2015

Will Remote Sensing Shape the Next Generation of Species Distribution Models?

Kate S. He
Murray State University, xhe@murraystate.edu

Bethany A. Bradley
University of Massachusetts Amherst, bbradley@eco.umass.edu

Anna F. Cord
Helmholtz Centre for Environmental Research

Duccio Rocchini
Fondazione Edmund Mach, S. Michele all’Adige

Mao-Ning Tuanmu
Yale University

See next page for additional authors

Follow this and additional works at: https://scholarworks.umass.edu/nrc_faculty_pubs

Part of the Natural Resources and Conservation Commons

Recommended Citation
He, Kate S.; Bradley, Bethany A.; Cord, Anna F.; Rocchini, Duccio; Tuanmu, Mao-Ning; Schmidtlein, Sebastian; Turner, Woody; Wegmann, Martin; and Pettorelli, Nathalie, "Will Remote Sensing Shape the Next Generation of Species Distribution Models?" (2015). Remote Sensing in Ecology and Conservation. 376.
https://doi.org/10.1002/rse2.7

This Article is brought to you for free and open access by the Environmental Conservation at ScholarWorks@UMass Amherst. It has been accepted for inclusion in Environmental Conservation Faculty Publication Series by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.
INTERDISCIPLINARY PERSPECTIVES

Will remote sensing shape the next generation of species distribution models?

Kate S. He1, Bethany A. Bradley2, Anna F. Cord3, Duccio Rocchini4, Mao-Ning Tuanmu5, Sebastian Schmidtlein6, Woody Turner7, Martin Wegmann8&9 & Nathalie Pettorelli10

1Department of Biological Sciences, Murray State University, Murray, Kentucky 42071, USA
2Department of Environmental Conservation, University of Massachusetts, Amherst, Massachusetts 01003, USA
3Department of Computational Landscape Ecology, Helmholtz Centre for Environmental Research – UFZ, Permoserstraße 15, 04318, Leipzig, Germany
4GIS and Remote Sensing Unit, Department of Biodiversity and Molecular Ecology, Research and Innovation Center, Fondazione Edmund Mach, Via E. Mach 1, S. Michele all’Adige, Trento 38010, Italy
5Department of Ecology and Evolutionary Biology, Yale University, New Haven, Connecticut 06520, USA
6Institute of Geography and Geocology, Karlsruhe Institute of Technology (KIT), 76131 Karlsruhe, Germany
7Earth Science Division, NASA, Washington, District of Columbia, USA
8Department of Remote Sensing, University of Wuerzburg, 97074, Wurzburg, Germany
9CEOS Biodiversity at German Remote Sensing Data Centre, German Aerospace Centre DLR, 82234 Wessling, Germany
10Institute of Zoology, Zoological Society of London, Regent’s Park, London, NW1 4RY, UK

Keywords
Ecological niche modeling, habitat suitability modeling, hyperspectral and multispectral data, LiDAR and RADAR metrics, predictor and response variables, spatial and temporal resolution

Abstract
Two prominent limitations of species distribution models (SDMs) are spatial biases in existing occurrence data and a lack of spatially explicit predictor variables to fully capture habitat characteristics of species. Can existing and emerging remote sensing technologies meet these challenges and improve future SDMs? We believe so. Novel products derived from multispectral and hyperspectral sensors, as well as future Light Detection and Ranging (LiDAR) and RADAR missions, may play a key role in improving model performance. In this perspective piece, we demonstrate how modern sensors onboard satellites, planes and unmanned aerial vehicles are revolutionizing the way we can detect and monitor both plant and animal species in terrestrial and aquatic ecosystems as well as allowing the emergence of novel predictor variables appropriate for species distribution modeling. We hope this interdisciplinary perspective will motivate ecologists, remote sensing experts and modelers to work together for developing a more refined SDM framework in the near future.

Introduction
Over the past two decades, a tremendous amount of work has been undertaken to map species’ distributions and use the collected information to identify suitable habitats (Austin, 2002; Araújo et al. 2005; Franklin 2010). An array of sophisticated modeling tools are available to ecologists interested in predicting species occurrence (Elith and Leathwick 2009; Kissling et al. 2012) and species distribution models (SDMs) are now commonly used for pursuing diverse research endeavors, such as testing ecological theories (e.g. Petitpierre et al. 2012); predicting species range dynamics in response to environmental change (e.g. Schurr et al. 2012; Fordham et al. 2013, 2014; Dolos et al. 2015);
assessing invasion risks of introduced species (e.g. Bradley et al. 2009); and facilitating the design and selection of nature reserves (e.g. Kremen et al. 2008).

