



## Note

# Landscape Capability Predicts Upland Game Bird Abundance and Occurrence

ZACHARY G. LOMAN,<sup>1</sup> *Department of Wildlife, Fisheries and Conservation Biology, University of Maine, 5755 Nutting Hall, Orono, ME 04469, USA*  
 ERIK J. BLOMBERG, *Department of Wildlife, Fisheries and Conservation Biology, University of Maine, 5755 Nutting Hall, Orono, ME 04469, USA*  
 WILLIAM V. DELUCA, *Department of Environmental Conservation, University of Massachusetts, Amherst, 160 Holdsworth Way, Amherst, MA 01003, USA*  
 DANIEL J. HARRISON, *Department of Wildlife, Fisheries and Conservation Biology, University of Maine, 5755 Nutting Hall, Orono, ME 04469, USA*  
 CYNTHIA S. LOFTIN, *U.S. Geological Survey, Maine Cooperative Fish and Wildlife Research Unit, Orono, ME, USA*  
 PETRA B. WOOD, *U.S. Geological Survey, West Virginia Cooperative Fish and Wildlife Research Unit, West Virginia University, Morgantown, WV, USA*

**ABSTRACT** Landscape capability (LC) models are a spatial tool with potential applications in conservation planning. We used survey data to validate LC models as predictors of occurrence and abundance at broad and fine scales for American woodcock (*Scolopax minor*) and ruffed grouse (*Bonasa umbellus*). Landscape capability models were reliable predictors of occurrence but were less indicative of relative abundance at route (11.5–14.6 km) and point scales (0.5–1 km). As predictors of occurrence, LC models had high sensitivity (0.71–0.93) and were accurate (0.71–0.88) and precise (0.88 and 0.92 for woodcock and grouse, respectively). Models did not predict point-scale abundance independent of the ability to predict occurrence of either species. The LC models are useful predictors of patterns of occurrences in the northeastern United States, but they have limited utility as predictors of fine-scale or route-specific abundances. © 2017 The Wildlife Society.

**KEY WORDS** American woodcock, bird-habitat modeling, *Bonasa umbellus*, Designing Sustainable Landscapes, landscape ecology, Northeast Region, ruffed grouse, *Scolopax minor*.

Maximizing effectiveness of conservation planning requires an understanding of the capability of landscapes to support animals. To accomplish this goal and inform landscape conservation design, the North Atlantic Landscape Conservation Cooperative has supported the Designing Sustainable Landscapes project to develop regional, species-specific models of landscape capability (LC) throughout the northeastern United States (McGarigal et al. 2016). Landscape capability models have potential for use in conservation planning (Will et al. 2005) because they are broad-scale models that explicitly map species-habitat relationships and independently model climate suitability, habitat capability, and prevalence (McGarigal et al. 2016). Landscape capability should not be confused with habitat suitability indices (HSI). Unlike HSI, LC is a measure of the capability of the landscape to provide suitable and accessible habitat and climate that supports multiple home ranges of a focal organism for a given life-history stage (McGarigal et al. 2016). This definition may be a better guide for conservation planning because it explicitly incorporates information about relevant scales for focal organisms, which are typically species-dependent (Mitchell et al. 2001), and measures landscapes' abilities to support local populations rather than

individuals. Landscape capability models do not require demographic information, which is typically difficult and expensive to collect, and do not model source-sink dynamics, thereby creating output closer to presence-absence data, and facilitating validation (Bonnot et al. 2011).

A goal of the Designing Sustainable Landscapes project is to develop models for species of conservation concern to be used as surrogates for species that share similar habitat relationships. Two such species are American woodcock (*Scolopax minor*) and ruffed grouse (*Bonasa umbellus*), which, because of their associations with successional vegetation (Dessecker and McAuley 2001), share habitat requirements with a host of species associated with young forests (Dettmers 2003, Fuller and DeStefano 2003). American woodcock are of conservation concern; eastern populations have undergone long-term declines of 0.93/year during 1968–2016 (Seamans and Rau 2016), and are a harvested migratory gamebird managed by the United States Fish and Wildlife Service (USFWS). Ruffed grouse have experienced similar regional declines during 1985–2010 (Backs and Castrale 2010), and are also a conservation priority as a popular state-managed game bird that imparts regional economic benefits (Knoche and Lupi 2013).

Validation is a critical component of the modeling process and ensures model predictions reliably inform conservation planning. Ruffed grouse and American woodcock are poorly sampled by surveys targeting breeding songbirds, partly because the breeding period for both species occurs earlier than for most songbirds (Rusch et al. 2000,

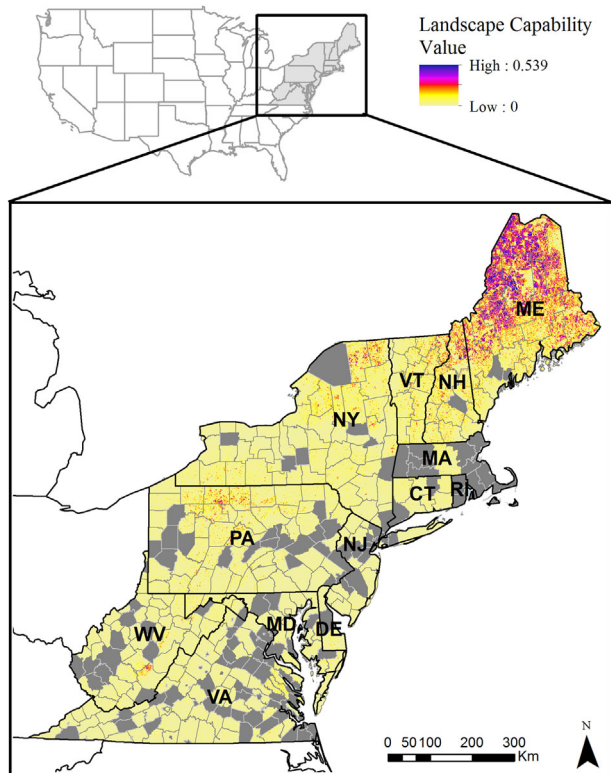
Received: 22 August 2016; Accepted: 24 February 2017

