities across taxa in Madagascar with high-resolution planning tools. *Science* 320, 222–226.
- Lefsky, M.A., Cohen, W.B., Parker, G.G., Harding, D.J., 2002. Lidar remote sensing for ecosystem studies. *Bioscience* 52, 19–30.
- Leyequien, E., Verrelst, J., Slot, M., Schaepman-Strub, G., Heitkonig, I.M.A., Skidmore, A., 2007. Capturing the fugitive: applying remote sensing to terrestrial animal distribution and diversity. *International Journal of Applied Earth Observation and Geoinformation* 9, 1–20.
- Morán-Ordóñez, A., Suárez-Seoane, S., Elith, J., Calvo, L., de Luis, E., 2012. Satellite surface reflectance improves habitat distribution mapping: a case study on heath and shrub formations in the Cantabrian Mountains (NW Spain). *Diversity and Distributions* 18, 588–602.
- Morisette, J.T., Jarnevich, C.S., Ullah, A., Cai, W.J., Pedelty, J.A., Gentle, J.E., Stohlgren, T.J., Schnase, J.L., 2006. A tamarisk habitat suitability map for the continental United States. *Frontiers in Ecology and the Environment* 4, 11–17.
- Morisette, J.T., Richardson, A.D., Knapp, A.K., Fisher, J.L., Graham, E.A., Abatzoglou, J., Wilson, B.E., Breshears, D.D., Henebry, G.M., Hanes, J.M., Liang, L., 2009. Tracking the rhythm of the seasons in the face of global change: phenological research in the 21st century. *Frontiers in Ecology and the Environment* 7, 253–260.
- Nagendra, H., 2001. Using remote sensing to assess biodiversity. *International Journal of Remote Sensing* 22, 2377–2400.
- Olsson, A.D., Betancourt, J.L., Crimmins, M.A., Marsh, S.E., 2012. Constancy of local spread rates for buffelgrass (*Pennisetum ciliare* L.) in the Arizona Upland of the Sonoran Desert. *Journal of Arid Environments*, <http://dx.doi.org/10.1016/j.jaridenv.2012.06.005>.
- Osborne, P.E., Alonso, J.C., Bryant, R.G., 2001. Modelling landscape-scale habitat use using GIS and remote sensing: a case study with great bustards. *Journal of Applied Ecology* 38, 458–471.
- Padalia, H., Bharti, R.R., Pundir, Y.P.S., Sharma, K.P., 2010. Geospatial multiple logistic regression approach for habitat characterization of scarce plant population: a case study of *Pittosporum eriocarpum* Royle (an endemic species of Uttarakhand, India). *Journal of the Indian Society of Remote Sensing* 38, 513–521.
- Pearson, R.G., Dawson, T.P., 2003. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecology and Biogeography* 12, 361–371.
- Pearson, R.G., Dawson, T.P., Liu, C., 2004. Modelling species distributions in Britain: a hierarchical integration of climate and land-cover data. *Ecography* 27, 285–298.
- Peterson, A.T., Sanchez-Cordero, V., Martinez-Meyer, E., Navarro-Sigüenza, A.G., 2006. Tracking population extirpations via melding ecological niche modeling with land-cover information. *Ecological Modelling* 195, 229–236.
- Peterson, E.B., 2005. Estimating cover of an invasive grass (*Bromus tectorum*) using tobit regression and phenology derived from two dates of Landsat ETM plus data. *International Journal of Remote Sensing* 26, 2491–2507.
- Pettorelli, N., Ryan, S., Mueller, T., Bunnefeld, N., Jedrzejewska, B., Lima, M., Kausrud, K., 2011. The Normalized Difference Vegetation Index (NDVI): unforeseen successes in animal ecology. *Climate Research* 46, 15–27.
- Pettorelli, N., Vik, J.O., Mysterud, A., Gaillard, J.M., Tucker, C.J., Stenseth, N.C., 2005. Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends in Ecology and Evolution* 20, 503–510.
- Prates-Clark, C.D., Saatchi, S.S., Agosti, D., 2008. Predicting geographical distribution models of high-value timber trees in the Amazon Basin using remotely sensed data. *Ecological Modelling* 211, 309–323.
- Resasco, J., Hale, A.N., Henry, M.C., Gorchov, D.L., 2007. Detecting an invasive shrub in a deciduous forest understory using late-fall Landsat sensor imagery. *International Journal of Remote Sensing* 28, 3739–3745.