In most of the SDMs published in the last decade, the response variable (species occurrence data) is derived from herbaria or atlases, whereas predictor variables are mostly derived from spatially interpolated data (e.g. climate variables of climate research unit (CRU), New et al. 2002 and Worldclim, Hijmans et al. 2005), or categorical data (e.g. land cover and vegetation type). Occurrence data derived from remote sensing technology have started to be used in SDM studies (e.g. Bradley and Mustard 2006; Andrew and Ustin 2009), yet the utilities of remotely derived occurrence or abundance data remain largely unexplored. Environmental predictor variables derived from remote sensing data are more common in SDMs; this is particularly true when thinking of topographical information derived from the Shuttle Radar Topography Mission (SRTM) (see examples in Franklin 2010) and land cover maps (e.g. Pearson et al. 2004; Thuiller et al. 2004; Luoto et al. 2007; Newton-Cross et al. 2007; Morán-Ordóñez et al. 2012; Rickbeil et al. 2014). Continuous remotely sensed metrics as predictors of habitat condition, such as the normalized different vegetation index (NDVI) and leaf area index (LAI), both effective proxies for vegetation productivity (Zimmermann et al. 2007; Buermann et al. 2008; Pettorelli 2013), are still relatively under used. Yet, these and other remotely sensed products are becoming increasingly available for ecological analyses. We believe that continuous remote sensing metrics have become an integral part of SDM studies and will contribute significant amount of spatially explicit data for multi-scales and multi-taxon distribution models given recent development in remote sensing technologies and products.

Here, we describe examples of response and predictor variables derived from remote sensing that could provide novel information for species distribution modeling. We focus our attention on spaceborne and airborne systems, targeting both passive and active sensors. Passive sensors considered in this work range from panchromatic (e.g. high-resolution aerial photography with a single grayscale spectral band) to multispectral (e.g. moderate resolution sensors like Landsat collecting information in 4–11 bands) and hyperspectral (e.g. airborne high to moderate resolution data from AVIRIS with over one hundred narrow spectral bands). Active sensors include laser-light remote sensing Light Detection and Ranging (LiDAR) and microwave RADARs. Specific information on these sensors and others is provided in Table 1. We demonstrate how remotely derived variables have helped improve our understanding of species distribution over the past decade with a few case studies, while pointing out the uncertainty and constraints related to the use of remote sensing variables in SDMs. Lastly, we discuss how new technologies and products may shape the next generations of SDMs (NG-SDMs).

Remote Sensing of Species Distributions: The Response Variable

In SDMs, presence data are the most common response variable, with presence/absence or abundance data only occasionally available (Elith and Leathwick 2009). Occurrence records are generally derived from herbarium and museum collections, national atlases, large-scale field surveys, regional checklists, expert range maps and collections from citizen science groups (Jetz et al. 2012). However, these data can be associated with a variety of limitations, including sampling biases, inaccuracies in geo-referencing and taxonomy (Dickinson et al. 2010). Species occurrence data such as presence and absence records from museum and herbarium collections and field sampling can indeed be quite biased. This can sometimes be traced back to the distribution of collection sites, with some sites being under-sampled due to accessibility and other logistics issues. Reliable species absence data can be even more problematic to acquire since some species can be present in the considered site, but undetected. As demonstrated below, these limitations can be overcome in certain cases by using remotely derived species occurrence records.

Plants

Remote detection of plant species is most likely to be viable if the target plant species has a unique growth form or pheno-ology. Many ecologists are familiar with global or national land cover classifications derived from satellite reflectance data (e.g. Friedl et al. 2002). Even with a few spectral bands, it is possible to separate functional types of vegetation (i.e. grasslands, forests, deserts, salt marshes, etc.) across broad spatial extents (He et al. 2009). A similar approach could enable species-level detection in cases where the target plant is the dominant form or a homogenous stand. For example dominant tree species in shrublands or grasslands have been identified based on unique vegetation index time series signatures (extracted from MODIS; Morissette et al. 2006) as well as through object-based identification of tree crowns given high enough spatial resolution (based on aerial photos; Weisberg et al. 2007). In a perennial shrubland, invasive annual grasses were detectable using Landsat imagery (Peterson 2005).

In addition to identifying distinct plant functional types through growth form, multispectal remote sensing can be used to identify plants with unique phenologies. This approach has been used most often to identify invasive plants (Bradley 2014). For example inter-annual variability in phenology has been used to identify annual grasses
in desert ecosystems, including cheatgrass (*Bromus tectorum*) (Bradley and Mustard 2005) and Lehmann lovegrass (*Eragrostis lehmanniana*) (Huang and Geiger 2008). Early growth and late senescence has been used to map dominant forest understory species including two bamboo species (*Bashania faberi* and *Fargesia robusta*) (Tuanmu et al. 2010) and honeysuckle (*Lonicera maackii*) (e.g. Wilfong et al. 2009).