<sup>1</sup>E-mail: zachary.loman@maine.edu

McAuley et al. 2013). Ruffed grouse often are surveyed with roadside drumming surveys (Hansen et al. 2011); however, as a non-migratory bird managed by individual state wildlife agencies, there are no regionally coordinated surveys, and state-specific survey methods vary. American woodcock have been systematically surveyed since 1968 by the USFWS's Singing-Ground Survey (SGS), which uses roadside surveys to monitor singing male woodcock during their spring display period (Cooper and Rau 2014). Our objective was to validate LC models, developed under the Designing Sustainable Landscapes project, as predictors of occurrence and abundance of ruffed grouse and American woodcock using species-specific survey methods while explicitly considering relevant spatial and temporal scales.

## STUDY AREA

Our study area was the extent of the LC models, which is the 13-state Northeast Region, USA, of the USFWS (Fig. 1). This region covered 614,235 km<sup>2</sup> of land area, roughly 6.7% of the total United States, and had a human population of 69,175,320 as of 2010 (U.S. Census Bureau 2012). This region has a strongly seasonal temperate climate with mean monthly maximum daily temperatures in Virginia sometimes exceeding 32°C in August and as low as -9.5°C in winter in Maine (National Oceanic and Atmospheric Administration



**Figure 1.** States and counties in the Northeast United States examined in this study. Color gradient shows American woodcock landscape capability (LC), a spatial model reflecting the current capability of the landscape to provide accessible and suitable habitat at biologically relevant scales within counties used in this study; we obtained LC values from McGarigal et al. (2016). Counties in gray had no usable Singing-Ground Survey route data and were only used for the eBird validation of the LC model.

2017). This region includes many mesic habitat macrogroups including urban development, agriculture, alpine forest, boreal upland forest, northern mixed hardwood and conifer forest, central and southern oak-pine (*Quercus-Pinus*) forests, longleaf pine (*Pinus palustris*) forest, cliff and talus slopes, coastal grassland and shrublands, glade, barren and savanna habitats, and many hydric habitat macrogroups including swamps, peatlands, marshes, and bottomland and floodplain forests (Gawler 2008). We limited data used in the study to 2005–2015.

## METHODS

McGarigal et al. (2016) developed LC models to incorporate independent models of habitat capability and climate suitability. Habitat capability projects the quantity, quality, and accessibility (via spatial configuration) of habitat across a landscape. Importantly, it assesses each cell within the landscape at an appropriate home-range scale for the species, accumulating local resource values and then returning a summarized value to the focal cell (McGarigal et al. 2016). Landscape capability model outputs are spatially fine-scale (30 m × 30 m grid) and temporally broad-scale (~5–10 yr). We tested LC outputs created by McGarigal et al. (2016) as predictors of metrics of ruffed grouse and American woodcock breeding distribution for both broad and fine spatial scales. For point-level analyses, we treated survey data as counts to test relative abundance and as presence-absence data to test occurrence. We considered models validated if we found a significant positive relationship between models and abundance or occurrence (Brooks 1997). We included several model performance descriptors in our evaluations (Table 1).

We used the New York State Department of Environmental Conservation's (NYSDEC) Ruffed Grouse Drumming Survey (NYSDEC 2015), collected by turkey (*Meleagris gallopavo*) hunters and covering the majority of New York. We also used survey data collected by the Maine Department of Inland Fisheries and Wildlife and the New Hampshire Fish and Game Department, with each dataset covering representative portions of those states. We used American woodcock data collected from SGS routes and eBird data from throughout the USFWS Northeast Region (Sullivan et al. 2009). We conducted analyses using data derived from studies of vertebrate animals conducted by others as part of state and federal agency monitoring programs, and not collected by the authors directly. The authors assert that the original data collection followed protocols and guidelines related to use of vertebrate animals in effect at the time data were collected.

### Regional Scales

We used broadly aggregated data for analyses of the ability of LC models to explain variation in regional abundance. For ruffed grouse, we used data collected by NYSDEC from hunter reports during the turkey hunting season, the last week in April through May during 2007–2015. These data are numbers of ruffed grouse flushed per hour of hunting effort aggregated within each Wildlife Management Unit (WMU; NYSDEC 2016). The surveys provide an index of spring grouse distribution and abundance independent

**Table 1.** Predictive performance statistics of the best-fit woodcock occurrence model at the point level from Singing-Ground Survey (SGS) data collected in the northeastern United States, 2005–2014, from eBird data for ruffed grouse and American woodcock, and from ruffed grouse road surveys in New Hampshire and Maine, USA, 2014–2015, using 2 buffer widths around survey routes, 550 m and 2 km, to calculate landscape capability scores used in predicting occurrence.