- Rew, L.J., Maxwell, B.D., Aspinall, R., 2005. Predicting the occurrence of nonindigenous species using environmental and remotely sensed data. *Weed Science* 53, 236–241.
- Rogers, D.J., Randolph, S.E., Snow, R.W., Hay, S.I., 2002. Satellite imagery in the study and forecast of malaria. *Nature* 415, 710–715.
- Saatchi, S., Buermann, W., Ter Steege, H., Mori, S., Smith, T.B., 2008. Modeling distribution of Amazonian tree species and diversity using remote sensing measurements. *Remote Sensing of Environment* 112, 2000–2017.
- Sellers, J.D., Jolls, C.L., 2007. Habitat modeling for *Amaranthus pumilus*: an application of light detection and ranging (LIDAR) data. *Journal of Coastal Research* 23, 1193–1202.

- Stöckli, R., Vidale, P.L., 2004. European plant phenology and climate as seen in a 20-year AVHRR land-surface parameter dataset. *International Journal of Remote Sensing* 25, 3303–3330.
- Stohlgren, T.J., Ma, P., Kumar, S., Rocca, M., Morissette, J.T., Jarnevich, C.S., Benson, N., 2010. Ensemble habitat mapping of invasive plant species. *Risk Analysis* 30, 224–235.
- Tan, B., Morissette, J.T., Wolfe, R.E., Gao, F., Ederer, G.A., Nightingale, J., Pedelty, J.A., 2011. An enhanced TIMESAT algorithm for estimating vegetation phenology metrics from MODIS data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 4, 361–371.
- Thomas, C.D., Cameron, A., Green, R.E., Bakkenes, M., Beaumont, L.J., Collingham, Y.C., Erasmus, B.F.N., de Siqueira, M.F., Grainger, A., Hannah, L., Hughes, L., Huntley, B., van Jaarsveld, A.S., Midgley, G.F., Miles, L., Ortega-Huerta, M.A., Peterson, A.T., Phillips, O.L., Williams, S.E., 2004. Extinction risk from climate change. *Nature* 427, 145–148.
- Thuiller, W., Araujo, M.B., Lavorel, S., 2004. Do we need land-cover data to model species distributions in Europe. *Journal of Biogeography* 31, 353–361.
- Thuiller, W., Richardson, D.M., Pysek, P., Midgley, G.F., Hughes, G.O., Rouget, M., 2005. Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. *Global Change Biology* 11, 2234–2250.
- Tuanmu, M.-N., Viña, A., Bearer, S., Xu, W., Ouyang, Z., Zhang, H., Liu, J., 2010. Mapping understory vegetation using phenological characteristics derived from remotely sensed data. *Remote Sensing of Environment* 114, 1833–1844.
- Turner, W., Spector, S., Gardiner, N., Fladeland, M., Sterling, E., Steininger, M., 2003. Remote sensing for biodiversity science and conservation. *Trends in Ecology and Evolution* 18, 306–314.
- 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.
- Willems, E.P., Barton, R.A., Hill, R.A., 2009. Remotely sensed productivity, regional home range selection, and local range use by an omnivorous primate. *Behavioral Ecology* 20, 985–992.
- Woodcock, C.E., Allen, R., Anderson, M., Belward, A., Bindschadler, R., Cohen, W., Gao, F., Goward, S.N., Helder, D., Helmer, E., Nemani, R., Oreopoulos, L., Schott, J., Thenkabail, P.S., Vermote, E.F., Vogelmann, J., Wulder, M.A., Wynne, R., Landsat Sci. T., 2008. Free access to Landsat imagery. *Science* 320, 1011.
- Xie, Y., Sha, Z., Yu, M., 2008. Remote sensing imagery in vegetation mapping: a review. *Journal of Plant Ecology* 1, 9–23.
- Zimmermann, H., Von Wehrden, H., Damascos, M.A., Bran, D., Welk, E., Renison, D., Hensen, I., 2011. Habitat invasion risk assessment based on Landsat 5 data, exemplified by the shrub *Rosa rubiginosa* in southern Argentina. *Austral Ecology* 36, 870–880.
- Zimmermann, N.E., Edwards, T.C., Moisen, G.G., Frescino, T.S., Blackard, J.A., 2007. Remote sensing-based predictors improve distribution models of rare, early successional and broadleaf tree species in Utah. *Journal of Applied Ecology* 44, 1057–1067.