The previous examples of broad-scale plant detection rely on unique functional or phenological properties. But,

<table>
<thead>
<tr>
<th>Platform and sensor</th>
<th>Spatial resolution (pan)</th>
<th>Spatial resolution (multi)</th>
<th>Spatial resolution (thermal)</th>
<th>Swath width</th>
<th>Revisiting time (theoretical maximum) Temporal availability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passive sensors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multispectral</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Worldview (-3)</td>
<td>0.31 m</td>
<td>1.24 m</td>
<td>13 km</td>
<td>1–4 days</td>
<td>2009(14)-present</td>
</tr>
<tr>
<td>Quickbird</td>
<td>0.61 m</td>
<td>2.4 m</td>
<td>17 km</td>
<td>1–3 days</td>
<td>2001-present</td>
</tr>
<tr>
<td>Pleiades</td>
<td>0.7 m</td>
<td>2 m</td>
<td>20–120 km</td>
<td>Daily</td>
<td>2000-present</td>
</tr>
<tr>
<td>Ikonos</td>
<td>0.82 m</td>
<td>3.2 m</td>
<td>11 km</td>
<td>1–3 days</td>
<td>1999-present</td>
</tr>
<tr>
<td>TopSAT</td>
<td>2.8 m</td>
<td>5.6 m</td>
<td>10–15 km</td>
<td>Daily</td>
<td>2005-present</td>
</tr>
<tr>
<td>RapidEye</td>
<td>6.5 m</td>
<td></td>
<td>77 km</td>
<td>5.5 days</td>
<td>2008-present</td>
</tr>
<tr>
<td>RapidEye+</td>
<td>&lt;1 m</td>
<td>Similar to RapidEye</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPOT (5)</td>
<td>5 m</td>
<td>10–20 m</td>
<td>60 km</td>
<td>2–3 days</td>
<td>1986 (2002)-present</td>
</tr>
<tr>
<td>SENTINEL 2A</td>
<td>10–60 m</td>
<td></td>
<td>290 km</td>
<td>5 days</td>
<td>Launch in 2015</td>
</tr>
<tr>
<td>ASTER</td>
<td>15–30 m</td>
<td>90 m</td>
<td>60 km</td>
<td>16 days</td>
<td>2000-present</td>
</tr>
<tr>
<td>CBERS</td>
<td>20 m</td>
<td>20–260 m</td>
<td>80 m</td>
<td>120–890 km</td>
<td>26 days</td>
</tr>
<tr>
<td>Landsat</td>
<td>30 m</td>
<td>30 m</td>
<td>185 km</td>
<td>16 days</td>
<td>1982-present</td>
</tr>
<tr>
<td>TM 4/5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Landsat 7/8</td>
<td>15 m</td>
<td>30 m</td>
<td>60/100 m</td>
<td>185 km</td>
<td>16 days</td>
</tr>
<tr>
<td>MODIS</td>
<td>250–1000 m</td>
<td>1000 m</td>
<td>2330 km</td>
<td>1–2 days</td>
<td>2000-present</td>
</tr>
<tr>
<td>AVHRR</td>
<td>1090 m</td>
<td>1090 m</td>
<td>2600 km</td>
<td>Daily</td>
<td>1978-present</td>
</tr>
<tr>
<td>Hyperion</td>
<td>30 m</td>
<td></td>
<td>7.7 km</td>
<td>Tasked</td>
<td>2000-present</td>
</tr>
<tr>
<td>HyMap</td>
<td></td>
<td></td>
<td>Spatial resolution depending on flight altitude (c. 3–20 m), availability on request, up to several hundred bands</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HySpex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AVIRIS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EnMAP</td>
<td>Hyspectral space-borne mission by DLR, c. 30 m spatial resolution, launch planned before c. 2020</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OMI</td>
<td>Hyspectral space-borne mission for atmospheric parameters by NIrV and FMI, planned with 13–24 km spatial resolution</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Platform and sensor</th>
<th>Spatial resolution</th>
<th>Revisiting time (theoretical maximum)</th>
<th>Temporal availability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active sensors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COSMO-Skymed</td>
<td>1–15 m</td>
<td>1–15 days</td>
<td>2007-present</td>
</tr>
<tr>
<td>TerraSAR-X</td>
<td>1–18 m</td>
<td>2.5 days</td>
<td>2007-present</td>
</tr>
<tr>
<td>Tandem-X</td>
<td>1–18 m</td>
<td>2.5 days</td>
<td>2010-present</td>
</tr>
<tr>
<td>RADARSAT-2</td>
<td>3–100 m</td>
<td>24 days</td>
<td>2007-present</td>
</tr>
<tr>
<td>Sentinel 1</td>
<td>5–40 m</td>
<td>c. 12 days</td>
<td>2014-present</td>
</tr>
<tr>
<td>ENVISAT</td>
<td>30–1000 m</td>
<td>35 days</td>
<td>2002–2012</td>
</tr>
<tr>
<td>LiDAR</td>
<td>Airborne, availability on request</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GEDI</td>
<td>Space-borne LiDAR, planned</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRMM</td>
<td>5 km</td>
<td>16 times per day</td>
<td>1997-present</td>
</tr>
<tr>
<td>GPM</td>
<td>5 km</td>
<td>c. 3 hours</td>
<td>2014-present</td>
</tr>
<tr>
<td>SMAP</td>
<td>1–3 km</td>
<td>2–3 days</td>
<td>2015-present</td>
</tr>
<tr>
<td>SRTM</td>
<td>Global DEM, 30–90 m spatial resolution</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASTER-DEM</td>
<td>Global DEM, 30 m spatial resolution</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TerraSAR-X/Tandem-X</td>
<td>Global DEM, 12 m spatial resolution, forthcoming</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
detection of plant species is also possible thanks to the higher thematic details provided by hyperspectral data. With over a hundred spectral bands being monitored, hyperspectral sensors can detect subtle differences in reflectance resulting from unique plant chemistries. This could help reduce misidentification and taxonomic biases found in field surveys. Numerous case studies of successful plant species detection using hyperspectral information can be found for exotic and invasive plants (Huang and Asner 2009; He et al. 2011). For example Andrew and Ustin (2008) used HyMap to identify unique white flowers of invasive pepperweed (Lepidium latifolium) near Sacramento, California. Similarly, Mitchell and Glenn (2009) also used HyMap to identify the unique yellow bracts of invasive leafy spurge (Euphorbia esula) in south-east Idaho.