Statistic	SGS woodcock (500 m)	Roadside grouse (550 m)	Roadside grouse (2 km)	eBird woodcock	eBird grouse
$\kappa^a$	0.68	0.42	0.13	0.66	0.72
Precision or positive predictive value	0.28	0.92	0.90	0.87	0.89
Negative predictive value	0.94	0.67	0.50	0.78	0.81
Specificity	0.71	0.95	0.61	0.86	0.90
Sensitivity	0.71	0.93	0.85	0.79	0.78
False positive rate (type I error)	0.29	0.33	0.39	0.14	0.10
False negative rate (type II error)	0.29	0.07	0.15	0.21	0.22
Accuracy	0.71	0.88	0.71	0.83	0.84
Prevalence	0.14	0.80	0.80	0.54	0.50
Markedness	0.22	0.63	0.40	0.65	0.70

<sup>a</sup> Kappa is an unreliable measure with skewed prevalence.

of harvest. We used data collected within 77 of the 92 WMUs ( $\bar{x} \pm \text{SE}$  hunter hours/WMU =  $500.5 \pm 57.4$  from  $130.6 \pm 14.7$  reports) that encompass the entirety of New York State, using those with  $>80$  hours of hunter effort. We used a generalized linear model (GLM) to test if mean LC values within each WMU explained observed grouse abundance. We made comparisons using detections/unit effort (grouse/hunter hr) as rank and continuous data.

We tested the ability of the LC models to explain variation in broadly aggregated regional American woodcock abundance, using SGS data to compare observed woodcock counts to the LC model prediction. We used data collected during 2005–2015. Each survey consisted of 10 stops along a 6-km survey route along secondary roads, during which an observer would count calling male woodcock during a 2-minute count period (Cooper and Rau 2014). We excluded routes from our analysis that were designated constant zero routes by USFWS at any time during the 2005–2015 time window because these were not surveyed and thus provided no data (Sauer and Bortner 1991). We aggregated route-level data within 275 counties (Fig. 1), and we used a generalized linear mixed model (GLMM) approach to test whether mean LC values within the respective county explained observed woodcock counts (log transformed) per route. We used county as the unit of aggregation, because not all routes were spatially referenced; however, USFWS provided the county for each route. We included other variables in models likely to affect observed counts including survey, ordinal date (linear and quadratic effects), and cloud cover. All models included 2 random intercept terms for county and survey route to address uneven sampling among counties. Our candidate model set included 15 models (Table A1, available online in Supporting Information).

We used eBird data to provide a regional assessment of each species-specific LC model's ability to broadly predict occurrence for both species. This component of the validation was not completely independent because the LC model uses eBird data in the climate niche portion of Designing Sustainable Landscapes models. However, the climate niche envelope is very broad-scale and with low spatial resolution, and it provides little fine-scale predictive ability, which is what we used eBird for in our validation

process. We used observations during 2005–2012, between 0500 and 2000 hours, with  $<3$  hours duration and during the breeding season for each species (15 Apr–30 Jun; Fink et al. 2010). We constrained selection of observations so that we used only 1 random location within a 250-m radius. Woodcock had 1,129 detections and 958 non-detections, and grouse had 328 detections and 327 non-detections. We made a 25-km grid of the study area, and weighted each observation by the reciprocal of the number of sample points in its respective cell, to reduce broad sampling heterogeneity. We used logistic regression with detections (1) and non-detections (0) as the response and maximum LC within 250 m of the survey location, ordinal date, and time of day (linear and quadratic) as predictors, with effort (hours) as an offset.

### Fine Scales

New Hampshire Fish and Game Department provided road-based spring drumming survey data for ruffed grouse. Observers counted drumming grouse from 10 stops along survey routes ( $14.6 \pm 2.2$  km;  $\bar{x} = 1.6$  km between stops). We obtained data from drumming surveys conducted by the Maine Department of Inland Fisheries and Wildlife. We downsampled Maine data from 15 to 10 stops/route, to make data comparable to New Hampshire data ( $11.5 \text{ km} \pm 0.6 \text{ km}$ , downsampled  $\bar{x} = 1.3$  km between stops). Maine, but not New Hampshire, surveys were replicated 3 times per year. Thus, we used the first complete survey from each Maine route. We calculated mean ruffed grouse LC within ecologically relevant fixed-width buffers of each route. We used 380-m and 550-m buffers corresponding to maximum distances at which grouse could be heard, depending on habitat (Hansen et al. 2011), and a 2-km buffer corresponding to maximum daily movement distances for non-dispersing individuals (Yoder et al. 2004). We dropped the 380-m buffer, which was highly correlated with the 550-m buffer. We tested if mean LC was a predictor of grouse occurrence along routes using logistic regression within a GLM framework.