In Hawaii, a combination of differences in pigmentation and leaf water content enabled the detection of non-native trees using AVIRIS (Asner et al. 2008a). Other tree species were also successfully mapped with hyperspectral data within the tropics and subtropics (Clark et al. 2005; Carlson et al. 2007; Lucas et al. 2008; Yang et al. 2009; Féret and Asner 2013) as well as in temperate forest ecosystems (Fassnacht et al. 2014). Given sufficient expertise, effective classification algorithms and available data, many more plant species could be detectable using hyperspectral data.

LiDAR coupled with multispectral or hyperspectral data has also been used for identifying tree species (Jones et al. 2010; Heinzel and Koch 2011; Dalponte et al. 2012; Alonzo et al. 2014; Ghosh et al. 2014). This approach takes the advantage of using complementary information gathered from spectral reflectance and vertical structure of target species. Using a multi-sensor system (hyperspectral AISA, multispectral GeoEye-1, and high point density LiDAR), Dalponte et al. (2012) identified eight tree species in the Southern Alps with accuracies ranging from 76.5 to 93.2%. Similar conclusions were also made when mapping eleven tree species in coastal south-western Canada thanks to a combination of hyperspectral imagery and LiDAR data (Jones et al. 2010). In Hawaii, Asner et al. (2008b) employed a hybrid airborne system, combining the Carnegie Airborne Observatory small-footprint LiDAR system with AVIRIS to map the three-dimensional spectral and structural properties of three highly invasive trees. In this particular study, the authors separated the tree species based on their unique biophysical properties with a multi-stage spectral mixture analysis.

**Animals**

Tracking the presence of animal species using satellite remote sensing is feasible given fine enough pixel resolution and large enough animals under an unobstructed view. For example Yang et al. (2014) used both expert interpretation and an automated object-based classification to estimate populations of zebra (Equus quagga burchelli) and wildebeest (Connochaetes taurinus) and investigate their migration patterns in the open savannah of the Maasai Mara National Reserve, Kenya, thanks to very high-resolution GeoEye-1 satellite images (0.5 m resolution). Fretwell et al. (2014) identified 55 southern right whales (Eubalaena australis) in a breeding ground off the coast of Argentina based on brighter reflectance from WordView-2 (50 cm resolution). Similar approaches have been used to identify polar bears (Ursus maritimus) (Stapleton et al. 2014), walrus (Odobenus rosmarus) (Platonov et al. 2013) and emperor penguins (Aptenodytes forsteri) (Fretwell et al. 2012).

Spaceborne and airborne remote sensing can be very effective in supplementing species occurrence data (presence-absence, presence-only and point events), but getting very high-resolution remote sensing imagery is still costly in general even though no-cost imagery and open-source software for imagery processing are an increasingly common practice worldwide (Wegmann et al. 2015). At times, these high costs can be reduced by employing light-weight unmanned aerial vehicles (UAVs; Anderson and Gaston 2013).

For example UAVs mounted with off the shelf cameras and GPS were used to count marine mammals (dugong, Dugong dugon) in western Australia (Hodgson et al. 2013), along with a variety of other marine species. In a terrestrial case study, UAV images were used to identify orangutan (Pongo abelii) and elephant (Elephas maximus) in Sumatra, Indonesia (Koh and Wich 2012).

High spatial resolution remote sensing of terrestrial and marine animals is an excellent tool for measuring populations and identifying important habitat (e.g. stopovers on migratory routes, breeding grounds). Thus far, most animal detection studies have focused on a small area due to the reliance on very high spatial resolution data. However, increasingly available high-resolution imagery and inexpensive UAVs coupled with object-based identification of animals might enable much broader scale identification of animal occurrence. The use of UAVs can also be a particularly cost-efficient way to collect input data for model calibration and validation.

Similarly to species occurrence data collected from field surveys, remotely derived response variables come with uncertainty and errors. These errors are typically introduced during data acquisition and processing, and through associated analytical algorithms. Studies have used various metrics to estimate classifiers’ performance, ideally based on independent validation data, such as the Cohen’s kappa statistics, confusion matrix, F-scores, overall accuracy and the receiver operating characteristic (ROC) curve (obtained by plotting fraction of true posi-
tive against fraction of false positive). The typically acceptable accuracies range from 60 to 90% for plant species and error rates for validation or training data are not reported in most studies (He et al. 2011). Furthermore, false-positive and false-negative detections can be identified by using plot-scale surveys guided by remotely sensed data (Asner et al. 2008c). When modeling species potential ranges with presence-only data, omission errors (identifying a species as absent when it is actually present) can lead to underestimates of potential range. In contrast, commission errors (identifying a species as present when it is actually absent) can lead to overestimates of potential range. The relative amount of omission versus commission error in the initial remote sensing map can inform interpretation of a resulting SDM output.

Regardless of the relative error rates, remote sensing of plants and animals provides both presence and absence data, which is more informative for SDMs than presence-only data, a relative measure that only represents a partial estimate of species occurrence (Fithian and Hastie 2013). Recent developments in SDMs have shown that a Poisson process model can combine presence-absence and presence-only data for correcting sampling bias in SDMs (Fithian et al. 2015). Furthermore, treating presence-only data as point events and estimating the intensity of the spatial location of presence points with a point process modeling framework, may also reduce uncertainty in SDMs (Renner et al. 2015). In the latter case, remote sensing can effectively provide point event data at various spatial scales to complement survey data. Most importantly, remote sensing may provide an estimate of absolute population density rather than relative density and this can be achieved with LiDAR and RADAR systems, particularly for tree species. Lastly, we want to point out that response variables derived from remote sensing can be updated every year or at desired time span, which allows a more dynamic approach to understanding habitat suitability and species range expansion or contraction. As time series data become more readily available, phenological events in plants and animals can be tracked and linked to fine-scale species distribution studies.

Remote Sensing of Environmental Conditions: The Predictor Variable

Good distribution models require spatial predictor variables that are ecologically relevant (Franklin 1995) for the modeled organisms. In some cases, remote sensing metrics can be challenging to translate into meaningful ecological entities, particularly those that provide indirect measures of ecosystem processes (e.g. surface roughness from RADAR; Buermann et al. 2008) making it unclear why to consider them in SDMs in the first place.

Abiotic predictor variables

Climate data have been commonly used to predict potential species distributions across broad spatial scales (e.g. Franklin 2010), CRU (New et al. 2002), WorldClim (Hijmans et al. 2005), CliMond (Kriticos et al. 2012) and PRISM (US only; Daly et al. 2002), are all examples of spatially explicit datasets of climatic conditions. These datasets encompass information on modeled temperature, precipitation, solar radiation and soil moisture (along with several derived bioclimatic combinations of temperature and precipitation), which are based on interpolations from global weather stations. However, interpolations are only as good as the underlying data, and uneven geographical coverage leads to high model uncertainty, especially in developing countries where few weather stations are in place (Daly 2006; Bedia et al. 2013; Waltari et al. 2014). When uncertainty in spatial climate variables is not accounted for, coefficient estimates tend to be biased which lead to poor performances of SDMs as shown with recent simulations (Stoklosa et al. 2015).

Remotely sensed climate data are continuously observed without interpolation and geographical biases. Therefore, satellite-based temperature, precipitation and radiation measurements could improve climate predictor variables. For example land surface temperature (LST) is measured globally four times per day by the MODIS Terra and Aqua satellites (Wan et al. 2004; Sims et al. 2008) and a derived product at 250 m spatial resolution is freely available at http://gis.cri.fmach.it/eurolst for Europe (Metz et al. 2014). MODIS LST data are increasingly being used in SDMs to understand and predict ecological processes (Buermann et al. 2008; Bisrat et al. 2012; Neter et al. 2013; Pau et al. 2013; Still et al. 2014). Recently, efforts have been made to use LST to facilitate interpolation of weather station data as weather station data have a long temporal span, which cannot be fully covered by remote sensing data (Parmentier et al. 2015). In addition, the global UV-B radiation dataset from NASA Aura-OMI (Beckmann et al. 2014), designed for macroecological studies, offers exciting opportunities for both correlative and process-based species distribution modeling.

Precipitation estimates from satellite are available historically from TRMM (Tropical Rainfall Measuring Mission; Huffman et al. 2007) at a 4 km spatial resolution covering the tropical region (20°N–20°S) and extended to 50°N–50°S (Wang et al. 2014). New rainfall products are just becoming available from the Global Precipitation Measurement (GPM) mission, which has replaced TRMM (TRMM data collection stopped in 2105). With the recent launch of NASA’s Soil Moisture Active Passive (SMAP) Mission, high-resolution soil moisture data (3 and 9 km) with global coverage will also soon be available. Finally,
an analysis of global cloud cover from MODIS can serve as a proxy for average precipitation (A. Wilson and W. Jetz, pers. comm. 2015). Satellite measurements of temperature, precipitation and soil moisture may thus soon provide better wall to wall estimates of climatic conditions than weather station interpolations and are becoming increasingly accessible to ecologists for building SDMs with less uncertainty.