We used point-level SGS data from a subset of SGS routes that had spatially referenced stops confirmed by USFWS to test the LC model's ability to predict American woodcock occurrence. We included stops only from years with confirmed

locations, because route locations sometimes changed between years. We calculated mean LC within fixed radii at each stop: 250 m (19.6 ha), corresponding to mean diurnal home range size (Sepik and Derleth 1993, Masse et al. 2014) and median distance (209.5 m) between courtship and nocturnal locations for singing males (Hudgins et al. 1985); 500 m for mean movement between successive diurnal to nocturnal points in adult woodcock (Sepik and Derleth 1993); and 1 km (Thogmartin et al. 2007), the range of distances between singing and nocturnal locations for singing males (Hudgins et al. 1985) and greatest monthly median distance between diurnal and nocturnal locations for adult males (Sepik and Derleth 1993). We tested if mean LC within each buffer was a predictor of woodcock occurrence with logistic regression using GLMM and survey route as a random intercept. We used model selection to select the best models explaining observed woodcock occurrence (Burnham and Anderson 2010) using the least corrected Akaike's Information Criterion ( $AIC_c$ ) value. Candidate models sets included combinations of LC within the optimal buffer radius, linear and quadratic ordinal date and year trends, and cloud cover in 25% increments. We considered competing models  $<2 \Delta AIC_c$  throughout all model selection analyses.

We evaluated log counts of grouse heard per survey route in New Hampshire and Maine to test the ability of LC models to explain fine-scale abundance. We used GLM including buffered mean LC variables as a fixed effect in 4 models (1 for each combination of year and buffer width), selecting buffer width using least corrected Akaike's Information Criterion ( $AIC_c$ ). We tested if LC explained count data independent of occurrence by rerunning models and excluding zero counts. We tested for spatial autocorrelation in model residuals with all data (zeros included) using Moran's  $I$  (Zuur et al. 2009).

We evaluated counts of woodcock detected per survey stop for each SGS route as a response metric to test the LC model's ability to predict fine-scale woodcock abundance. We used GLMM with a negative binomial distribution (Zuur et al. 2009). We built candidate model sets including route as a random intercept term. We also tested the ability of LC models to explain count data dependent and independent of occurrence by running count models both with and without zero counts. We additionally tested a hierarchical N-mixture model (Kéry and Royle 2016) to separately analyze Maine grouse data, because the repeated surveys allowed detection-corrected point abundance estimates at the survey-point level (Appendix A, available online in Supporting Information). We conducted all calculations in R, version 3.2 (R Core Team 2015) using packages lme4 (Bates et al. 2015), lmerTest (Kuznetsova et al. 2016), nlme (Pinheiro et al. 2017), AUC (Ballings and Van den Poel 2013), and ape (Paradis et al. 2004).

## RESULTS

Landscape capability was positively correlated with ruffed grouse detections per unit effort at the WMU-level in New York. The ruffed grouse LC model was verified, that is, significant with ranked data (slope  $\pm$  SE =  $0.57 \pm 0.09$ ,  $R^2 = 0.33$ ,  $F_{1,75} = 36.73$ ,  $P < 0.001$ ), and validated, significant with raw data (slope  $\pm$  SE =  $1.11 \pm 0.19$ ,  $R^2 = 0.31$ ,

$F_{1,75} = 33.97$ ,  $P < 0.001$ ). The LC covariate was the sole predictor of woodcock abundance and accounted for the majority of  $AIC_c$  weight (0.88; Table A1). No competing models were within  $2 \Delta AIC_c$ . The LC was positively correlated with observed log woodcock/stop (0.21, 95% CI = 0.16–0.26,  $R^2_c = 0.24$ ,  $R^2_m = 0.64$ ,  $F_{7,99} = 186.9$ ,  $P < 0.001$ ; Table A2).

Landscape capability was a significant positive predictor for regional woodcock and grouse occurrence ( $P < 0.001$ ,  $R^2_D = 0.31$  and  $0.32$ , respectively) using eBird data. The woodcock and grouse LC models had good occurrence discrimination (area under receiver operating characteristic curve [AUC] = 0.89 and 0.91, respectively), and the kappa statistic for each was greater than the null expectation ( $P < 0.01$ ; Table 1).

Ruffed grouse LC positively predicted route-level occurrence in New Hampshire and Maine in both years with the 550-m (2014 estimate:  $1.60 \pm 0.70$ ,  $Z = 2.31$ ,  $P = 0.021$ ; 2015 estimate:  $1.65 \pm 0.65$ ,  $Z = 2.54$ ,  $P = 0.011$ ) and 2-km buffers (2014 estimate:  $1.40 \pm 0.59$ ,  $Z = 2.35$ ,  $P = 0.019$ ; 2015 estimate:  $1.24 \pm 0.51$ ,  $Z = 2.45$ ,  $P = 0.014$ ). The 550-m buffer outperformed the 2-km buffer. Landscape capability models were accurate, precise, and sensitive (Table 1). For American woodcock, the 500-m LC buffer had better model fit than the 1-km ( $\Delta AIC_c = 2.84$ ) and 250-m buffers ( $\Delta AIC_c = 4.37$ ). Landscape capability within 500 m predicted woodcock occurrence with 71% accuracy ( $\bar{x} \pm SE = 1.80 \pm 0.56$ ,  $Z = 3.2$ ,  $P = 0.001$ ; Table 1). We observed a better negative predictive value (0.94) than positive predictive value (0.28) because of low prevalence (14%).

Landscape capability positively correlated with log ruffed grouse counts per survey in New Hampshire and Maine using the 550-m (2014 estimate:  $0.32 \pm 0.13$ ,  $R^2 = 0.14$ ,  $F_{1,41} = 6.75$ ,  $P = 0.01$ ; 2015:  $0.34 \pm 0.11$ ,  $R^2 = 0.17$ ,  $F_{1,47} = 10.8$ ,  $P = 0.003$ ) and 2-km buffers (2014:  $0.33 \pm 0.13$ ,  $R^2 = 0.15$ ,  $F_{1,41} = 6.94$ ,  $P = 0.012$ ; 2015:  $0.35 \pm 0.11$ ,  $R^2 = 0.17$ ,  $F_{1,47} = 9.68$ ,  $P = 0.003$ ). When we removed zero values (representing abundance independent of occupancy), LC did not correlate with count data using the 550-m (2014 estimate:  $0.07 \pm 0.11$ ,  $R^2 = 0.01$ ,  $F_{1,33} = 0.39$ ,  $P = 0.539$ ; 2015:  $0.07 \pm 0.09$ ,  $R^2 = 0.02$ ,  $F_{1,40} = 0.71$ ,  $P = 0.403$ ) and 2-km buffers (2014:  $0.08 \pm 0.11$ ,  $R^2 = 0.01$ ,  $F_{1,33} = 0.50$ ,  $P = 0.484$ ; 2015:  $0.11 \pm 0.09$ ,  $R^2 = 0.04$ ,  $F_{1,37} = 1.58$ ,  $P = 0.217$ ). Global models showed no spatial autocorrelation in residuals (Moran's  $I = 0.01$ – $0.03$ ). Using the grouse hierarchical model, LC was a marginal predictor of abundance at the stop level for repeated within-year surveys in Maine (2.47, 95% credible interval =  $-0.93$  to  $5.51$ ), with the posterior probability of the parameter  $> 0$  equal to 0.93. The model had adequate fit (Bayesian  $P = 0.74$ ). Detection probability decreased with date ( $\beta_{DATE} = -0.45$  [ $-0.71$  to  $-0.20$ ]), and time had a marginal effect on detection probability ( $\beta_{TIME} = -0.18$  [ $-0.4$  to  $0.02$ ]). Inclusion of LC compared to a null model including only the random effect reduced the variance in point estimates by 7% (partial  $R^2 = 0.07$ ).

Top American woodcock models included LC, date, and date<sup>2</sup>, and there was little variation in parameter estimates for LC. The 1-km buffer width was the radius with the best fit

(500-m buffer +3.78  $\Delta\text{AIC}_c$ , 250-m buffer +9.41  $\Delta\text{AIC}_c$ ). We found absolute model fit for all fixed effects in the top model was low ( $R^2_D = 0.05$ ). When we refit models to the data excluding counts with zero-detections and using a log-linear model of woodcock counts, we found no model including LC to have  $\Delta\text{AIC}_c < 2$ , and low LC parameter importance (0.28), suggesting LC did not correlate well with abundance beyond predicting presence-absence. We found no spatial autocorrelation in model residuals (Moran's  $I = 0.04$ ).

## DISCUSSION

Landscape capability explained observed breeding season woodcock abundance from SGS at regional scales, and observed occurrence at point scales, particularly from eBird data. Landscape capability was not strongly correlated with route-level woodcock relative abundance. It is unlikely that LC is sufficient to capture a large portion of variation within single-visit relative abundance data over such a broad (10 yr) survey window, given the reliance of woodcock on relatively ephemeral early successional vegetation, which is difficult to map remotely. Landscape capability models were better at ranking areas based on likelihood of occurrence, rather than by the number of animals expected in each occupied habitat patch. A significant complicating factor consistent with our results is that SGS may be a better measure of woodcock occurrence than abundance at a point scale, because subdominant adult males tend to remain at display sites rather than disperse, but they are non-vocal. As populations vary, there consistently is a singing male, but numbers of non-vocal subdominant males change, reducing heterogeneity in observed occurrence with increased heterogeneity in actual abundance (Dwyer et al. 1988, Sauer and Bortner 1991). At broad scales, LC reliably explained a significant portion of variation in log woodcock counts. Patterns of greater woodcock LC are qualitatively similar to woodcock relative abundance mapping developed from SGS data by Thogmartin et al. (2007). A regional comparison of woodcock LC to eBird data supports this finding. Greater LC values tend to more frequently occur moving northward (Fig. 1), with notable correspondence between LC and relative abundance in large portions of northern New York and New England. Intermediate LC values correspond to intermediate woodcock relative abundance in western Pennsylvania, portions of West Virginia, and eastern Virginia.

Predictive ability of the LC models for breeding season ruffed grouse qualitatively matched that of the American woodcock models, particularly with the regional assessment using eBird data. In route-level analysis, LC predicted grouse abundance only in so much as it predicted occupancy. The model was not a useful predictor of abundance after removal of zero counts in route-level models. Landscape capability was positively associated with abundance when noise was reduced through temporal or spatial aggregation, as in the New York data, or by accounting for unequal detection probability using repeated surveys in the N-mixture model. Still, this ability to predict abundance was a function of the model's predictive ability for occurrence or occupancy because most survey stops reported hearing 0, 1, or 2

grouse. Assessing measures of abundance requires some aggregation at least to the survey-route level, which begs the question of whether a regionally developed model is useful for predicting abundance at a local scale.

Predicting ruffed grouse occurrence is complicated by several factors. Grouse often are associated with aspen (*Populus tremuloides*; Svoboda and Gullion 1972, Hansen et al. 2011) in regions where it is prevalent, but they vary in their response to aspen based on nutrient content of individual aspen (Jakubas and Gullion 1991) and as a function of range (Stafford and Dimmick 1979, Jones et al. 2008). Grouse space use requirements vary at different stages of breeding, non-breeding, and brooding (Haulton et al. 2003), and habitat selection varies with demographic class (Tirpak et al. 2010), season, among years (Zimmerman et al. 2009), early successional vegetation availability, habitat interspersions, and hard mast availability (Blomberg et al. 2009). Given these caveats, model predictions compared with broader scale data likely perform better, as sources of heterogeneity average out in the aggregation process (Gotway and Young 2002).