Topographic features of land surface derived from SRTM digital elevation data (the DEM products) and GDEM (Global Digital Elevation Map) from ASTER are already commonly used as predictor variables in SDMs (Franklin 2009). For finer-scale studies of local and micro-topography, airborne LiDAR and stereographic DEMs from WorldView-2 are both options. Very high-resolution topographic data derived from LiDAR have been incorporated in SDMs while assessing habitat suitability of eleven at-risk plant species in Hawaii (Questad et al. 2014) and for assessing diversity and composition of a temperate montane forest in Germany (Leutner et al. 2012). However, both datasets are costly and LiDAR data are currently limited in temporal coverage and spatial extent. An emerging alternative source of very high-resolution DEMs are UAVs (Anderson and Gaston 2013), which are rapidly becoming more reliable, lightweight and cost effective for LiDAR instrumentation (Watts et al. 2012).

In the marine realm, sea-surface temperature derived from Aqua MODIS (https://po-daac.usgs.gov/SeaSurfaceTemperature), with global resolutions as fine as 1 km, has been one of the most influential predictors in SDMs for identifying productivity hotspots and seascape modeling (Louzao et al. 2011; Ramírez et al. 2014). Furthermore, the Bio-ORACLE (ocean rasters for analyses of climate and environment at http://www.bio-oracle.ugent.be), a marine counterpart of the WorldClim database has been developed, consisting of 23 environmental rasters, derived from both satellite-based and in situ data for modeling the distribution of shallow water marine species at a global scale (Tyberghein et al. 2012). A comprehensive review on using remotely derived variables to inform marine habitat mapping and monitoring can be found in Kachelriess et al. (2014).

**Biotic predictor variables**

Vegetation characteristics can be important predictors of species’ habitat, acting as a proxy for sources of food availability or shelter. Many studies have used remotely sensed variables to model habitat suitability for animals, in particular using satellite-derived land cover classifications (Leyequien et al. 2007) as well as continuous metrics of vegetation productivity such as NDVI (Petorelli et al. 2011). For example NDVI data from MODIS was used as a predictor of food availability in a model of vervet monkey (Chlorocebus pygerythrus) habitat in Africa (Willems et al. 2009). Similarly, researchers used NDVI data from AVHRR to assess habitat availability of the Iberian mole (Talpa occidentalis) along a biogeographic gradient in Spain (Suárez-Seoane et al. 2014). Although being increasingly used (Bradley et al. 2012; Cord et al. 2013), incorporating remotely sensed metrics of vegetation into models of habitat suitability requires a more careful approach, particularly when it comes to plants. Bradley et al. (2012) indeed caution that the use of NDVI metrics in plant models could create biases in cases where they measure properties of the target species directly.

Vegetation structure derived from RADAR or LiDAR could also be an important predictor of habitat (Vierling et al. 2008). For example Buermann et al. (2008) used RADAR data from QuikSCAT as a proxy for Amazonian forest canopy roughness and found that it improved habitat suitability models for several species of birds. Similarly, Farrell et al. (2013) concluded that models incorporating LiDAR-derived metrics, such as tree height, improved model predictions of bird habitat in Texas.

In addition, vegetation phenology derived from satellite time series can provide important information about timing of biological events (Morisette et al. 2009) and serve as a proxy for habitat. For example NDVI-based estimate of the length of summer was an important predictor of moose (Alces alces) body weight, and therefore habitat quality, in Norway (Herfindal et al. 2006). Tuanmu et al. (2011) suggested that multi-year phenology metrics derived from MODIS can reduce model complexity and multicollinearity among predictor variables and thus improve model transferability (i.e. the ability of a model developed in one time period/area to be applied to a different time period/area) for giant panda (Ailuropoda melanoleuca) habitat change in China. Furthermore, remotely sensed seasonal variation in vegetated land surfaces can be influential predictor variables when modeling species distribution and habitat suitability (Osborne and Suárez-Seoane 2007).

Multi-year NDVI and its predicted values under climate change scenarios have been used to assess likely impacts of environmental change on future species distributions and extinction risks, which are a major motivation for SDM research. Singh and Milner-Gulland (2011) for example used 25 years of temperature and NDVI data to identify the changing drivers of migratory saiga (Saiga tatarica) distribution in Central Kazakhstan under a range of scenarios, including changes in temperature and productivity. In this study, projected NDVI values were proven as one of the critical predictors in modeling future saiga distribution and changes in population density.
Thus, if remotely sensed predictors, such as NDVI in this case, improve SDMs, then predicted extinction risks from environmental change are going to become more reliable.

Information on biotic conditions can also be derived from UAVs. For example Koh and Wich (2012) note that imagery from a UAV in Sumatra, Indonesia could detect evidence of small-scale human disturbance, including logging and local oil palm plantations. Similarly, Getzin et al. (2014) used a UAV to identify small forest gaps in Germany, which could be an important predictor of early successional species occurrence.