Placement of SGS routes for woodcock and drumming survey routes for grouse are designed to be in areas that likely contain habitat (e.g., primarily forested areas, at least semi-natural environments). Predicting occurrence at sites that would almost certainly be unoccupied (likely an easy task for the model in areas with high-intensity urban development or open water) was precluded for both species, reducing the reported ability of the model to predict true-negatives. Thus, specificity, negative predictive value,  $\kappa$ , and accuracy are underestimated in assessing landscapes that include unsuitable areas for the species.

Landscape capability models are not designed as predictors of occurrence or abundance explicitly. Instead, LC predicts a landscape's capability to support home ranges of multiple individuals. Although LC models do not explicitly incorporate demographic information, reproductive success, or source-sink dynamics, they are intended to reflect the landscape contribution to these factors. Landscape capability would be expected, therefore, to be correlated with occupancy and abundance to an extent, as our analyses show. However, we urge caution when drawing inference about the strength of these correlations, and in not assuming LC translates into a spatial mapping of population viability.

We illustrated the relationship between landscape capability and occurrence and abundance metrics useful for informing management, monitoring, and conservation planning. However, in a strict sense the relevant metric of model performance as it was designed is the model's predictive sensitivity (or recall). False positive predictions of occurrence are not by default false predictions of potential suitability, as sites may be unoccupied for any number of reasons including dispersal limitation or historical contingency (Pearson et al. 2007). Unfortunately, differentiating the 2 is not possible in practice. Although the paradigm has shifted away from seeking a universal validation of correlative species models for both practical and philosophical reasons (Van Horne 2002, Jarnevich et al. 2015), the sensitivity demonstrated in conjunction with the use of LC models as a predictor of

occurrence suggests the models perform well for their designed purpose, but more importantly they are useful.

Species-habitat model verification provides an incomplete picture of a model's utility as a predictor of bird occurrence. Sensitivity to spatial modeling methods can lead to low conformity in identifying key areas for conservation (Jones-Farrand et al. 2011). Therefore, when it is impractical or cost prohibitive to develop multiple models for ensemble prediction, the best strategy may be a know-thy-model approach, whereby explicit model performance at specified temporal and spatial scales is reported (Jones-Farrand et al. 2011). Despite explicit spatial and temporal granularity of model outputs, scales at which models perform best are not implicit in outputs, and careful consideration of scale is required (Will et al. 2005). Many model assumptions (e.g., system equilibrium, adequate sampling across all relevant ecological gradients) are essentially untestable (Jarnevich et al. 2015). These caveats preclude easy validation of spatial models for predictions of distribution independent of intended model uses.

## MANAGEMENT IMPLICATIONS

Landscape capability models were predictive of upland game bird occurrences; however, they were not useful in predicting abundance. Therefore, LC models may be surrogates of species distribution, but they are unlikely to inform us about population densities. Because of their design, LC performed better at identifying habitat where animals are likely to occur (i.e., potential habitat) rather than predicting non-occurrence. There continues to be a need for high-quality, broad-extent data of occurrence and abundance for upland game birds in conservation planning, especially for developing regional models.

## ACKNOWLEDGMENTS

Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government. We thank K. McGarigal, E. B. Plunkett, B. W. Compton, and J. Grand for developing and providing DSL models, and M. V. Schiavone, B. Allen, K. Sullivan, K. A. Boredeau, and R. D. Rau for providing or assisting with validation data. Data were provided by the Designing Sustainable Landscapes Project at the University of Massachusetts Amherst-Landscape Ecology Lab, NYS-DEC, Maine Department of Inland Fisheries and Wildlife, New Hampshire Fish and Game Department, Cornell Lab of Ornithology via eBird, and the USFWS Migratory Bird Program. The NYSDEC Ruffed Grouse Drumming Survey data was supported by Federal Aid in Wildlife Restoration Grant W-173-G. The datasets are available by request to the owners. Other funding was provided by the U.S. Geological Survey-Science Support Program and the University of Maine.

## LITERATURE CITED

Backs, S. E., and J. S. Castrale. 2010. The distribution and conservation status of ruffed grouse in Indiana: 25 years of decline. *Proceedings of the Indiana Academy of Science* 119:101–104.

- Ballings, M., and D. Van den Poel. 2013. AUC: threshold independent performance measures for probabilistic classifiers. R package version 0.3.0. <https://CRAN.R-project.org/package=AUC>. Accessed 10 Feb 2017.
- Bates, D., M. Maechler, B. Bolker, and S. Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67:1–48.
- Blomberg, E. J., B. C. Tefft, E. G. Endrulat, and S. R. McWilliams. 2009. Predicting landscape-scale habitat distribution for ruffed grouse *Bonasa umbellus* using presence-only data. *Wildlife Biology* 15:380–394.
- Bonnot, T. W., F. R. Thompson III, and J. J. Millsaugh. 2011. Extension of landscape-based population viability models to ecoregional scales for conservation planning. *Biological Conservation* 144:2041–2053.
- Brooks, R. P. 1997. Improving habitat suitability index models. *Wildlife Society Bulletin* 25:163–167.
- Burnham, K. P., and D. R. Anderson. 2010. *Model selection and multimodel inference: a practical information-theoretic approach*. Second edition. Springer, New York, New York, USA.
- Cooper, T. R., and R. D. Rau. 2014. American woodcock population status, 2014. U.S. Fish and Wildlife Service, Laurel, Maryland, USA.
- Dessecker, D. R., and D. G. McAuley. 2001. Importance of early successional habitat to ruffed grouse and American woodcock. *Wildlife Society Bulletin* 29:456–465.
- Dettmers, R. 2003. Status and conservation of shrubland birds in the northeastern United States. *Forest Ecology and Management* 185:81–93.
- Dwyer, T. J., G. F. Sepik, E. L. Derleth, and D. G. McAuley. 1988. Demographic characteristics of a Maine woodcock population and effects of habitat management. U.S. Fish and Wildlife Service, Fish and Wildlife Research 4, Washington, D.C., USA.
- Fink, D., W. M. Hochachka, B. Zuckerberg, D. W. Winkler, B. Shaby, M. A. Munson, G. Hooker, M. Riedewald, D. Sheldon, and S. Kelling. 2010. Spatiotemporal exploratory models for broad-scale survey data. *Ecological Applications* 20:2131–2147.
- Fuller, T. K., and S. DeStefano. 2003. Relative importance of early successional forests and shrubland habitats to mammals in the northeastern United States. *Forest Ecology and Management* 185:75–79.
- Gawler, S. C. 2008. Northeastern Terrestrial Wildlife Habitat Classification. Report to the Virginia Department of Game and Inland Fisheries on behalf of the Northeast Association of Fish and Wildlife Agencies and the National Fish and Wildlife Foundation. NatureServe, Boston, Massachusetts, USA.
- Gotway, C. A., and L. J. Young. 2002. Combining incompatible spatial data. *Journal of the American Statistical Association* 97:632–648.
- Hansen, C. P., J. J. Millsaugh, and M. A. Rumble. 2011. Occupancy modeling of ruffed grouse in the Black Hills National Forest. *Journal of Wildlife Management* 75:71–77.
- Haulton, G. S., D. F. Stauffer, R. L. Kirkpatrick, and G. W. Norman. 2003. Ruffed grouse (*Bonasa umbellus*) brood microhabitat selection in the southern Appalachians. *American Midland Naturalist* 150:95–103.
- Hudgins, J. E., G. L. Storm, and J. S. Wakeley. 1985. Local movements and diurnal-habitat selection by male American woodcock in Pennsylvania. *Journal of Wildlife Management* 49:614–619.
- Jakubas, W. J., and G. W. Gullion. 1991. Use of quaking aspen flower buds by ruffed grouse: its relationship to grouse densities and bud chemical composition. *Condor* 93:473–485.
- Jarnevich, C. S., T. H. Stohlgren, S. Kumar, J. T. Morissette, and T. R. Holcombe. 2015. Caveats for correlative species distribution modeling. *Ecological Informatics* 29:6–15.
- Jones, B. C., J. L. Kleitch, C. A. Harper, and D. A. Buehler. 2008. Ruffed grouse habitat use in a mixed hardwood forest: implications for forest management in the Appalachians. *Forest Ecology and Management* 255:3580–3588.
- Jones-Farrand, D. T., T. M. Fearer, W. E. Thogmartin, F. R. Thompson III, M. D. Nelson, and J. M. Tirpak. 2011. Comparison of statistical and theoretical habitat models for conservation planning: the benefit of ensemble prediction. *Ecological Applications* 21:2269–2282.
- Kéry, M., and J. A. Royle. 2016. *Applied hierarchical modeling in ecology*. Academic Press, New York, New York, USA.
- Knoche, S., and F. Lupi. 2013. Economic benefits of publicly accessible land for ruffed grouse hunters. *Journal of Wildlife Management* 77: 1294–1300.
- Kuznetsova, A., P. B. Brockhoff, and R. H. B. Christensen. 2016. lmerTest: tests in linear mixed effects models. R package version 2.0–32. <https://CRAN.R-project.org/package=lmerTest>. Accessed 10 Feb 2017.