For a few satellite missions (e.g. Landsat), data archives are now decades long, enabling the tracking of temporal changes in ecosystems. While land use/land cover change is a long recognized discipline (e.g. Meyer and Turner 1994), including change metrics as predictors in SDMs is exceedingly rare. Yet, temporal trends in NDVI (e.g. Verbyla 2008) and other satellite measurements may be key indicators of ecological changes likely to influence the distribution of species. With MODIS archives reaching 15 years and Landsat 40, plus the recently launched Sentinel 1 and 2 missions, the combination of higher spatial and longer-term temporal analyses is increasingly possible.

**Future Opportunities**

**New space missions and sensor networks**

Remote sensing products provide dynamic information that is increasingly relevant to the fields of ecology and conservation biology (Pettorelli et al. 2014a; Turner 2014). In recent years, the potential of remote sensing to support ecological research has been boosted by the prospects of new technological developments and new space missions, including a number of very high spatial and spectral resolution passive optical satellites as well as active optical (LiDAR) and RADAR imaging systems equipped with state-of-the-art technology (Pettorelli et al. 2014b).

New optical satellite missions include the European Sentinel-2 satellites, the Pleiades of CNES, the TopSat (UK), CBERS (China and Brazil) and the Resourcesat series (India) along with a host of private sector missions seeking to offer high spatial resolution imagery of the sunlit Earth essentially everywhere at all times. Recently launched and planned RADAR and LiDAR missions include Sentinel 1 with C-Band, the TerraSAR-X/TanDEM-X mission, the RadarSat program of the Canadian Space Agency, the RADAR mission by the JAXA space agency in Japan and the NASA Global Ecosystem Dynamics Investigation (GEDI) LiDAR planned for the International Space Station (Koch 2010). Detailed information on sensors and missions is listed in Table 1. In addition, there are continued new developments in low-cost, light-weight, and long-duration UAVs (Lucieer et al. 2014). New missions and new sensors will allow mapping and monitoring of global ecosystems at an unprecedented level of detail (sub-meter spatial resolution and 3-D profiles are now possible), potentially providing invaluable data for improving the predictive power of SDMs.

**Novel predictor variables bring new possibilities**

Biophysical, biochemical and physiological predictors derived from modern remote sensing have huge potential when it comes to improving the predictive power of SDMs. Advances over the more widely used NDVI include LAI3g (third generation of LAI data with best-quality and significantly improved post-processing algorithms) and fraction of absorbed photosynthetically active radiation (fPAR, especially fPAR3g), both of which are available from MODIS data. In addition to detecting plant pigmentation, hyperspectral data can be used to measure leaf water content and leaf nitrogen content, along with other unique chemical signals. Active LiDAR and RADAR can estimate canopy/tree height and stem density, canopy moisture and 3-D habitat structural profiles (a vertical description of the habitat, such as the position of leaves, branches and ground) (Simonson et al. 2014). These and other potential predictor variables are outlined in Table 2.

With new high-resolution sensors, remotely sensed data could add insight into the spatial patterns of plant interactions at local to landscape scales. These sorts of biotic interactions are absent from SDMs due to lack of data (Kissling et al. 2012). But, hyperspectral data can be used to classify vegetation communities into plant functional types based on optical reflectance values (Ustin and Gamon 2010), creating very high-resolution maps of plant assemblages, which can provide information on interactions among species in terms of competition for water and light. Hyperspectral data have also been used to map plant communities based on competitive, stress tolerant, ruderal strategy (Schmidleinit et al. 2012) where C strategists are highly competitive, S strategists are stress tolerant and R strategists are ruderals with rapid growth and short life spans (sensu Grime 1974, 1977). At broader scales, remote measurements including fPAR, NDVI, LAI and tree/canopy height can combine to estimate overall ecological diversity (van Ewijk et al. 2014), a proxy for competition. These sorts of remote sensing products could enable assessments of intensity and spatial location of biotic interactions across thousands of hectares, much larger than current plot studies.
The NG-SDMs

Given the opportunities provided by remote sensing presented above, along with data collected from well-designed experiments, field plots and in situ sensors, the NG-SDMs could develop rapidly, building upon recent development in SDMs (Lurgi et al. 2015; Renner et al. 2015). NG-SDMs could be integrative models that (1) operate at different areas along the correlative-process model continuum (sensu Dormann et al. 2012). The majority of current SDMs either fall at one extreme end of the continuum for being correlative (with explanatory variables which may or may not be casual factors for species occurrence) or fit at the other extreme end of the continuum for being process-based (with clearly defined ecological meaningful parameters); (2) form a hierarchically nested predictive framework, allowing for assessing species distribution at multiple biological levels and spatial scales; and (3) explicitly consider biotic interactions and variation in demographic rates with a process-oriented approach driven by underlying mechanisms (Schurr et al. 2012; Kissling 2013; Wisz et al. 2013; Renner et al. 2015). We provide a comparative modeling framework between the current SDMs and the NG-SDMs proposed in this perspective in Figure 1.