- Masse, R. J., B. C. Tefft, and S. R. Williams. 2014. Multiscale habitat selection by a forest-dwelling shorebird, the American woodcock: implications for forest management in southern New England, USA. *Forest Ecology and Management* 325:37–48.
- McAuley, D., D. M. Keppie, and R. M. Whiting Jr. 2013. American woodcock (*Scolopax minor*). No. 100 in A. Poole, editor. *The birds of North America*. Cornell Lab of Ornithology, Ithaca, New York, USA.
- McGarigal, K. W. V. Deluca, B. W. Compton, E. B. Plunkett, and J. Grand. 2016. Designing Sustainable Landscapes: modeling focal species. Report to the North Atlantic Landscape Conservation Cooperative, U.S. Fish and Wildlife Service, Northeast Region. [http://jamba.provost.ads.umass.edu/web/lcc/DSSL\\_documentation\\_species.pdf](http://jamba.provost.ads.umass.edu/web/lcc/DSSL_documentation_species.pdf). Accessed 4 Feb 2016.
- Mitchell, M. S., R. A. Lancia, and J. A. Gerwin. 2001. Using landscape-level data to predict the distribution of birds on a managed forest: effects of scale. *Ecological Applications* 11:1692–1708.
- New York State Department of Environmental Conservation [NYSDEC]. 2015. Ruffed Grouse Drumming Survey. New York State Department of Environmental Conservation, Albany, New York, USA. [http://www.dec.ny.gov/docs/wildlife\\_pdf/grousemdrump15.pdf](http://www.dec.ny.gov/docs/wildlife_pdf/grousemdrump15.pdf). Accessed 24 Nov 2015.
- New York State Department of Environmental Conservation [NYSDEC]. 2016. Wildlife Management Units. New York State Department of Environmental Conservation, Albany, New York, USA. <http://www.dec.ny.gov/outdoor/8302.html>. Accessed 3 Feb 2016.
- National Oceanic and Atmospheric Administration. 2017. National Centers for Environmental Information, Climate at a Glance: U.S. Time Series, Maximum Temperature. National Oceanic and Atmospheric Administration, National Centers for Environmental Information, Asheville, North Carolina, USA. <http://www.ncdc.noaa.gov/cag/>. Accessed on 16 Mar 2017.
- Paradis E., J. Claude, and K. Strimmer. 2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20:289–290.
- Pearson, R. G., C. J. Raxworthy, M. Nakamura, and A. T. Peterson. 2007. Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *Journal of Biogeography* 34:102–117.
- Pinheiro J, D. Bates, S. DebRoy, D. Sarkar and R Core Team. 2017. nlme: linear and nonlinear mixed effects models. R package version 3.1-130. <https://CRAN.R-project.org/package=nlme>. Accessed 10 Feb 2017.
- R Core Team. 2015. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Rusch, D. H., S. Destefano, M. C. Reynolds, and D. Lauten. 2000. Ruffed grouse (*Bonasa umbellus*). No. 515 in A. Poole, editor. *The birds of North America*. Cornell Lab of Ornithology, Ithaca, New York, USA.
- Sauer, J. R., and J. B. Bortner. 1991. Population trends from the American woodcock Singing-Ground Survey, 1970–88. *Journal of Wildlife Management* 55:300–312.
- Seamans, M. E., and R. D. Rau. 2016. American woodcock population status, 2016. U.S. Fish and Wildlife Service, Laurel, Maryland, USA.
- Sepik, G. F., and E. L. Derleth. 1993. Habitat use, home range size, and patterns of moves of the American woodcock in Maine. Pages 41–49 in *Proceedings of the Eighth American Woodcock Symposium*. Biological Report 16. U.S. Fish and Wildlife Service, 29 Oct–2 Nov 1990, West Lafayette, Indiana, USA.
- Stafford, S. K., and R. W. Dimmick. 1979. Autumn and winter foods of ruffed grouse in the southern Appalachians. *Journal of Wildlife Management* 43:121–127.
- Sullivan, B. L., C. L. Wood, M. J. Iliff, R. E. Bonney, D. Fink, and S. Kelling. 2009. eBird: a citizen-based bird observation network in the biological sciences. *Biological Conservation* 142:2282–2292.
- Svoboda, F. J., and G. W. Gullion. 1972. Preferential use of aspen by ruffed grouse in northern Minnesota. *Journal of Wildlife Management* 49:605–610.
- Thogmartin, W. E., J. R. Sauer, and M. G. Knutson. 2007. Modeling and mapping abundance of American woodcock across the midwestern and northeastern United States. *Journal of Wildlife Management* 71:376–382.
- Tirpak, J. M., W. M. Giuliano, T. J. Allen, S. Bittner, J. W. Edwards, S. Friedhof, C. A. Harper, W. K. Igo, D. F. Stauffer, and G. W. Norman. 2010. Ruffed grouse-habitat preference in the central and southern Appalachians. *Forest Ecology and Management* 260:1525–1538.
- U.S. Census Bureau. 2012. 2010 Census of population and housing, population and housing unit counts, CPH-2-1, United States Summary: 2010. U.S. Government Printing Office, Washington, D.C., USA.
- Van Horne, B. 2002. Approaches to habitat modeling: the tensions between pattern and process and between specificity and generality. Pages 63–72 in J. M. Scott, P. Helgund, M. L. Morrison, J. B. Haufler, M. G. Raphael, W. A. Wall, and F. B. Samson, editors. *Predicting species occurrences: issues of accuracy and scale*. Island Press, Washington, D.C., USA.
- Will, T. C., J. M. Ruth, K. V. Rosenberg, D. Krueper, D. Hahn, J. Fitzgerald, R. Dettmers, and C. J. Beardmore. 2005. The five elements process: designing optimal landscapes to meet bird conservation objections. *Partners in Flight Technical Series No. 1*, Laurel, Maryland, USA. <https://www.fort.usgs.gov/publication/21581>. Accessed 8 Feb 2016.
- Yoder, J. M., E. A. Marschall, and D. A. Swanson. 2004. The cost of dispersal: predation as a function of movement and site familiarity in ruffed grouse. *Behavioral Ecology* 15:469–476.
- Zimmerman, G. S., R. J. Gutiérrez, W. E. Thogmartin, and S. Banerjee. 2009. Multiscale habitat selection by ruffed grouse at low population densities. *Condor* 111:294–304.
- Zuur, A., E. N. Ieno, N. Walker, A. A. Saveliev, and G. M. Smith. 2009. *Mixed effects models and extensions in ecology with R*. Springer, New York, New York, USA.

Associate Editor: Wayne Thogmartin.

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's website.