The concrete contributions to the development of NG-SDMs from remote sensing could include a variety of ecologically meaningful predictors. First, ecophysiological relevant variables, such as remotely derived earth surface temperature, precipitation and MODIS phenologi-
cal metrics as discussed in previous sections, can be the basis for mapping a species’ tolerance to abiotic constraints. These variables are mostly suited for broad-scale models developed at the correlative end of the continuum for both plant and animal species. Second, demographic parameters capturing differences in life histories, such as a population’s growth rates obtained from LiDAR (e.g. tree height, stem density for both canopy and sub-canopy layers at different points in time) and species’ biophysical traits derived from hyperspectral sensors (e.g. leaf water and nitrogen contents, pigment characteristics and other biochemistry traits), provide opportunities for developing process-based models at the local scale. Third, biotic predictors, including plant functional types, fPAR (an indirect proxy for light competition and growth) and 3-D habitat structure (capable of depicting reliance among species), can be related to biotic interactions at both local and broader scales. The response variables in NG-SDMs will be multi-level in nature, and could include the presence/absence of a single taxon; species fitness metrics; trait diversity information; and occurrence or abundance of aggregated taxa, functional groups and community assemblages. Being integrative models as suggested by Lurgi et al. (2015), the NG-SDMs can handle a wide range of data types and resolutions, and model uncertainty, while being capable of revealing the underlying causal factors of shaping species distribution and abundance.

Finally, we also want to stress the limitations and challenges of remote sensing in NG-SDMs. First of all, not all plant and animal species can be detected by remote sensors. Understory species and species with lesser distinctive spectral features are difficult to detect. To this end, sensor networks and data fusion (optical/RADAR/LiDAR) may play a key role in tracking species distribution (Koch 2010; Dalponte et al. 2012). Second, there is always a trade-off between spatial and temporal resolutions and a trade-off between spatial and spectral resolution. For
example high temporal resolution data usually have low spatial resolution (such as time series of multispectral sensors, Landsat or MODIS). Low spatial resolution can hardly discriminate objects on the ground resulting in lower classification accuracy. In general, finer spatial resolution increases classification accuracy, but at the same time, smaller pixels increase spectral variance resulting in decreased spectral separability of classes (Nagendra and Rocchini 2008). Third, remote sensing data are limited by the short time span of their availability and their contributions to modeling future projections of species distributions under climate change scenarios are limited at this stage. However, current archives of remote sensing data provide important baseline information such as changes in plant physiology and phenology for future climate change studies. New sensors with high temporal resolutions will become an integral part of monitoring instrument for tracking and predicting future species distribution under global climate change. Fourth, using species distributions that have been derived from remote sensing as responses in image-based SDMs bears the risk of circularity. Even if we have an independent response and aim at using remote sensing as predictor we should consider that the response may have had an influence on reflectance. Fifth, to fully utilize remote sensing data, one needs expertise in data processing and software development. In recent years, this has been facilitated by open-source algorithms and software as well as powerful computing capacities. Lastly, accessibility to free remote sensing data with global coverage can be challenging and this is particularly true in developing countries where data processing, storage and sharing are still hampered by information technology and archiving capability (Pettorelli et al. 2014b).

To overcome these limitations and constraints, we call for (1) the creation of sensor networks and improved interoperability between remotely sensed information and in situ biological data collections, as in situ data provide powerful information for accurate imagery interpretation; (2) development of ecologically meaningful predictors and application of cross-scale approaches; and (3) targeted coordination of field campaigns and the acquisition of remote sensing data.

Conclusions

Remote sensing has been one of the most powerful approaches to provide observations of key species distribution patterns in terms of reduced time and costs. Novel analytical techniques, increasing computational capacity, enhanced sensor fusion and networking capability as well as free access to satellite data (Turner 2014) have greatly promoted the use of remote sensing in species distribution modeling and provide the opportunity to develop a novel modeling framework as we propose here, the NG-SDMs. This modeling framework will bring new possibilities for hypothesis testing and further exploration of generalized patterns of biodiversity and underlying environmental drivers in both terrestrial and aquatic ecosystems. We hope that this interdisciplinary perspective will stimulate more discussions on species distribution modeling and motivate ecologists, remote sensing experts and modelers to work together for developing a more refined modeling framework in the near future.

Acknowledgments

This paper emerged from a Plenary Symposium ‘Tracking Changes from Space: Advances of Remote Sensing in Biogeography’ at the 7th Biennial Conference of the International Biogeography Society (IBS), held in Bayreuth, Germany from 8 to 12 January 2015. The symposium organizers, Kate S. He, Anna F. Cord and Mao-Ning Tuanmu, thank the IBS for making the symposium possible. We are also very fortunate to have had very thorough and insightful reviewers; the current paper has benefited greatly from their comments and suggestions.

Conflict of Interest

None declared.

References


Ficetola, G. F., A. Bonardi, C. A. M. 2015 The Authors

K. S. He et al. Remote Sensing and Species Distribution Models


combined-sensor precipitation estimates at fine scales. J. Hydrometeorol. 8:38–55.
Osborne, P., and S. Suárez-Seoane. 2007. Identifying core areas in a species’ range using temporal suitability analysis: an


